

# nature



## PROTEIN POWER

AI network predicts highly accurate 3D structures for the human proteome

**Troubled waters**  
The race to save the Great Barrier Reef from climate change

**Coronavirus**  
Time is running out to find the origins of SARS-CoV-2

**Storage hunting**  
Quantifying carbon held in Africa's montane forests

# Nature.2021.08.30

[Mon, 30 Aug 2021]

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

# This Week

- [\*\*Control methane to slow global warming — fast\*\*](#) [ 25 August 2021]  
Editorial • Carbon dioxide reductions are key, but the IPCC's latest report highlights the benefits of making cuts to other greenhouse gases, too.
- [\*\*Too many scientists still say Caucasian\*\*](#) [ 24 August 2021]  
World View • Racist ideas of categories for human identity continue to warp research and medicine.
- [\*\*Can Chinese diners embrace potatoes? The answer could affect Earth's climate\*\*](#) [ 18 August 2021]  
Research Highlight • A switch to spuds in a country where rice is prominent could cut greenhouse-gas emissions.
- [\*\*The animal that doesn't forget has tentacles, not a trunk\*\*](#) [ 17 August 2021]  
Research Highlight • Cuttlefish, an octopus relative, retain their cognitive powers as they grow older.
- [\*\*A tough prosthetic hand obeys the muscles' commands\*\*](#) [ 18 August 2021]  
Research Highlight • The inexpensive and lightweight prosthesis also provides feedback akin to a sense of touch.
- [\*\*What's the mystery source of two potent greenhouse gases? The trail leads to Asia\*\*](#) [ 18 August 2021]  
Research Highlight • Atmospheric levels of two powerful heat-trapping gases are rising quickly — and are higher than official emissions records suggest.
- [\*\*A staggering store of water is revealed in Earth's crust\*\*](#) [ 17 August 2021]  
Research Highlight • Modelling work shows that crustal groundwater accounts for more water than the world's ice caps and glaciers.
- [\*\*Are 20 seconds of handwashing really necessary? Physics says yes\*\*](#) [ 20 August 2021]  
Research Highlight • A simple model suggests that there's no fast way to rid hands of virus-sized particles.

- **Clever orangutans invent nutcrackers from scratch** [ 18

August 2021]

Research Highlight • Chimpanzees are not the only great apes to develop tools without tutoring.

---

| [Next section](#) | [Main menu](#) |

- EDITORIAL
- 25 August 2021

# Control methane to slow global warming — fast

Carbon dioxide reductions are key, but the IPCC's latest report highlights the benefits of making cuts to other greenhouse gases, too.



Oil and gas operations — such as Inglewood Oil Field in Los Angeles, California — are a key source of methane. Credit: Citizens of the Planet/Education Images/Universal Images Group/Getty

There's no substitute for eliminating fossil fuels and halting the release of carbon dioxide into the atmosphere to avoid the painful and disruptive effects of global warming. The latest report from the Intergovernmental

Panel on Climate Change (IPCC) leaves no doubt about this. But CO<sub>2</sub> is not the only greenhouse gas. The climate panel also highlights the problem — and opportunity — posed by methane, which has contributed as much as 0.5 °C of warming since pre-industrial times, second only to CO<sub>2</sub>.

Methane is the main component of natural gas, whose popularity as a relatively clean source of fossil energy has soared by more than 50% over the past two decades. In the United States, cheap and plentiful natural gas supplies — obtained from controversial hydraulic fracking — have helped to displace coal. But there has also been a cost to the climate: fossil fuels have helped to boost atmospheric methane concentrations, which have more than doubled since pre-industrial times, from around 700 parts per billion by volume to nearly 1,900 p.p.b. in 2020.

Methane is of concern because it has an outsized impact on the climate. The gas makes up a tiny fraction of our atmosphere — CO<sub>2</sub> levels are more than 200 times higher. But in the first 20 years after release, methane is around 80 times more powerful than CO<sub>2</sub> at trapping heat in Earth's atmosphere. It also breaks down much more quickly than CO<sub>2</sub>, with an average lifetime of around a decade, compared with centuries for CO<sub>2</sub>. This means that curbing methane emissions could provide short-term relief while governments and businesses negotiate the more difficult transition from fossil fuels to clean energy.



## IPCC climate report: Earth is warmer than it's been in 125,000 years

In an effort to reduce methane emissions, scientists have been investigating two linked questions. First, what are the major sources of methane? Second, where are the worst offenders? Livestock is the largest source, responsible for 31% of the global total, according to Ilissa Ocko at the non-profit Environmental Defense Fund (EDF) in New York City and her colleagues<sup>1</sup>. Oil and gas operations rank a close second, releasing 26%. Other sources include landfills, coal mines, rice paddies and water-treatment plants.

Reducing methane from livestock represents a particular challenge. People could eat less meat, but persuading people to change their diet is rarely straightforward. Moreover, meat consumption is rising in low- and middle-income countries — in line with rising incomes. It should be easier to curb emissions from other sectors. In many cases, doing so wouldn't cost anything — and it could even be profitable.

Global methane emissions could be cut by 57% by 2030 using existing technologies, Ocko and her colleagues report. And almost one-quarter of the global methane total could be eliminated at no net cost. The oil and gas industry could make the biggest difference here, having both the infrastructure and the incentive to minimize methane losses: more methane in their pipelines means more revenue. In other sectors, the operators of landfills, coal mines and wastewater-treatment plants could capture the gas and use it to generate electricity. And rice producers could minimize emissions with better irrigation and soil-management practices. If these measures were implemented worldwide, projected increases in global warming could be reduced by 0.25°C by 2050, and 0.5°C by 2100, the study finds. Those are significant numbers considering that the world has already warmed by 1.1°C, and global leaders have committed to limiting the total to 1.5–2°C.



### The hard truths of climate change — by the numbers

The main problem is pinpointing precisely where methane emissions are coming from. More than a decade ago, researchers monitored methane from aircraft surveys using infrared sensors that can detect gases in sunlight reflected back from Earth<sup>2</sup>. Today, satellites are part of this monitoring effort. Research shows that a relatively small number of ‘super-emitters’ are responsible for a significant share of methane emissions, particularly in the oil and gas industry<sup>3</sup>.

The ability to identify major methane sources is poised to advance in the next two years. In 2022, the EDF will launch a satellite designed to identify emissions across large swathes of land. Carbon Mapper, a non-profit partnership that includes NASA’s Jet Propulsion Laboratory and the San Francisco-based company Planet, will follow up in 2023 with two prototype satellites designed to track methane and CO<sub>2</sub> at the scale of individual facilities.

In March, the United Nations Environment Programme and the European Commission launched the International Methane Emissions Observatory to help coordinate these efforts and help policymakers and companies take action. The observatory will also have access to estimated emissions inventories from governments and industry. Around 70 oil and gas producers, including giants such as Shell and BP, have committed to setting

clear emissions-reductions targets and reporting emissions under an initiative led by the Climate & Clean Air Coalition, an international initiative involving governments, non-profit organizations, businesses and others. This work will also help to inform new methane-reduction commitments due to be made at the UN climate conference in Glasgow, UK, in November.

The world will continue to warm as long as CO<sub>2</sub> is being pumped into the atmosphere. But curbing emissions of methane and other powerful greenhouse gases might reduce the sting. That is why governments and businesses should seize the opportunity, buying humanity a bit more time to do what needs to be done.

*Nature* **596**, 461 (2021)

doi: <https://doi.org/10.1038/d41586-021-02287-y>

## References

1. 1.

Ocko, I. B. *et al.* *Environ. Res. Lett.* **16**, 054042 (2021).

2. 2.

Pétron, G. *et al.* *J. Geophys. Res.* **117**, D04304 (2012).

3. 3.

Cusworth, D. H. *et al.* *Environ. Sci. Technol. Lett.* **8**, 567–573 (2021).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02287-y>

- WORLD VIEW
- 24 August 2021

## Too many scientists still say Caucasian



Racist ideas of categories for human identity continue to warp research and medicine.

- [Alice B. Popejoy](#) 0

Of the ten clinical genetics labs in the United States that share the most data with the research community, seven include ‘Caucasian’ as a multiple-choice category for patients’ racial or ethnic identity, despite the term having no scientific basis. Nearly 5,000 biomedical papers since 2010 have used ‘Caucasian’ to describe European populations. This suggests that too many scientists apply the term, either unbothered by or unaware of its roots in racist taxonomies used to justify slavery — or worse, adding to pseudoscientific claims of white biological superiority.

I work at the intersection of statistics, evolutionary genomics and bioethics. Since 2017, I have co-led a diverse, multidisciplinary working group funded by the US National Institutes of Health to investigate diversity measures in clinical genetics and genomics ([go.nature.com/3su2t8n](https://go.nature.com/3su2t8n)).

Many working in genomics do have a nuanced understanding of the issues and want to get things right. Still, I have been dismayed by how often the academics and clinicians I've encountered shy away from examining, or even acknowledging, how racism warps science. Decades of analyses have shown that 'racial groups' are defined by societies, not by genetics. Only the privileged have the luxury of opining that this is not a problem. As a white woman, I too have blind spots that need constant examination.



### [Don't ignore genetic data from minority populations](#)

Pioneering works in social science such as Dorothy Roberts' *Fatal Invention* (2012), Kim Tallbear's *Native American DNA* (2013) and *The Social Life of DNA* (2016) by Alondra Nelson, have eloquently pointed out many of the flawed assumptions and approaches that plague human genomics.

A common theme of this scholarship is that groupings depend more on dominant culture than on ancestry. In Singapore, the government mandates that individuals are identified explicitly as Chinese, Malay, Indian or Other,

which affects where they can live and study. In the United States, people with ancestry from the world's two most populous countries, India and China, along with every other country on the continent, are collapsed into a single racial category called 'Asian'. Similarly, the term 'Hispanic' erases a multitude of cultural and ancestral identities, especially among Indigenous peoples of the Americas.

Erroneous ideas about genetic 'races' live on in the broad, ambiguous 'continental ancestry' groups such as 'Black, African' or 'African American', that are used in the US Census and are ubiquitous in biomedical research. These collapse incredible amounts of diversity and erase cultural and ancestral identities. Study participants deemed not to fit within such crude buckets are often excluded from analyses, despite the fact that fewer and fewer individuals [identify with a single population of origin](#).

One practical way forwards is to move away from having people identify themselves using only checkboxes. I am not calling for an end to the study of genetic ancestry or socio-cultural categories such as self-identified race and ethnicity. These are useful for tracking and studying equity in justice, health care, education and more. The goal is to stop conflating the two, which leads scientists and clinicians to attribute differences in health to innate biology rather than to [poverty and social inequality](#).



[Facing up to injustice in genome science](#)

We need to acknowledge that systemic racism, not genetics, is dominant in creating health disparities. It shouldn't have taken the inequitable ravages of a pandemic to highlight that. Furthermore, every researcher and physician should be aware of the racial bias that abounds in medical practice: some pulse oximeters give more accurate readings for light-skinned people than for those with dark skin; Black Americans are undertreated for pain; and historical biases in data used to train algorithms to make medical decisions can lead to worse outcomes for vulnerable groups. Hence the ongoing revisions to the subsection on race and ethnicity in the American Medical Association's *Manual of Style*, and why medical schools are examining how their curricula reinforce harmful misconceptions about race.

Thankfully, more researchers are collecting self-reported data on geographical family origins, languages spoken at home and cultural affiliations. I'd like to see data-collection forms with open-ended questions, rather than those that force fixed choices or reduce identity to a box labelled 'other'. These self-reported indicators could be combined with genetic data to improve on current approaches to mapping the dimensions of diversity in our populations.

Approaches to genetic ancestry based on known reference populations are inadequate, in part because so much global diversity is missing from our data. I am working with the Human Pan-genome Reference Consortium, which aims to generate a more accurate and inclusive resource for global genomic diversity. It will include communities, especially Indigenous peoples, in developing protocols for data collection, storage and use. This respects Indigenous data sovereignty, and makes for more accurate and inclusive studies.

The more precisely we can measure genetic and non-genetic contributors to health and disease, the less researchers will rely on biologically meaningless designations that reinforce faulty assumptions and cause harm. The use of sequence data in clinical care could, for instance, facilitate recommendations for drug dosage that are genotype-based, rather than race-based.

Simply picking another word to replace 'Caucasian' won't be enough to root out racism in research and medicine. But all should be aware of the

harms the word represents.

*Nature* **596**, 463 (2021)

*doi:* <https://doi.org/10.1038/d41586-021-02288-x>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02288-x>

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



Rice planting in China releases the potent greenhouse gas methane. Those emissions contribute to making rice grown in China a generally less climate-friendly crop per calorie than potato. Credit: Costfoto/Barcroft Media/Getty

Sustainability

18 August 2021

# Can Chinese diners embrace potatoes? The answer could affect Earth's climate

A switch to spuds in a country where rice is prominent could cut greenhouse-gas emissions.

China's ambitious plan to make the potato a staple crop could slash greenhouse-gas emissions — if the country's diners sacrifice some of their rice consumption.

Rice cultivation requires heavy use of water and fertilizer, and gradually degrades the soil. Seeking a sustainable and nutritious way to feed China's ballooning population, the government implemented a policy in 2015 to double potato yields and add potatoes to the Chinese diet.

Jun Bi at Nanjing University in China and his colleagues examined the policy's environmental implications and found that potatoes grown in China have lower greenhouse-gas emissions and water demand per calorie than do other crops, such as maize. Extrapolating from historical data, the researchers estimate that, from 2015 to 2030, greenhouse-gas emissions from Chinese farming of staple crops could fall nearly 9% if the government meets its potato goals and plants the crop in the most suitable places.

But if potato fields displace rice paddies and the population doesn't shift to a more potato-heavy diet, rice imports could increase. That would mean higher environmental costs because of transportation emissions, the authors warn.

*Nature Food* (2021).

- [Sustainability](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02230-1>



Age does not wither the mental powers of the common cuttlefish. Credit: Reinhard Dirscherl/Alamy

Neuroscience

17 August 2021

## The animal that doesn't forget has tentacles, not a trunk

Cuttlefish, an octopus relative, retain their cognitive powers as they grow older.

Even in old age, cuttlefish remain as mentally sharp as ever.

Ageing mammals — especially humans — gradually lose their ability to remember specific events. To find out whether the same is true of molluscs, an invertebrate group that includes clams, snails and octopuses, Alexandra Schnell at Cambridge University, UK, and her colleagues taught six older common cuttlefish (*Sepia officinalis*) that a seafood snack in their tanks changed location depending on the time of day. The old cuttlefish learnt to associate the time and location just as well as six young cuttlefish did.

Next, the researchers tested how well the molluscs could recall a specific memory about where and when they had eaten a specific snack for breakfast and use that memory to search for the food. Older cuttlefish were slightly better than young animals at this task, leading the researchers to conclude that they are the first known animal to not undergo mental decline with age.

It's still unclear which unique features of the cuttlefish brain anatomy and physiology lead to this phenomenon.

[Proc. R. Soc. B \(2021\)](#)

- [Neuroscience](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02240-z>



This soft prosthetic hand allows the wearer to grasp delicate objects such as cakes and berries. Credit: G. Gu *et al./Nature Biomed. Eng.*

Engineering

18 August 2021

## A tough prosthetic hand obeys the muscles' commands

The inexpensive and lightweight prosthesis also provides feedback akin to a sense of touch.

A lightweight prosthetic hand can grasp strawberries without crushing them and still works after being hit with a hammer or run over by a 1.5-tonne

vehicle.

Xiangyang Zhu at Shanghai Jiao Tong University in China and his colleagues developed the resilient prosthesis for people with amputations below the elbow. The device boasts a 3D-printed palm and five jointed digits containing silicone tubes that are moved with pressurized air. Electrical signals from forearm muscles control the hand, which has five configurations.

Other prosthetic hands controlled through nervous-system signals can cost more than US\$10,000. The components used in the team's version cost less than \$500. The hand is also considerably more lightweight than its competitors, thanks to a bag worn at the waist that contains the battery and the electronics used to turn the arm's signals into movement instructions. The bag also carries the pumps and valves used to supply air.

Study participants needed just 15 minutes to learn how to control the prosthesis, which provides feedback using signals sent from pressure-sensitive sensors on the fingertips to electrodes on the upper arm.

*Nature Biomed. Eng. (2021)*

- [Engineering](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02238-7>



An aluminium smelter in Zouping, China. Two greenhouse gases whose atmospheric levels have soared in recent years have been traced to such smelters and to semiconductor factories in Japan and South Korea. Credit: Brent Lewin/Bloomberg/Getty

Atmospheric science

18 August 2021

# What's the mystery source of two potent greenhouse gases? The trail leads to Asia

Atmospheric levels of two powerful heat-trapping gases are rising quickly — and are higher than official emissions records suggest.

The powerful greenhouse gases tetrafluoromethane and hexafluoroethane have been building up in the atmosphere from unknown sources. Now, modelling suggests that China's aluminium industry is a major culprit.

The gases are thousands of times more effective than carbon dioxide at warming the atmosphere. Official tallies of tetrafluoromethane and hexafluoroethane emissions from factories are too low to account for the levels in the air, which began to rise in 2015 after seven years of relative stability.

Seeking to pinpoint the sources of those emissions, Jooil Kim at the University of California, San Diego, and his colleagues analysed air samples collected roughly every 2 hours between November 2007 and December 2019 on South Korea's Jeju Island. The scientists also modelled the weather patterns that transported air across the island during that period, to track the gases' origins.

The results suggest that aluminium smelters in China account for a large proportion of these chemicals in the atmosphere. Semiconductor factories in South Korea and Japan are probably also to blame.

*[JGR Atmos. \(2021\)](#)*

- [Atmospheric science](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02231-0>



People gather at a pump in India to collect groundwater. Accessible, fresh water makes up only a fraction of the water in Earth's crust. Credit: Jack Laurenson/Lnp/Shutterstock

Water resources

17 August 2021

## A staggering store of water is revealed in Earth's crust

Modelling work shows that crustal groundwater accounts for more water than the world's ice caps and glaciers.

The depths of Earth's crust hold a huge volume of ancient, salty water that has been undetected until now.

Grant Ferguson at the University of Saskatchewan in Saskatoon, Canada, and his colleagues calculated how much of this underground water should exist. They analysed a global database of the types of rock that make up the uppermost 10 kilometres of the planet's continental crust. Nearly 88% is hard crystalline rock, and 12% is sedimentary rock, which has large spaces between its grains.

The scientists calculated how much water could exist between the grains of both of these rock types, and estimated that the uppermost 10 kilometres of Earth's crust holds nearly 44 million cubic kilometres of water. That's more than the amount frozen in glaciers and the ice sheets of Greenland and Antarctica.

Most of this vast reservoir lies at a depth of between 1 kilometre and 10 kilometres, beyond the reach of wells that could tap it. The groundwater used by many farmers for irrigation and by billions of people for drinking is at much shallower depths.

*[Geophys. Res. Lett. \(2021\)](#)*

- [Water resources](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02232-z>



No shortcuts, please: modelling supports the public-health directives to wash hands for 20 seconds. Credit: David Cliff/NurPhoto/Getty

Fluid dynamics

20 August 2021

## Are 20 seconds of handwashing really necessary? Physics says yes

A simple model suggests that there's no fast way to rid hands of virus-sized particles.

Handwashing must be done at a minimum speed to dislodge viruses and bacteria, and it should last roughly 20 seconds, on par with the time that

public-health experts recommend, according to an analysis of the fluid dynamics of soaping up.

The simple act of handwashing masks some complex physics. Two rough surfaces — hands — slide past one another, separated by a thin layer of water and soap.

To illuminate the physical details, Paul Hammond, a consultant based in Bourn, UK, turned to a 135-year-old branch of fluid dynamics called lubrication theory, which excels at describing the physics of thin layers of fluids wedged between surfaces. Hammond used its formulas to devise a simple model that could be used to estimate how long it takes to dislodge any virus-sized particles.

The results confirmed that it does indeed require about 20 seconds of hand-rubbing to knock off pathogens. Although the analysis did not take the chemistry and biology of handwashing into account, the author says the results are an encouraging starting point for further study.

*[Phys. Fluids \(2021\)](#)*

- [Fluid dynamics](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02241-y>.



A Sumatran orangutan named Padana raps nuts with a log ‘hammer’ atop a tree stump ‘anvil’, a technique she invented herself. Credit: Claudio Tennie

Animal behaviour

18 August 2021

## Clever orangutans invent nutcrackers from scratch

Chimpanzees are not the only great apes to develop tools without tutoring.

Chimpanzees have long been thought to be the only non-human great apes to regularly smash nuts with stones or wooden hammers — one of the most complex forms of tool use observed in nature. Now, researchers have found

that orangutans, too, can use hammers to crack open nuts, and they learn to do so without copying others.

Elisa Bandini at the University of Tübingen in Germany and her colleagues observed 12 zoo-dwelling orangutans (*Pongo abelii* and *Pongo pygmaeus*) that were given hard nuts and small wooden logs as potential hammers. None of the animals had previously broken nuts open with tools.

Some of the apes never tried the potential tools. But most wielded the logs as hammers to crack open the nuts, and three used a tree stump or another object as an anvil to stabilize the nuts.

Four animals started to use the tools without observing more experienced individuals, which suggests that orangutans can spontaneously learn to use objects as efficient nutcrackers, the researchers say.

[Am. J. Primatol. \(2021\)](#)

- [Animal behaviour](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02239-6>

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

# News in Focus

- **[Babbling bats, autocorrected genes and COVID antibodies](#)**

[ 25 August 2021]

News Round-Up • The latest science news, in brief.

- **[African languages to get more bespoke scientific terms](#)** [ 18

August 2021]

News • Many words common to science have never been written in African languages. Now, researchers from across Africa are changing that.

- **[Decades-old SARS virus infection triggers potent response to COVID vaccines](#)** [ 18 August 2021]

News • Dramatic antibody production in people infected during the 2002–04 outbreak furthers hopes of a vaccine against many coronaviruses.

- **[The mutation that helps Delta spread like wildfire](#)** [ 20

August 2021]

News • A key amino-acid change might underlie the coronavirus variant's ferocious infectivity.

- **[Brazilian road proposal threatens famed biodiversity hotspot](#)** [ 17 August 2021]

News • Scientists and environmentalists say the road, slated to pass through Iguaçu National Park, could harm research projects and precious ecosystems.

- **[Iran hopes to defeat COVID with home-grown crop of vaccines](#)** [ 17 August 2021]

News Q&A • Nature talks to vaccine developer Kayhan Azadmanesh about efforts in Iran to develop ten or more COVID jabs, two of which have been approved for use.

- **[Can artificially altered clouds save the Great Barrier Reef?](#)**

[ 25 August 2021]

News Feature • Australian scientists are rushing to develop new technologies — such as ways to block sunlight — to help preserve corals in the face of climate change.

- **[COVID vaccines and blood clots: what researchers know so far](#)** [ 24 August 2021]

News Feature • Scientists are trying to understand why a small number of people develop a mysterious clotting disorder after receiving a COVID jab.

---

| [Next section](#) | [Main menu](#) | [Previous section](#) |

- NEWS ROUND-UP
- 25 August 2021

# Babbling bats, autocorrected genes and COVID antibodies

The latest science news, in brief.



A greater sac-winged bat (*Saccopteryx bilineata*) pup babbling. Researchers think the young bats make these sounds to practise their vocal skills.Credit: Michael Stifter

## Baby bats babble like human infants

Pups of the greater sac-winged bat (*Saccopteryx bilineata*) develop their vocal skills by [babbling in a similar way to human babies](#), a study shows. The research is the first to identify baby babble produced by a mammal that isn't a primate ([A. Fernandez et al. Science 373, 923–926; 2021](#)).

Researchers recorded 216 babbling bouts in 20 wild bat pups in Costa Rica and Panama. They used ultrasonic sound equipment to capture the individual ‘syllables’ of the pups’ high-pitched squeals, and identified most of the 25 syllables produced by adult bats.

The team converted these audio snippets into images that show the pitch and intensity of the sound over time. This allowed them to search for eight key features that characterize babbling in human babies, including repetition of syllables and rhythm in the sounds. The bats’ babble had all of these features.

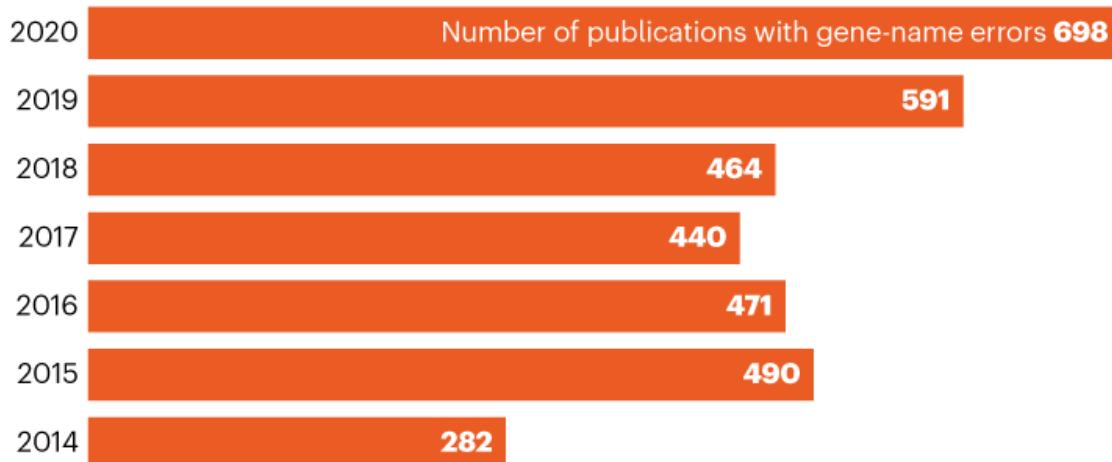
“Even though there are millions of years of different evolutionary pathways between bats and humans, it’s astonishing to see such a similar vocal-practice behaviour leading to the same result — acquiring a large vocal repertoire,” says study co-author Ahana Fernandez, an animal behavioural ecologist at the Berlin Museum of Natural History.

## **Excel autocorrect errors still creating genomics headache**

Five years after a study showed that autocorrect problems were widespread in the genetics literature, [error-riddled spreadsheets persist](#), according to an analysis of published gene lists.

## A GROWING PROBLEM

A 2016 analysis found that 20% of papers featuring gene names had errors created by spreadsheet autocorrect functions, but a bigger survey now finds the proportion is up to 30%. Since 2014, the number of papers with errors has increased significantly.



©nature

Source: Abeysooriya, M., Soria, M., Kasu, M. S. & Ziemann, M. *PLoS Comput. Biol.* **17**, e1008984 (2021).

The issue often occurs when the abbreviated form of a gene's name — or symbol — is incorrectly recognized as a date and autocorrected as such by Excel or Google Sheets. For example, *SEPT4* and *MARCH1* will be automatically changed to *4-Sep* and *1-Mar*. "It can have a significant impact," says biologist Auriol Purdie at the University of Sydney in Australia, who works with gene-microarray and gene-transcription data sets.

The problem was first documented in 2004, and in 2016 Mark Ziemann and his colleagues at the Baker IDI Heart and Diabetes Institute in Melbourne, Australia, quantified it ([M. Ziemann et al. \*Genome Biol.\* \*\*17\*\*, 177; 2016](#)). They found that 20% of papers in leading genomics journals contained conversion errors in Excel spreadsheets in supplementary data. These data sets are frequently used by other geneticists, so errors can be perpetuated and distort further analyses.

However, despite the issue being flagged with researchers — and steps being taken to fix it — the problem is still rife, according to a larger analysis

led by Ziemann, now at Deakin University in Geelong, Australia ([M. Abeysooriya et al. PLoS Comput. Biol. 17; e1008984; 2021](#)). His team found errors in almost one-third of 11,000 articles published between 2014 and 2020.



A garment worker in Gazipur, Bangladesh, receives a dose of the Moderna vaccine against COVID-19. Credit: Ahmed Salahuddin/NurPhoto/Getty

## Antibodies predict protection by Moderna COVID jab

Antibody levels in blood can predict the level of protection provided by Moderna's COVID-19 vaccine. After receiving the vaccine, people with relatively low levels of antibodies were [more likely to develop symptomatic infections](#) than were those who mounted a stronger antibody response, according to an analysis of such 'breakthrough' infections during the vaccine's efficacy trial ([P. B. Gilbert et al. Preprint at medRxiv https://doi.org/grz3; 2021](#)).

The study, posted as a preprint earlier this month, adds to the growing evidence that ‘neutralizing’ antibodies, which can block viral infection of cells, are a marker of vulnerability to COVID-19. A team including David Benkeser, a biostatistician at Emory University in Atlanta, Georgia, compared levels of these antibodies in the nearly 50 vaccinated trial participants who developed breakthrough infections with levels in matched controls who were not diagnosed with COVID-19.

The authors’ modelling found that people with undetectable levels of neutralizing antibodies were 10 times more likely to develop COVID-19 than were individuals whose antibody levels placed them in the 90th percentile of all study participants. The findings have not been peer reviewed.

*Nature* **596**, 467 (2021)

doi: <https://doi.org/10.1038/d41586-021-02289-w>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02289-w>

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

- NEWS
- 18 August 2021

# African languages to get more bespoke scientific terms

Many words common to science have never been written in African languages. Now, researchers from across Africa are changing that.

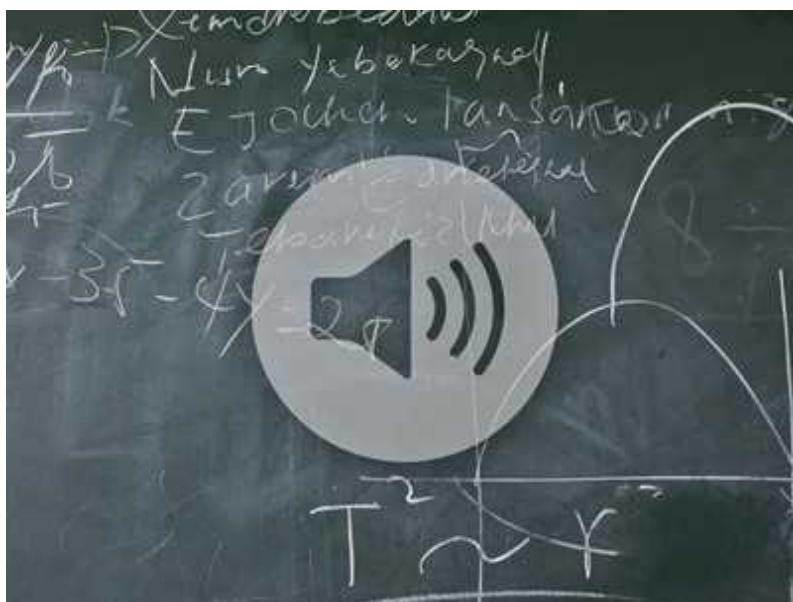
- [Sarah Wild](#)



Researchers want to expand scientific terms in African languages including Luganda, which is spoken in East Africa. Pictured: student-teachers in Kampala. Credit: Eye Ubiquitous/Alamy

There's no original isiZulu word for dinosaur. Germs are called *amagciwane*, but there are no separate words for viruses or bacteria. A quark is *ikhwakhi* (pronounced kwa-ki); there is no term for red shift. And researchers and science communicators using the language, which is spoken by more than 14 million people in southern Africa, struggle to agree on words for evolution.

IsiZulu is one of approximately 2,000 languages spoken in Africa. Modern science has ignored the overwhelming majority of these languages, but now a team of researchers from Africa wants to change that.



### What's the isiZulu for dinosaur? How science neglected African languages

A research project called [Decolonise Science](#) plans to translate 180 scientific papers from the AfricArXiv preprint server into 6 African languages: isiZulu and Northern Sotho from southern Africa; Hausa and Yoruba from West Africa; and Luganda and Amharic from East Africa.

These languages are collectively spoken by around 98 million people. Earlier this month, AfricArXiv [called for submissions](#) from authors interested in having their papers considered for translation. The deadline is 20 August.

The translated papers will span many disciplines of science, technology, engineering and mathematics. The project is being supported by the Lacuna Fund, a data-science funder for researchers in low- and middle-income countries. It was launched a year ago by philanthropic and government funders from Europe and North America, and Google.

## Languages left behind

The lack of scientific terms in African languages has real-world consequences, particularly in education. In South Africa, for example, less than 10% of citizens speak English as their home language, but it is the main teaching language in schools — something that scholars say is an obstacle to learning science and mathematics.

African languages are being left behind in the online revolution, says Kathleen Siminyu, a specialist in machine learning and natural language processing for African languages based in Kenya. “African languages are seen as something you speak at home, not in the classroom, not showing up in the business setting. It is the same thing for science,” she says.

Siminyu is part of Masakhane, a grass-roots organization of researchers interested in natural language processing in African languages. Masakhane, which means ‘we build together’ in isiZulu, has more than 400 members from about 30 countries on the continent. They have been working together for three years.

The Decolonise Science project is one of many initiatives that the group is undertaking; others include detecting hate speech in Nigeria and teaching machine-learning algorithms to recognize African names and places.

Eventually, Decolonise Science aims to create freely available online glossaries of scientific terms in the six languages, and use them to train machine-learning algorithms for translation. The researchers hope to complete this project by the beginning of 2022. But there’s a wider ambition: to reduce the risk of these languages becoming obsolete by giving them a stronger foothold online.

## Terminology creation

Decolonise Science will employ translators to work on papers from AfricArXiv for which the first author is African, says principal investigator Jade Abbott, a machine-learning specialist based in Johannesburg, South Africa. Words that do not have an equivalent in the target language will be flagged so that terminology specialists and science communicators can develop new terms. “It is not like translating a book, where the words might exist,” Abbott says. “This is a terminology-creating exercise.”

But “we don’t want to come up with a new word completely”, adds Sibusiso Biyela, a writer at ScienceLink, a science-communication company based in Johannesburg that is a partner in the project. “We want the person who reads that article or term to understand what it means the first time they see it.”

Biyela, who writes about science in isiZulu, [often derives new terms](#) by looking at the Greek or Latin roots of existing scientific words in English. Planet, for example, comes from the ancient Greek *planētēs*, meaning ‘wanderer’, because planets were perceived to move through the night sky. In isiZulu, this becomes *umhambi*, which also means wanderer. Another word for planet, used in school dictionaries, is *umhlaba*, which means ‘Earth’ or ‘world’. Other terms are descriptive: for ‘fossil’, for example, Biyela coined the phrase *amathambo amadala atholakala emhlabathini*, or ‘old bones found in the ground’.

In some scientific fields, such as biodiversity research, researchers trying to find the right terms will need to tap into spoken sources. Lolie Makhubu-Badenhorst, acting director of the Language Planning and Development Office at the University of KwaZulu-Natal in Durban, says that the absence of a scientific word from written data sets does not mean that it does not exist. “You’re written-centred, I’m oral centred. The knowledge is there, but it is not well-documented,” says Makhubu-Badenhorst, who is not part of the Decolonise Science project.

Decolonise Science’s terminology specialists will come up with a framework for developing isiZulu scientific terms, says Biyela. Once that’s complete, they will apply it to the other languages.

The team will offer its glossaries as free tools for journalists and science communicators, as well as national language boards, universities and technology companies, which are increasingly providing automated translation. “If you create a term and it isn’t being used by others, it isn’t going to permeate into the language,” says Biyela.



Google is calling for help to improve the quality of its African language translations. Credit: Cristina Aldehuela/AFP via Getty

## **Big tech: 'we need your help'**

Masakhane’s researchers say that global technology companies have historically ignored African languages, but in recent years, they have begun funding research in the field.

“We’re aware that the many thousands of African languages are currently under-represented in translation software,” a Google spokesperson told *Nature*. The tech giant wants to expand Google Translate to include more African languages, including Twi, Ewe, Baoulé, Bambara, Fula, Kanuri,

Krio, Isoko, Luganda, Sango, Tiv and Urhobo, they added. However, it needs “speakers of those languages to help us improve the quality of our translations” so they can be integrated into the service.

“The big idea is cultural ownership of science,” Biyela explains. Both he and Abbott say it is crucial to decolonize science by allowing people to do research and speak about science in their own languages. At the moment, it is possible to use African languages to talk about politics and sport, but not science, says Biyela.

Similarly, English is the dominant language of environmental stewardship and conservation — but unless people understand the meaning of specific terms and concepts and can talk about them in their home languages, they can feel disconnected from government efforts to preserve ecosystems and species, says Bheka Nxele, a programme manager for restoration ecology, environmental planning and climate protection in the eThekweni municipality of South Africa.

The researchers are concerned that if African languages are not included in online algorithms, they could, eventually, become obsolete and forgotten. “These are languages [people] speak. These are languages they use every day, and they live with and see the reality that in  $x$  number of years, their language might be dead because there is no digital footprint,” says Siminyu.

*Nature* **596**, 469-470 (2021)

*doi:* <https://doi.org/10.1038/d41586-021-02218-x>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02218-x>

- NEWS
- 18 August 2021

# Decades-old SARS virus infection triggers potent response to COVID vaccines

Dramatic antibody production in people infected during the 2002–04 outbreak furthers hopes of a vaccine against many coronaviruses.

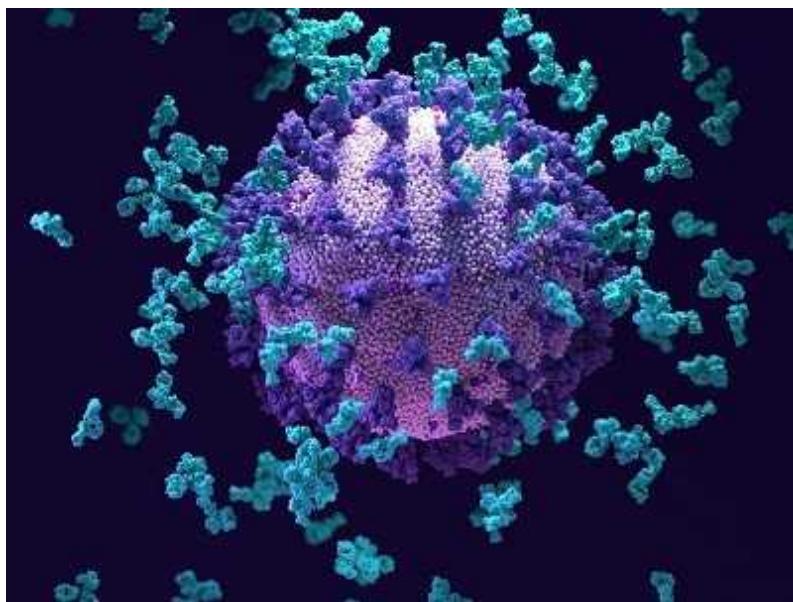
- [Smriti Mallapaty](#)



A person is treated in hospital in Singapore during the 2003 outbreak of severe acute respiratory syndrome. Credit: Paula Bronstein/Getty

People who were infected almost two decades ago with the virus that causes severe acute respiratory syndrome (SARS) generate a powerful antibody response after being vaccinated against COVID-19. Their immune systems can fight off multiple SARS-CoV-2 variants, as well as related coronaviruses found in bats and pangolins.

The Singapore-based authors of a small study published today in *The New England Journal of Medicine*<sup>1</sup> say the results offer hope that vaccines can be developed to protect against [all new SARS-CoV-2 variants](#), as well as other coronaviruses that have the potential to cause future pandemics.



### This ‘super antibody’ for COVID fights off multiple coronaviruses

The study is a “proof of concept that a pan-coronavirus vaccine in humans is possible”, says David Martinez, a viral immunologist at the University of North Carolina at Chapel Hill. “It’s a really unique and cool study, with the caveat that it didn’t include many patients.”

SARS-CoV-2 belongs to the sarbecovirus group of coronaviruses, which includes the virus that caused SARS (called SARS-CoV), as well as closely related bat and pangolin coronaviruses.

Sarbecoviruses use what are known as spike proteins to bind to ACE2 receptors in the membranes of host cells and enter them. They can jump

from animals to humans, as they did before in both the current pandemic and the 2002–04 outbreak of SARS, which spread to 29 countries. “The fact that this has happened twice in the last two decades is strong rationale that this is a group of viruses that we really need to pay attention to,” says Martinez.

## Neutralizing antibodies

Last year, Linfa Wang, a virologist at Duke–NUS Medical School in Singapore who led the latest study, went looking for people who had survived SARS to see whether they offered any clues about how to develop vaccines and drugs for COVID-19. He detected ‘neutralizing’ antibodies in their blood that blocked the original SARS virus from entering cells, but did not affect SARS-CoV-2 — which he found surprising, because the viruses are closely related.

But when Singapore rolled out the Pfizer–BioNTech COVID-19 vaccine this year, Wang decided to interrogate how the SARS infection affected responses to the vaccine. What he discovered was striking. Eight vaccinated study participants, who had recovered from SARS almost two decades ago, produced very high levels of neutralizing antibodies against both viruses, even after just one dose of the vaccine.



[Variant-proof vaccines — invest now for the next pandemic](#)

They also produced a broad spectrum of neutralizing antibodies against three SARS-CoV-2 variants of concern in the current pandemic — Alpha, Beta and Delta — and five bat and pangolin sarbecoviruses. No such potent and wide-ranging antibody response was observed in blood samples taken from fully vaccinated individuals, even those who had also [had COVID-19](#).

The researchers suggest that such broad protection could arise because the vaccine jogs the immune system's 'memory' of regions of the SARS virus that are also present in SARS-CoV-2, and possibly many other sarbecoviruses.

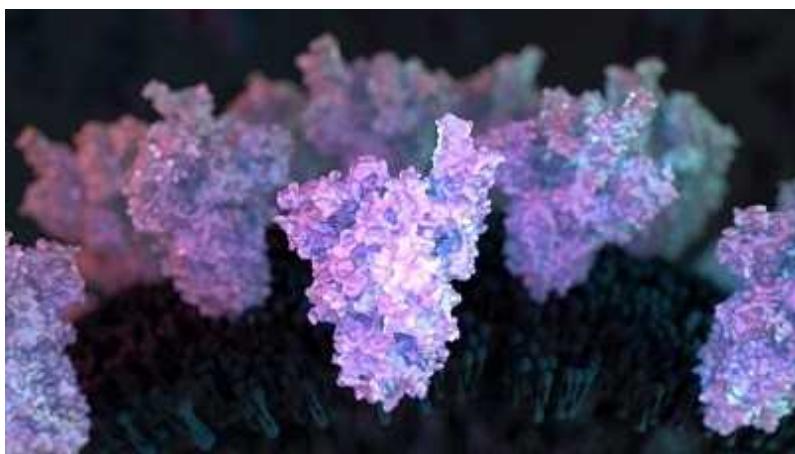
Coronaviruses found in bats have the potential to cause future pandemics, so the fact that a broad spectrum of neutralizing antibodies is generated that protects against some of them "is encouraging", says Daniel Lingwood, an immunologist at the Ragon Institute of MGH, MIT and Harvard in Boston, Massachusetts. But researchers say it is not clear how long this protection lasts.

A vaccine that is widely effective against sarbecoviruses could be administered to the general population in high-risk areas close to animals that harbour them, limiting the potential spread of these viruses in people, adds Christopher Barnes, a structural biologist at Stanford University in California.

## Which part of the virus

Barton Haynes, an immunologist at Duke University School of Medicine in Durham, North Carolina, says the study raises the question of whether a similar response could be generated if people vaccinated against COVID-19 were given a booster shot that targeted the original SARS virus. This might protect them against new variants of SARS-CoV-2 and other sarbecoviruses. Wang says preliminary studies in mice suggest that is possible.

But the latest study doesn't identify exactly which sections of the viruses induce the broad immune response, something that would be needed to develop vaccines. That's the "biggest question", says Martinez. If it is a region of the virus that is present not just in sarbecoviruses, but in the entire group of coronaviruses, there is potential for creating a vaccine against all of them, he says.



### Rare COVID reactions might hold key to variant-proof vaccines

Several research groups have identified [specific antibodies](#) that prevent SARS-CoV-2 and other sarbecoviruses from spreading in cells. Others are already working on pan-coronavirus vaccines, and have synthesized components that induce strong protection in monkeys and mice.

Haynes and his colleagues, for example, have developed<sup>2</sup> a protein nanoparticle studded with 24 pieces of a section of the SARS-CoV-2 spike protein called the receptor binding domain, a key target of antibodies. They

found that in monkeys, the nanoparticle induced much higher levels of antibodies against SARS-CoV-2 than did the Pfizer vaccine. It also induced cross-reactive antibodies against the original SARS virus and bat and pangolin sarbecoviruses.

Martinez and his colleagues have induced these widely reactive antibodies in mice, using a vaccine made from a combination of spike proteins from different coronaviruses<sup>3</sup>. But Martinez says the latest study suggests that this complex spike chimaera might not be necessary; a similar protective response could be induced simply by the original SARS virus's spike protein.

Wang says he is already working on potential vaccines that target multiple sarbecoviruses, and he now hopes to find additional survivors of the 2002–04 SARS outbreak to conduct a much larger study, including testing their responses to other COVID-19 vaccines.

*Nature* **596**, 471–472 (2021)

doi: <https://doi.org/10.1038/d41586-021-02260-9>

## References

1. 1.

Tan, C.-W. *et al.* *N. Engl. J. Med.*  
<https://doi.org/10.1056/NEJMoa2108453> (2021).

2. 2.

Saunders, K. O. *et al.* *Nature* **594**, 553–559 (2021).

3. 3.

Martinez, D. R. *et al.* *Science* <https://doi.org/10.1126/science.abi4506> (2021).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02260-9>

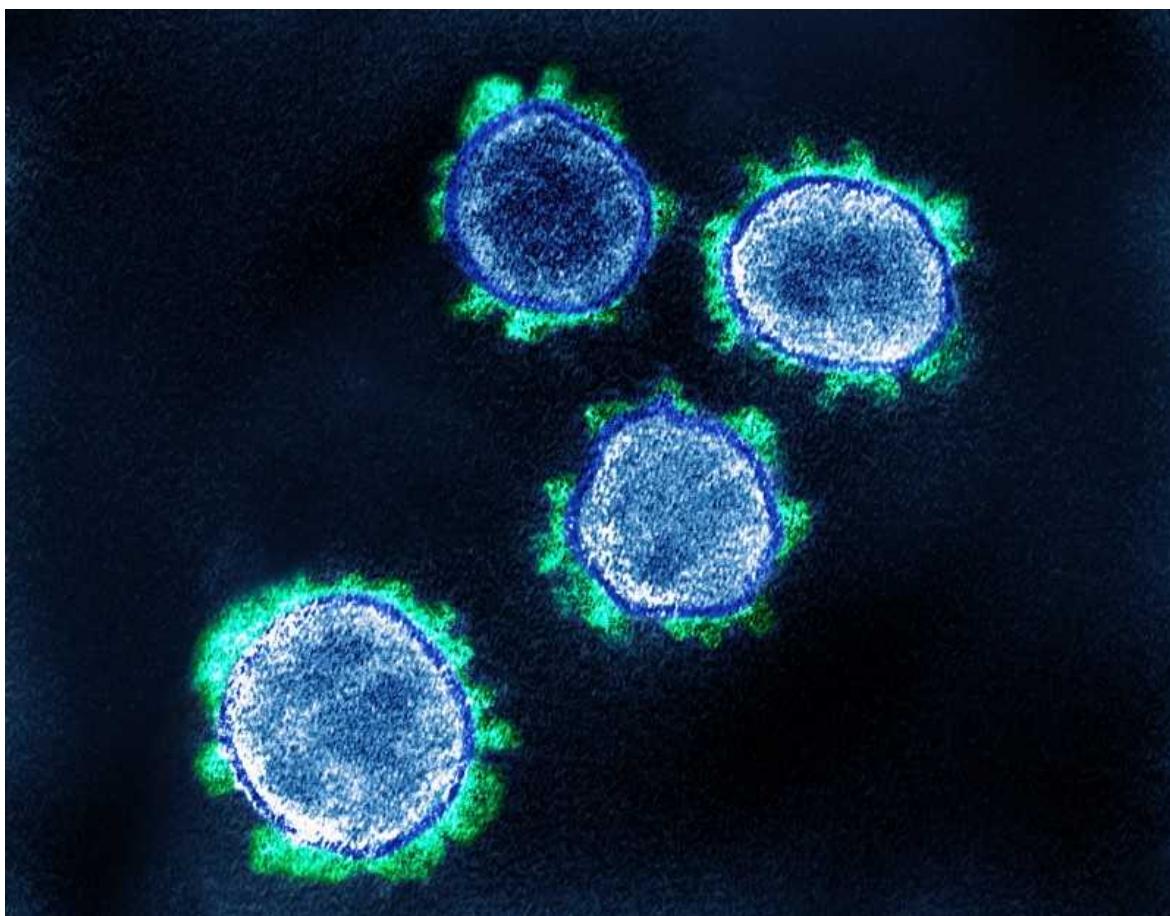
| [Section menu](#) | [Main menu](#) |

- NEWS
- 20 August 2021

# The mutation that helps Delta spread like wildfire

A key amino-acid change might underlie the coronavirus variant's ferocious infectivity.

- [Ewen Callaway](#)



SARS-CoV-2 coronavirus particles isolated from a US case of COVID-19. Scientists are trying to understand why the Delta variant spreads so quickly. Credit: National Institutes of Health/Science Photo Library

As the world grapples with the hyper-infectious Delta coronavirus variant, scientists are racing to understand the biological basis for its behaviour.

A slew of studies has highlighted an amino-acid change present in Delta that might contribute to its swift spread. Delta is at least 40% more transmissible than is the Alpha variant identified in the United Kingdom in late 2020, epidemiological studies suggest.

“The key hallmark of Delta is that transmissibility seems to be ramping up to the next notch,” says Pei-Yong Shi, a virologist at the University of Texas Medical Branch in Galveston. “We thought Alpha was pretty bad, very good at spreading. This one seems to be even more.”



### How do vaccinated people spread Delta? What the science says

Shi’s team and other groups have zeroed in on a mutation that alters a single amino acid in the SARS-CoV-2 spike protein — the viral molecule responsible for recognizing and invading cells. The change, which is called P681R and transforms a proline residue into an arginine, falls within an intensely studied region of the spike protein called the furin cleavage site.

The presence of this short string of amino acids set off alarm bells when SARS-CoV-2 was first identified in China, because it is associated with heightened infectivity in other viruses such as influenza, but had not previously been found in sarbecoviruses, the family of coronaviruses to which SARS-CoV-2 belongs. “This little insert sticks out and hits you in the face,” says Gary Whittaker, a virologist at Cornell University in Ithaca, New York.

## Pre-activated virus

To penetrate cells, the SARS-CoV-2 spike protein must be cut twice by host proteins. In the SARS-CoV-1 virus that causes severe acute respiratory syndrome (SARS), both incisions occur after the virus has locked on to a cell. But with SARS-CoV-2, the presence of the furin cleavage site means that host enzymes (including one called furin) can make the first cut as newly formed viral particles emerge from an infected cell. These pre-activated viral particles can then go on to infect cells more efficiently than do particles requiring two cuts, says Whittaker.

Delta wasn’t the first SARS-CoV-2 variant to gain a mutation that alters the furin cleavage site. The Alpha variant has a different amino-acid change at the same location as Delta. But the available evidence suggests that the mutation’s effect has been especially profound in Delta.

In a study reported as a preprint on 13 August<sup>1</sup>, Shi’s team found that the spike protein is cut much more efficiently in Delta-variant particles than in Alpha particles, echoing results reported in May by virologist Wendy Barclay at Imperial College London and her team, who compared Delta with an earlier strain<sup>2</sup>. Follow-up experiments by both groups showed that the P681R change was largely responsible for spike being clipped so much more efficiently. “This really nailed it, in terms of the mechanism,” says Shi.



## COVID vaccines slash viral spread – but Delta is an unknown

Researchers are also beginning to join the dots between P681R and Delta's ferocious infectivity. Shi's team found that, in cultured human-airway epithelial cells infected with equal numbers of Delta and Alpha viral particles, Delta rapidly outcompeted the Alpha variant, mimicking epidemiological patterns that have played out globally. But Delta's advantage disappeared when the researchers eliminated the P681R change.

The mutation might also speed up the spread of SARS-CoV-2 from cell to cell. A team led by Kei Sato, a virologist at the University of Tokyo, found that spike proteins bearing the P681R change fuse with the plasma membranes of uninfected cells — a key step in infection — almost three times faster than do spike proteins lacking the change<sup>3</sup>.

"I think the virus is succeeding on volume and speed," says Whittaker. "It's become a much more efficient virus. It's going through people and going through cells a lot quicker."

## **More than one mutation**

Although evidence is building that the P681R change is a crucial feature of Delta, researchers emphasize that it is unlikely to be the only mutation responsible for the variant's fast spread. Delta carries numerous other mutations to the spike protein, as well as to other less well-studied proteins,

that might be important. “It’s very simplistic to say it’s just this 681 change. I think it’s a sum of everything,” says Teresa Aydillo-Gomez, a virologist at Icahn School of Medicine at Mount Sinai in New York City.

The epidemiological and genetic context of the mutation is also important to Delta’s rise, say scientists. One of Delta’s siblings, a variant called Kappa that, like Delta, was first identified in India, carries many of the same mutations, including P681R, but its effects haven’t been as devastating as Delta’s. In a preprint posted on 17 August, a team led by structural biologist Bing Chen at Harvard Medical School in Boston, Massachusetts, reports that Kappa’s spike protein is cleaved less frequently and fuses to cell membranes much less efficiently than does Delta’s<sup>4</sup>. The researchers say this finding raises questions over the role of P681R.



### How the Delta variant achieves its ultrafast spread

Researchers in Uganda identified the P681R change in a variant that spread widely in the country in early 2021, but that never took off as Delta did, even though it displays many of the same properties in cell-based lab studies. Whittaker’s team inserted the P681R change into a spike protein from the coronavirus that was circulating in Wuhan, China, at the beginning of the pandemic, and found no increase in its infectivity<sup>5</sup>. “It takes more than one mutation to make a difference,” he adds.

Regardless of its role in Delta's dominance, Whittaker and other scientists say, the mutation has underscored the importance of understanding changes in the coronavirus's furin cleavage site. Whittaker doesn't expect P681R to be the last furin cleavage site mutation to cause concern. "I'm waiting to see what happens next."

*Nature* **596**, 472-473 (2021)

doi: <https://doi.org/10.1038/d41586-021-02275-2>

## References

1. 1.

Liu, Y. *et al.* Preprint at bioRxiv  
<https://www.biorxiv.org/content/10.1101/2021.08.12.456173v1> (2021).

2. 2.

Peacock, T. P. *et al.* Preprint at bioRxiv  
<https://www.biorxiv.org/content/10.1101/2021.05.28.446163v1> (2021).

3. 3.

Saito, A. *et al.* Preprint at bioRxiv  
<https://www.biorxiv.org/content/10.1101/2021.06.17.448820v2> (2021).

4. 4.

Zhang, J. *et al.* Preprint at bioRxiv  
<https://www.biorxiv.org/content/10.1101/2021.08.17.456689v1> (2021).

5. 5.

Lubinski, B. *et al.* Preprint at bioRxiv  
<https://www.biorxiv.org/content/10.1101/2021.06.30.450632v3> (2021).

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

- NEWS
- 17 August 2021

# Brazilian road proposal threatens famed biodiversity hotspot

Scientists and environmentalists say the road, slated to pass through Iguaçu National Park, could harm research projects and precious ecosystems.

- [Meghie Rodrigues](#)



Protesters oppose the Caminho do Colono at Iguaçu Falls. Credit: Marcos Labanca

Brazil's National Congress could soon vote on a bill proposing to construct a road through the country's Iguaçu National Park. If the proposal moves ahead, researchers fear that it will threaten the park's lush forest, a biodiversity hotspot that is home to almost 1,600 animal species, including endangered animals such as the purple-winged ground dove.

Environmentalists and researchers have fought off construction of the 17.5-kilometre road for years, arguing that it will bring not only pollution to the park, but also poachers, who would threaten animals such as jaguars and tapirs. Even research in the park could be affected. In a portion of the park that dips into Argentina, for example, "poachers often steal our cameras", says Julia Pardo, a mammal conservation and ecology researcher at the Subtropical Biology Institute in Misiones, Argentina.



### ['Apocalyptic' fires are ravaging the world's largest tropical wetland](#)

Under the leadership of President Jair Bolsonaro, Brazil's government has weakened protection of the country's forests in favour of industries such as mining, logging and ranching. The lower house of Brazil's Congress, the Chamber of Deputies, put the bill on a fast track in June, allowing it to skip regular debate among its committees and head straight for a vote — a move that has researchers worried.

If passed, the legislation would establish a dangerous precedent that could weaken environmental law in Brazil, says Sylvia Torrecilha, a biologist at the Secretariat of Environment, Economic Development, Production and Family Agriculture in the state of Mato Grosso do Sul. In addition to cutting Iguaçu Park in two with a road that will connect towns to its north and south (see ‘Contested route’), the bill seeks to create a new type of protected area — the *estrada-parque*, or park road — in Brazil’s System of Natural Conservation Units, which regulates environmentally protected areas. Approving the construction of the ‘Caminho do Colono’ (the Settler’s Road) in Iguaçu could literally pave the way for creating through-ways in other parks and conservation areas in Brazil, says Torrecilha.

Normally, the idea of a park road is to preserve the green areas along an already-existing scenic route, she says, not to bring commercial or economic advancement to a state — the argument lawmakers have made in favour of the road. The proposal, from its very beginning, is “inappropriate”, she adds.

## A historical route

Established in 1939, Iguaçu National Park is famous for the waterfall — one of the world’s largest — on the border with Argentina along its southwestern tip. But it is also notable because it contains the largest remaining patch of Atlantic Forest in southern Brazil. Although less well-known than the Amazon rainforest, the Atlantic Forest is rich in plant and animal species, and originally stretched along the coast of southeastern Brazil and down to Argentina and Paraguay. However, the forest is rapidly disappearing: it has lost almost 90% of its tree cover, accelerated by deforestation from urbanization, and agricultural and industrial activities in the twentieth century.

Because of these attributes, the park was designated as a World Heritage site by the United Nations cultural organization UNESCO in 1986.

## CONTESTED ROUTE

10 km

Brazil's National Congress will soon consider whether to pass legislation that would cut the country's Iguaçu National Park in two with a road. The route, called Caminho do Colono, would connect towns and stimulate economic growth, the bill says. But environmentalists worry it will harm wildlife.



If the legislation is successful, it will actually enable the creation of the Caminho do Colono for the second time. The government of Paraná, the state where Iguaçu National Park is located, transformed an existing walking path into an unpaved version of the road during the 1950s. "Nobody cared much at the time because there wasn't much difference between the inside and the outside of the park, as the Atlantic Forest stretched all over the place," says former park chief Ivan Baptiston. "With all the deforestation of the last decades, nowadays, the scenario is a lot different."

In 1986 — the same year the park received its UNESCO World Heritage Site designation — Brazil's Federal Prosecutor's Office filed a civil suit to close the road, and the following year, a federal judge officially closed it. Since then, vegetation has overtaken the route, and some local residents

have tried and failed to force it back open, claiming economic hardships associated with not being able to travel efficiently through the area.



### ['We are being ignored': Brazil's researchers blame anti-science government for devastating COVID surge](#)

The new bill states that re-establishing the road would offer a “solution to a logistical problem in Paraná state”. Sponsored by Nelsi Coguetto Maria, a member of the Chamber of Deputies, the proposal also says it “answers a decades-old outcry of Paraná inhabitants, salvaging the region’s history and its socioeconomic, environmental and tourism relations”.

Environmentalists have criticized Coguetto Maria for backing the bill. And [local media outlets have reported](#) that his family potentially stands to gain from the Caminho do Colono: two of his sons are partners in construction companies that could pave the road. Coguetto Maria’s office did not respond to *Nature*’s queries about this, or about researchers’ concerns over the road. When the Chamber of Deputies approved fast-tracking the bill, he argued that the Brazil of today is “responsible”, and has the “competence and capacity to build an ecologically correct road”, pointing out that the road existed as a walking path before the park was even created.

## **Research interrupted**

For many conservationists and researchers, the economic argument for opening the road doesn't hold water. The damage caused to the park's highly valued Atlantic Forest would far outweigh the potential economic gains for the surrounding towns<sup>1</sup>, they say. Furthermore, the species protected by the park are irreplaceable, they add. Iguaçu is the only location in the world where the jaguar population is increasing instead of declining. If the road opens, says Pardo, pressure on the animals will skyrocket. "Easy access is the main enabler for poachers," she says.



Iguaçu Falls is located along the border of Argentina and Brazil, on the Iguaçu River. Credit: Thiago Trevisan/Alamy

Cars using the road will also cause air, soil, water and even sound pollution, says Victor Prasnewski, a conservation biologist at the Federal University of Mato Grosso in Brazil. Sound pollution, in particular, changes communication patterns among a number of species. "Birds that attract females by singing will be forced to sing louder or longer to get noticed," says Prasnewski, who published a paper last year<sup>2</sup> listing the potential negative impacts of the Caminho do Colono.

“These changes can affect the reproduction and even the evolution of some birds,” says Carlos Araújo, a bioacoustics ecologist at Argentina’s Subtropical Biology Institute. “The building of a road would be catastrophic to research in my field,” he says.

He works on a large-scale monitoring project looking for the purple-winged ground-dove (*Paraclaravis geoffroyi*), the last confirmed sighting of which was more than three decades ago. “It’s a rare animal, and we leave recorders spread over the forest to try and catch her singing. We often capture helicopter noise, which disturbs our work.” Cars and trucks on the road would create similar low-frequency noise, he says. “It will be a lot harder to find birds like this dove.”



### [Brazil's lawmakers renew push to weaken environmental rules](#)

For some, the argument that the road will enhance tourism in Paraná doesn’t make sense either. Reopening the road, says Carmel Croukamp Davies, chief executive of Parque das Aves, a private bird sanctuary and shelter near the park, could threaten Iguaçu’s UNESCO World Heritage title if it damages the park’s biodiversity and severs the Atlantic Forest. Visitors come because they want to experience nature, she adds: “Whoever doesn’t understand the impact of a proposal like this doesn’t understand an inch of tourism nor biodiversity.”

With Brazil's Congress having returned from holiday earlier this month, the bill could soon be put to a vote. And when it is, environmentalists worry it will be passed, given how many representatives in the Chamber of Deputies currently align with Bolsonaro. Then it would face the Senate, and finally, Bolsonaro, who is ultimately expected to approve it.

*Nature* **596**, 473–474 (2021)

doi: <https://doi.org/10.1038/d41586-021-02199-x>

## References

1. 1.

Ortiz, R. A. *Ambientalia* **1**, 141–160 (2009).

2. 2.

Prasniewski, V. M. *et al.* *Ambio* **49**, 2061–2067 (2020).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02199-x>

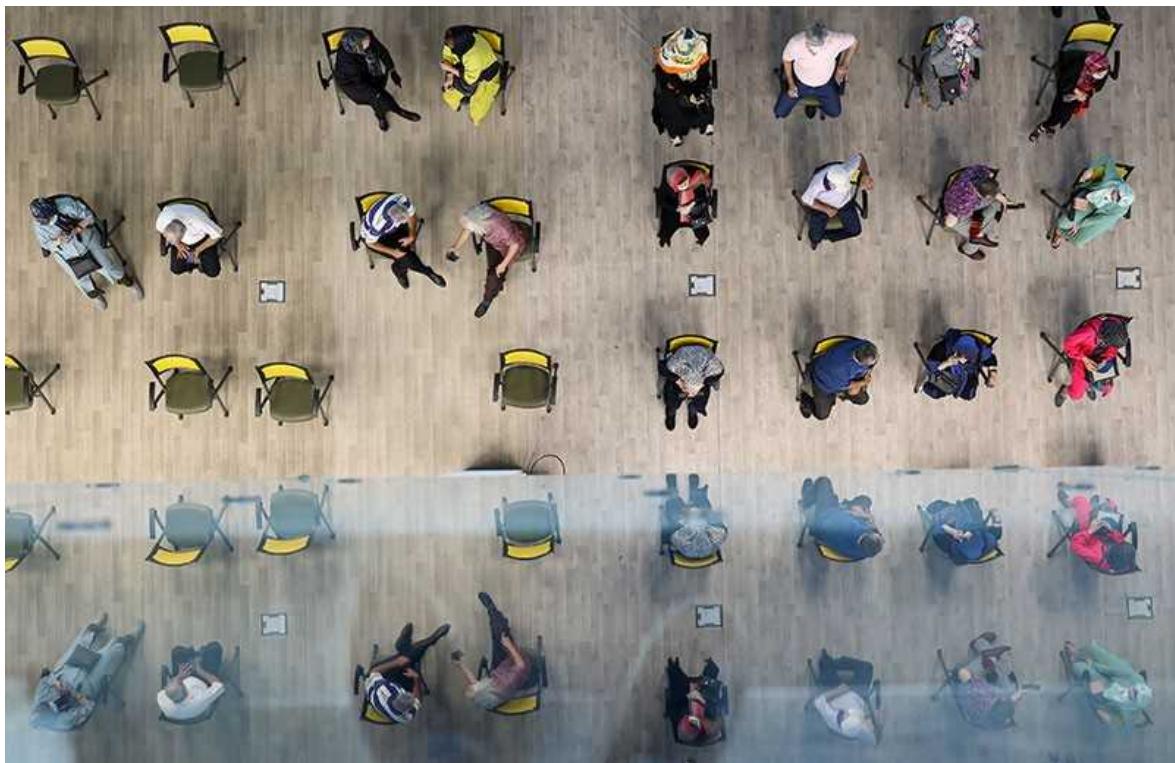
| [Section menu](#) | [Main menu](#) |

- NEWS Q&A
- 17 August 2021

# Iran hopes to defeat COVID with home-grown crop of vaccines

*Nature* talks to vaccine developer Kayhan Azadmanesh about efforts in Iran to develop ten or more COVID jabs, two of which have been approved for use.

- [Smriti Mallapaty](#)



Iranians wait to be inoculated at a mass vaccination centre in Tehran.Credit: Majid Asgaripour/WANA via Reuters

Iran was among the first countries to be hit with an outbreak of COVID-19 in early 2020. It is currently battling its fifth wave, probably driven by the Delta variant. Official figures suggest that more than 4.3 million people have been infected and 97,000 have died since the pandemic began, but the true toll is potentially much higher.

Scientists say Iran is one of few Middle Eastern nations with the capacity to develop vaccines. It has been doing so in earnest: around ten are under development and one is already bolstering its vaccination drive, but little is known about these vaccines outside Iran.

*Nature* speaks to Kayhan Azadmanesh, a medical doctor and biotechnologist who is head of the virology research division at the Pasteur Institute of Iran in Tehran, about the nation's vaccine landscape. Azadmanesh also advises the Iranian government and is developing two viral-vector vaccines through his spin-off company Humimmune Biotech.

## How badly has the pandemic affected Iran?

Since January 2020, we've had five separates waves. We're currently experiencing the highest number of new cases reported so far, with around 40,000 a day, and the most common variant we detect is Delta. But many more cases are likely going unreported. The outbreak is putting pressure on hospitals and the situation is not looking good.

## Which COVID-19 vaccines are available in Iran?

So far, 18 million or so doses have been administered: some 12 million were China's Sinopharm vaccine; 4 million were the Oxford–AstraZeneca vaccine; and one million were COVIRan Barekat, developed by the Iranian state-owned Shifa Pharmed Industrial Group in Tehran. The remainder include doses of Russia's Sputnik V and India's Covaxin. More than half a million doses are being administered a day, and some 17% of Iran's population of 85 million have received their first dose of a COVID-19 vaccine.

## Could you tell us about COVIran Barekat?

It is an inactivated vaccine and is still undergoing phase III trials, but it received emergency-use authorization in June. It was approved on the basis of the levels of antibodies it induces, including those that can ‘neutralize’ SARS-CoV-2, or block it from entering cells. In early trials, the researchers found that more than 93% of vaccinated people produced neutralizing antibodies. We don’t know how long this protection will last, but I assume that it will be similar to that provided by other inactivated vaccines — such as CoronaVac, produced by the Chinese firm Sinovac Life Sciences — for which antibody levels have been shown to drop after six months<sup>1</sup>, suggesting that boosters are likely to be required.



Kayhan Azadmanesh holding rapid-antibody tests that show a negative result before, and a positive result after, he received a COVID-19 vaccination.Credit: Kayhan Azadmanesh

## **What other vaccines are being developed in Iran?**

Pasteurcovac is a recombinant-protein vaccine developed in a collaboration between Cuba's Finlay Institute of Vaccines in Havana and the Pasteur Institute of Iran. The vaccine is known as [Soberana 02](#) in Cuba. It also received emergency-use approval in Iran in June, despite still being in phase III trials. There are several other inactivated vaccines and recombinant-protein vaccines in clinical trials, and there is at least one mRNA vaccine, two adenovirus-vector vaccines and one measles-virus-vector vaccine in earlier stages of development. Vaccines developed outside Iran are also currently in clinical trials and being produced locally.

## **Tell me about the vaccines you are designing?**

My company, Humimmune Biotech, has been working on two vaccine candidates. One uses the measles virus as a backbone to introduce a gene that encodes either the SARS-CoV-2 spike protein, which the virus uses to enter cells, or the nucleocapsid protein that it requires to replicate. That vaccine is being produced by the Iranian firm BioSun Pharmed in Tehran.

The other vaccine, which might be more promising, uses an adenovirus 5 backbone to deliver part of the sequence for the spike protein — a similar backbone to that used in the second dose of Sputnik V. We hope to start clinical trials early next year. Most of the COVID-19 vaccines used in Iran so far have been inactivated vaccines, which I expect will mean people will need booster shots next year. Our vaccine could be used as a booster, and a mix-and-match approach might even offer better protection. The technology can also be easily modified against new variants — we have already begun developing a version for Delta.

## **Why are Iranian scientists creating so many vaccines?**

We have a long history of vaccine production in Iran. The Pasteur Institute of Iran was established in 1920, and has produced vaccines against

tuberculosis and rabies. Vaccines have also been developed in Iran against measles, mumps and human papilloma virus.

We can't rely on help from the international community with the pandemic. We are living under sanctions imposed by the United States; in our opinion, these are unjustified. The United States says that sanctions don't affect humanitarian activities, but when your ability to transfer money is restricted, it is difficult to buy drugs and medicines. And we have the technology to produce vaccines, so why not use it? To ensure the safety of Iranians, it makes sense to develop a variety of vaccines using different research and development strategies, as China has done.



Iran's Supreme Leader Ayatollah Ali Khamenei receives a dose of the locally made COVIran Barekat vaccine in June. Credit: Iranian Supreme Leader's Office/ZUMA/Shutterstock

## **Why are Iranian researchers reluctant to publicize their work internationally?**

This could be another side effect of the sanctions. Researchers in Iran might not want to draw too much attention to their work in case they put potential partnerships in jeopardy before they have achieved a final product, or they run the risk of losing access to raw materials and technologies they need for vaccines.

Researchers are also extremely busy, helping in the effort to fight the pandemic in Iran. They might not have time to publish results in international journals. But some have started to share results. In June, the researchers developing COVIran Barekat published a preprint of their preclinical results<sup>2</sup>, and they will share clinical results very soon. We also plan to share the results of our adenovirus-vector vaccine soon.

## **What have been the biggest challenges in developing COVID-19 vaccines?**

The sanctions have caused a lot of difficulty, because they make it hard for us to buy materials and equipment. For example, chromatography resins we need to purify vaccines are mostly produced by multinational companies that are major suppliers to the United States, so they might be afraid of selling to us. The United States says that we can apply for exemptions, but, in our experience, that hasn't worked. But somehow, we find a way. We modify our methods, find other providers, or look for local solutions. We search for the best we can get, but sometimes quality and efficiency are affected.

Also, one of the biggest challenges globally is scale. Prior to the pandemic, Iran primarily had to produce vaccines for children, with a production requirement for each vaccine of around three million doses a year. Now we need about 170 million doses to fully vaccinate the whole population.

## **What does the future hold for vaccine development in Iran?**

The initial target for COVIran Barekat was to produce up to 30 million doses a month by September, which would have been enough to vaccinate the adult population. But they have not been able to achieve that, so we have

had to import millions of doses of other vaccines. As many people have said, this will not be the last coronavirus pandemic that we face. I expect the vaccine production capacity will be used for years to come to develop new vaccines and drugs, for both coronaviruses and other diseases.

*Nature* **596**, 475 (2021)

*doi:* <https://doi.org/10.1038/d41586-021-02216-z>

*This interview has been edited for length and clarity.*

## References

1. 1.

Li, M. *et al.* Preprint at medRxiv  
<https://doi.org/10.1101/2021.08.03.21261544> (2021).

2. 2.

Abdoli, A. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.06.10.447951> (2021).

---

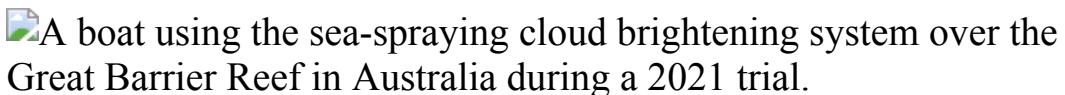
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02216-z>

- NEWS FEATURE
- 25 August 2021

# Can artificially altered clouds save the Great Barrier Reef?

Australian scientists are rushing to develop new technologies — such as ways to block sunlight — to help preserve corals in the face of climate change.

- [Jeff Tollefson](#)



During a field trial, a turbine generates plumes of seawater droplets that rise into the sky. Credit: Brendan Kelaher/SCU

In place of its normal load of cars and vans, the repurposed ferry boat sported a mobile science laboratory and a large fan on its deck as it left Townsville, Australia, in March. Researchers dropped anchor in a coral lagoon some 100 kilometres offshore and then fired up the cone-shaped turbine, which blew a mist of seawater off the back of the boat. What happened next came as a welcome surprise: after briefly drifting along the ocean surface, the plume ascended into the sky.

Looking a bit like a jet engine, this mist machine is at the centre of an experiment that, if successful, could help to determine the future of the Great Barrier Reef. Three-hundred and twenty nozzles spewed a cloud of nano-sized droplets engineered to brighten clouds and block sunlight — providing a bit of cooling shade for the coral colonies below. Scientists used sensors aboard the ferry, drones and a second boat to monitor the plume as it migrated skyward.

The experiment wasn't big enough to significantly alter the clouds. But preliminary results from the field tests — which were shared exclusively with *Nature* — suggest that the technology might perform even better than computer models suggested it would, says Daniel Harrison, an oceanographer and engineer at Southern Cross University in Coffs Harbour, Australia, who is heading up the research. "We are now very confident that we can get the particles up into the clouds," Harrison says. "But we still need to figure out how the clouds will respond."

Harrison's project is the world's first field trial of marine cloud brightening, one of several controversial [geoengineering technologies](#) that scientists have studied in the laboratory for decades. The research has been driven by fear that humans might one day be forced to deliberately manipulate the Earth's climate and weather systems to blunt the most severe impacts of global warming.

For many Australians, that day arrived in 2017, when a [marine heat wave spurred massive coral bleaching](#) and death across much of the 2,300-kilometre Great Barrier Reef. That crisis hit just a year after another bleaching event along the reef, which supports more than 600 species of coral and an estimated 64,000 jobs in industries such as tourism and fishing. Research suggests that the reef lost more than half of its coral between 1995 and 2017, as a result of warming waters, tropical storms and predatory starfish ([A. Dietzel et al. Proc. R. Soc. B. 287, 20201432; 2020](#)).



## These corals could survive climate change — and help save the world's reefs

The project has raised concerns among some scientists abroad, in part because the Australian group has published little about its work. Environmentalists outside Australia objected to the project last year after news of the first trial broke, and there could be similar criticism when details of the 2021 trial emerge.

Harrison stresses that the cloud-brightening project is about local adaptation to climate change, not global geoengineering, because its application would be limited in both space and time. It's also just one part of a larger Aus\$300 million (US\$220 million) Reef Restoration and Adaptation Program (RRAP) launched last year by Australia to investigate and develop techniques and technologies to save the country's reefs. Many of the proposals, from cloud brightening to [breeding heat-tolerant corals](#), would represent unprecedented human interventions in the natural reef system.

Ecological modelling suggests that a large-scale intervention involving multiple strategies — including a fleet of mist machines — could prolong the life of the reef while governments work to eliminate greenhouse-gas emissions. The goal now is to work out what's achievable in the real world, says Cedric Robillot, executive director of the RRAP.

“You need to consider every angle, from the fundamental science to the very pointy end of engineering, if you want to succeed,” Robillot says. “It’s not enough to just prove you could do it. You need to explain how you would do it.”

## **Into the clouds**

Harrison conducted his first field test in March 2020: a three-day proof-of-concept expedition on a small car ferry with four scientists, one representative from a local Indigenous group, and two shipping containers for equipment and sleeping quarters. The team had a minimal Aus\$400,000 budget and limited scientific instrumentation to monitor the mist, but it was enough to document that the plume flowing out of their mist machine rode a draught of warm air high into the sky.

It was the first time they had witnessed this phenomenon. Their models had suggested that evaporation of the brine droplets would cool the plume, which would then float across the surface of the ocean, only slowly mixing upwards into the low-lying marine clouds. The models also indicated a risk that the tiny droplets might merge and drop out of the air. Instead, brine droplets floated along the surface of the ocean for half a kilometre without coalescing, gradually losing water and weight to evaporation along the way. And then they shot upwards.



A marine heat wave in 2017 caused coral bleaching along much of Australia's Great Barrier Reef. Credit: Juergen Freund/Nature Picture Library

“We didn’t expect that at all,” Harrison says, “but it turned out we were doing this experiment in the middle of a rising air mass.”

The scientists feared it was a fluke. Although years of research and development have gone into the nozzles, initially led by a separate American team, this was the first time anybody had ever deployed them in the field with fresh seawater. The team also didn’t know what to expect from clouds

and aerosols in that region, because research on the reef has focused almost exclusively on what happens below the water, not the conditions above.

For Harrison, the 2020 experiment was more than enough to justify moving forward with another, larger trial in March 2021. But it did raise eyebrows among some scientists and observers abroad, where geoengineering research has met strong opposition and struggled to attract funding.



### [IPCC climate report: Earth is warmer than it's been in 125,000 years](#)

Most of the concern has centred on a form of solar geoengineering that involves injecting reflective material into the stratosphere to block sunlight at a global scale. But cloud brightening has also been studied as a potential global intervention, and it has attracted criticism from some environmental groups who argue that it carries inevitable ecological risks and detracts from efforts to limit greenhouse gases.

Some scientists, as well as environmental advocates who follow geoengineering research, told *Nature* that they were surprised to see the experiment move forward without more scrutiny — or without published research to justify such an investment.

Critics also worry that Australia is setting the wrong kind of precedent by rebranding a solar-geoengineering experiment that could have regional

impacts as a local adaptation project. “One could say that there should have been some level of consultation with the outside world,” says Janos Pasztor, who heads the Carnegie Climate Governance Initiative, an advocacy group in New York City that has been pushing for a global debate over geoengineering governance in the United Nations.

Harrison says scientists in the programme have consulted with regulatory authorities, as well as with the general public and Indigenous groups that have historic claims on the reef. He also readily acknowledges trying to avoid getting embroiled in a debate about solar geoengineering, arguing that the project would be more akin to cloud-seeding operations that are designed to promote rain and that are not considered to be geoengineering. One of the next modelling efforts, however, will be to explore any potential regional and global implications, he says.



A plume of seawater droplets rises up into the sky during a field trial in March 2021.Credit: Brendan Kelaher/SCU

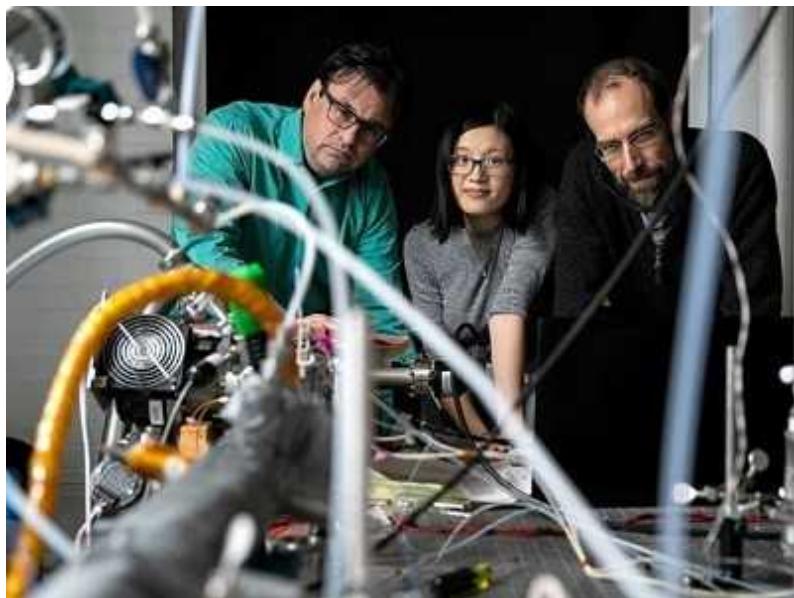
Others question the Australian government’s motivations in funding such work. Under the conservative prime minister Scott Morrison, the government has yet to strengthen its climate pledge under the 2015 Paris agreement, as many nations have done in the past year. Morrison has personally ruled out committing to net-zero emissions. Pushing for a

technological fix to global warming without moving to aggressively curb greenhouse gases is “sheer lunacy”, says Peter Frumhoff, chief climate scientist for the Union of Concerned Scientists, an advocacy group in Cambridge, Massachusetts.

Some researchers, however, are pleased to see marine cloud brightening move from theory to the field, including US scientists working on a similar project that has been struggling to get into the field for nearly a decade. “This is an early example of how climate disruption can drive interest in these things,” says Sarah Doherty, an atmospheric physicist who manages the Marine Cloud Brightening Project at the University of Washington in Seattle. Members of the team provided the initial nozzle design and have been tracking the Australian group’s progress.

## Coral crisis

The first time that scientists observed a major bleaching event along the Great Barrier Reef was in 1998, and the second event followed four years later. In both cases, corals expelled the algae that live within them and that provide colour and energy through photosynthesis. Most of the corals eventually recovered. But in 2016 and 2017, many corals bleached and then died across two-thirds of the reef.



## First sun-dimming experiment will test a way to cool Earth

“It was absolutely horrifying,” says David Wachenfeld, chief scientist at the Great Barrier Reef Marine Park Authority, which manages the reef. The clear message from those events was that the traditional approach to managing corals and coral reefs would not be enough, he adds. “Our hand was forced.”

In 2018, the Australian government allocated Aus\$6 million to a consortium of universities and government research institutes for a feasibility study focused on potentially radical strategies that could be applied across the reef. Researchers reviewed some 160 ideas, including putting live corals on ice for long-term preservation and synthetically engineering new varieties that can tolerate the warmer waters. Many approaches proved too costly and energy intensive, but 43 interventions were singled out for further study. Marine cloud brightening drew support in part because it theoretically provides direct relief precisely when and where corals need it most.

Much of the emphasis of the programme is on helping corals to adapt and repopulate the reef, including efforts to improve coral aquaculture operations so that they can produce millions of corals per year rather than thousands. For Madeleine van Oppen, a coral geneticist at the Australian Institute of Marine Science near Townsville, the RRAP programme helps to integrate her team’s work on assisting coral evolution to make them more heat tolerant.

Thanks to the RRAP, she says, data from those projects are now being fed directly into models that enable researchers to assess the potential benefits — as well as the risks — of releasing new strains of coral and microalga into the wild. The programme is also raising ecological questions, such as whether the introduction of new coral species can propagate disease, or whether a new variety of more heat-tolerant corals might displace corals struggling to survive.



Researchers are testing specialized nozzles that create jets of seawater mist. Credit: Alejandro Tagliafico/SCU

“It sort of speeds up the whole path from research to implementation in the field,” says van Oppen.

In the long run, the models indicate that without interventions, the extent of coral on the reef could shrink by well over 60% by 2070 compared with 2020 levels ([S. A. Condie et al. R. Soc. Open Sci. 8, 201296; 2021](#)). But simulations suggest that Australia could cut those losses in half with a three-pronged approach focused on propagating heat-tolerant corals, controlling outbreaks of the predatory crown-of-thorns starfish and brightening clouds to take the edge off of heat waves. Crucially, the latest modelling also suggests that without the cooling provided by Harrison’s cloud brightening project, the other interventions might not amount to much.

## Testing the wind

When Harrison’s group returned to the field this year, they had more-powerful drones as well as other aerosol sensors on a second boat. As in the previous year’s experiment, each time they created a plume, it rose into the sky after the droplets lost around 90% of their water to evaporation. The

likely explanation, Harrison says, is that the reef is creating its own weather as warm water along the shallow corals heats the air above.

Many more droplets are making it into the clouds than the scientists had initially calculated, but Harrison says their mist machine might need to be scaled up by a factor of 10 — from 320 to around 3,000 nozzles — to produce enough particles to brighten nearby clouds by around 30%. His team's modelling suggests that this could in turn reduce the incoming solar radiation on the reef locally by around 6.5%. Even then, the operation would require 800–1,000 stations to cover the length of the Great Barrier Reef.



### [Fevers are plaguing the oceans — and climate change is making them worse](#)

But it's unclear whether that spray of salty droplets will have the desired effect, says Lynn Russell, an atmospheric chemist at the Scripps Institution of Oceanography in La Jolla, California, who has studied cloud brightening. Russell has not seen the latest — and as-yet unpublished — results, but questions whether there are enough of the low layered clouds considered suitable for cloud brightening.

Harrison acknowledges such concerns and says that his team sees more of these clouds on the southern part of the reef. His team's modelling suggests the technology will also work on the clouds that are common across the rest of the reef in summer. Even then, he says, it remains unclear how much

coverage a full-scale cloud-brightening operation could provide across the entirety of the reef. More measurements, and detailed modelling, are needed to provide answers.

For now, Harrison has secured funding for another two years, and he needs to demonstrate progress. The RRAP is testing all 43 approaches and will redistribute resources to projects that show potential, Robillot says. But he stresses that no amount of science and engineering will preserve the reef in its current form. “Even if we do all of this, the system that you’ll end up with is not going to be the Great Barrier Reef that we know today,” Robillot says. “You might, however, retain a very functional ecosystem.”

That’s enough to keep Harrison going, and his team is already preparing for a trip into the field in 2022. The scientists plan to run the mist machine at higher pressure, which should produce a sixfold increase in the number of particles, and they will use new instrumentation to determine how particles alter clouds. They are also investigating an entirely different nozzle technology that could reduce the number of nozzles needed by a factor of 1,000.

Harrison is more confident today than he was even a year ago that cloud brightening might work over the reef, but he is also realistic about the future if governments fail to limit carbon emissions. “There are only so many clouds available, and there is only so much you can brighten them,” he says. “Eventually, climate change just overwhelms things.”

*Nature* **596**, 476-478 (2021)

doi: <https://doi.org/10.1038/d41586-021-02290-3>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02290-3>

- NEWS FEATURE
- 24 August 2021

# COVID vaccines and blood clots: what researchers know so far

Scientists are trying to understand why a small number of people develop a mysterious clotting disorder after receiving a COVID jab.

- [Heidi Ledford](#)



A health worker administers the Oxford–AstraZeneca COVID-19 vaccine in Mexico City. Some countries have restricted its use. Credit: Leonardo Montecillo/Agencia Press South/Getty

It was when the second person with unusual clots came in that Phillip Nicolson knew something was wrong. Blood clots are uncommon in young people, and it's even rarer to see a combination of blood clots and alarmingly low levels of platelets — cell fragments that help to form clots.

Yet in the space of one week in March, two young people with this pairing of symptoms had arrived at the Queen Elizabeth Hospital in Birmingham, UK, where Nicolson works as a haematology specialist. And both had recently been given the Oxford–AstraZeneca COVID-19 vaccine.

Nicolson had been on call at the hospital all weekend, and had been looking forward to a rest on Monday. Instead, he found himself rushing around to get consent to collect samples to study in the laboratory. By the time he arrived at the second patient's bedside, a third had been admitted.

That week, Nicolson was among the first to witness what researchers now call vaccine-induced immune thrombotic thrombocytopenia (VITT), a life-threatening and mysterious condition that affects a very small number of people who have received the Oxford–AstraZeneca or Johnson & Johnson (J&J) COVID-19 vaccines. It is now estimated that VITT occurred in about 1 in 50,000 people aged under 50 who received the Oxford–AstraZeneca vaccine<sup>1</sup>. This and similar observations in other countries have led some officials to delay and then scale back the deployment of these vaccines.

Despite fervent work by researchers such as Nicolson, the mechanism that links the vaccines and VITT is still uncertain. Establishing a mechanism could reveal ways to prevent and treat the condition, and improve the design of future vaccines. Over the past few months, researchers have gathered clues and developed a host of hypotheses.

Working through these possibilities is a daunting task. “You can have your hypothesis, but how do you find which is the one that caused an event in maybe 1 in 100,000 people?” asks John Kelton, a haematologist at McMaster University in Hamilton, Canada. “It’s really, really hard.”

## Clotting concepts

The unusual constellation of symptoms was immediately familiar to some haematologists, particularly those with experience of treating people with a rare reaction to the anti-clotting drug heparin. That syndrome, called HIT, is also characterized by low platelet counts and sometimes the presence of clots.

HIT is caused by heparin, a negatively charged molecule, binding to a positively charged protein called platelet factor 4 (PF4) that is produced by platelets to promote clotting. In some people, the immune system views this complex as foreign, and develops antibodies against it.



### [COVID vaccines and safety: what the research says](#)

These antibodies can also bind to and activate platelets, priming them to clump together and trigger clotting. The clots can clog up important blood vessels, and the condition can be fatal, although some treatments improve the chances of survival.

Only a handful of labs around the world study HIT, and those that do scrambled to get samples from the few people who had been diagnosed with VITT. When researchers analysed the samples, it was clear that vaccine recipients who had this mysterious clotting reaction were also producing antibodies against their own PF4<sup>2-4</sup>. But it was anyone's guess as to what had triggered these antibodies. Kelton, who has been studying HIT for decades, had to wait to get precious specimens from people with VITT, and

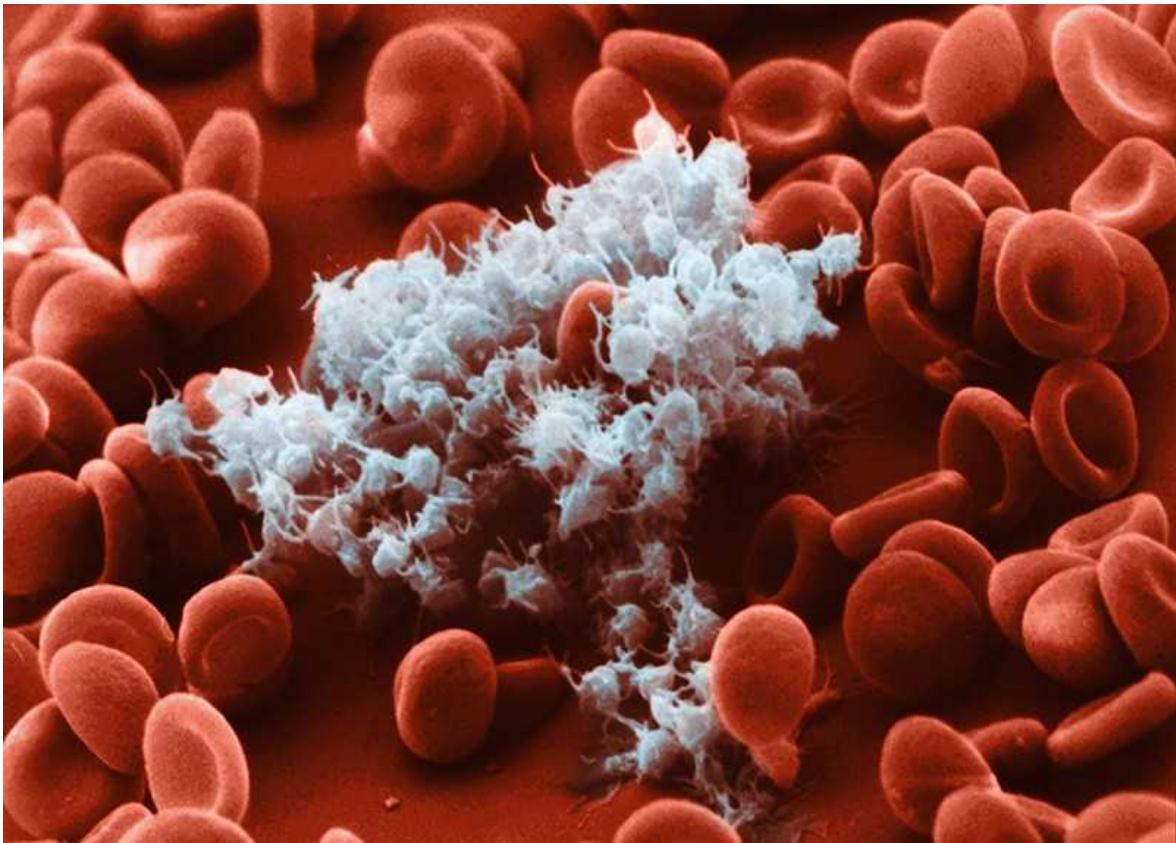
then his team had to wade through samples of varying quality. Some were contaminated by treatments the people with VITT had received. “Many, many samples were not what I would call pristine,” he says. “These people are as sick as can be, and the physicians throw the book at them. They have every kind of chemical in them.” And about two-thirds of the samples his team received lacked the PF4 antibodies altogether, suggesting that the patients did not have VITT, but instead had developed a clotting disorder that was probably unrelated to their vaccination, Kelton says.

Eventually, his team was able to get five samples taken from people before treatment for VITT. When researchers characterized antibodies in the samples, they found that some were binding to PF4 at the same site as the one used by heparin, and that they were also capable of activating platelets<sup>5</sup>. The results suggested that the mechanism behind the vaccine-linked syndrome was similar to that of HIT — but the trigger seemed to be the vaccine rather than heparin.

## Suspect ingredients

Something in the vaccine or the body’s response to it must be binding to PF4 — but what? VITT has been linked to two COVID-19 vaccines, both of which use disabled adenoviruses as a ‘vector’ to shuttle a gene encoding a coronavirus protein, called spike, into human cells. Once there, the gene is expressed and the protein is made. The immune system detects spike and generates antibodies against it that are crucial for protection against coronavirus infection.

Some researchers have proposed that impurities in the vaccines left over from the manufacturing process — such as snippets of DNA floating around in the solution, or proteins in the broth used to grow the virus — are interacting with PF4 to generate the clumps that are then targeted by antibodies<sup>6</sup>.



Platelets (white) are fragments of cells that encourage clots to form. Credit: Lennart Nillson, Boehringer Ingelheim International

Others think the culprit could be the adenovirus itself. Previous work has shown that adenoviruses can bind to platelets and trigger their depletion in mice<sup>7</sup>. It's conceivable that those mice might also have developed clots if they had been followed for longer, says Maha Othman, who studies blood clotting at Queen's University in Kingston, Canada, and was lead author of the study.

Before the COVID-19 pandemic, adenovirus-based vaccines were being developed against infections such as HIV and Ebola, but had not yet been used in large populations. There have been no reports that these vaccines produced a VITT-like condition; however, they were not tested in nearly as many people as have received the Oxford–AstraZeneca COVID-19 vaccine.

Haematologist Mitesh Borad at the Mayo Clinic in Phoenix, Arizona, and his colleagues have analysed the structure of the chimpanzee adenovirus

used in the Oxford–AstraZeneca vaccine and determined that it has a strong negative charge. Molecular simulations suggest that this charge, combined with aspects of the virus's shape, could allow it to bind to the positively charged PF4 protein<sup>8</sup>. If so, it could then set off a cascade much like the rare reaction to heparin, says Borad, although it remains to be seen whether this happens.



### How could a COVID vaccine cause blood clots? Scientists race to investigate

Even if the adenovirus is to blame, Borad says he would not advocate that vaccine developers stop using adenoviruses in vaccines. Some adenoviruses could be engineered to reduce their negative charge, he says, and some are less negatively charged than others; the Ad26 adenovirus used in the J&J COVID-19 vaccine does not have as much of a charge as the chimpanzee virus, which might explain why VITT seems to be less common in recipients of the J&J vaccine. And so far, no link to VITT has been reported for the Sputnik V COVID-19 vaccine, which uses both Ad26 and another adenovirus called Ad5 that has still less negative charge, he adds.

Then there's the spike protein itself. One team of researchers wondered whether the antibodies that bind to PF4 in people with VITT are an unintended by-product of the body's immune response to spike. But they

found that the PF4 antibodies can't bind to it, suggesting that they are not part of the immune response to the viral protein<sup>9</sup>.

But cancer researcher Rolf Marschalek at Goethe University Frankfurt in Germany and his colleagues have shown that the snippets of RNA that encode spike can be cut apart and stitched back together in different ways in human cells; some of these forms, called splice variants, can generate spike proteins that get into the blood and then bind to the surface of cells that line blood vessels<sup>10</sup>. There, they cause an inflammatory response that is also seen in some SARS-CoV-2 infections, which in severely affected people can lead to the formation of clots.

And the lower rate of clots in J&J's vaccine compared with Oxford–AstraZeneca's could be because the version of spike generated by the J&J vaccine was engineered to remove the sites that allow the RNA to be processed into splice variants, says Marschalek.

Marschalek thinks that if this idea is borne out, then the Oxford–AstraZeneca vaccine and other adenovirus-based vaccines could be rendered safer if their versions of spike were similarly engineered.

There are reports that the teams behind the Oxford–AstraZeneca and J&J vaccines are working to develop safer adenoviral vectors, and Marschalek says he would be surprised if companies abandoned adenoviral vectors altogether. Others agree. “I think they are very popular and will remain popular,” says Othman, citing the ease with which the vaccines can be produced and manipulated, and the wealth of data suggesting that, for most people, the vaccines are safe. Instead of abandoning them, she says, “we should study more about the immune responses to them.”



## Why is it so hard to investigate the rare side effects of COVID vaccines?

One possible factor affecting the safety of adenoviral vaccines is how they are administered. The COVID-19 vaccines are given as injections into muscle, but if the needle happens to puncture a vein, the vaccine could enter the bloodstream directly. Leo Nicolai, a cardiologist at Ludwig Maximilian University of Munich, Germany, and his colleagues found in a mouse study that platelets clump together with adenovirus and become activated when the Oxford–AstraZeneca vaccine is injected into blood vessels, but not when it is injected into muscle<sup>11</sup>.

It's possible, says Nicolai, that on rare occasions, a vaccine is inadvertently injected into a vein — as was done in the earlier mouse studies that found that adenovirus could bind to platelets. If so, many cases of VITT might be avoided by asking vaccinators to first draw a small amount of fluid from the injection site with the syringe to check for blood before they actually push the plunger to administer the vaccine. This is already standard practice in some countries, and Denmark has added it to its official guidelines for COVID-19 vaccine administration.

## **Improving treatments**

Better treatments are still needed for VITT, which according to a UK study<sup>1</sup> killed 49 of the 220 people who were diagnosed with the condition between

March and June 2021. Currently, doctors treat VITT by giving anti-clotting treatments other than heparin, and administering high doses of naturally occurring antibodies from blood-plasma donors. The antibodies compete with the anti-PF4 antibodies for binding sites on platelets, and reduce the latter's ability to promote blood coagulation. "The hope is to try to confuse the body and hide the dangerous antibodies within a huge fog of normal antibodies," says Kelton. "That's a very, very blunt tool."

In Birmingham, Nicolson has been working to develop more-specific approaches. He has tested blood serum from people with VITT to see whether he can repurpose drugs developed for other conditions to treat it. In particular, he is focusing on treatments that interfere with a protein on platelets, to see whether any drugs can prevent platelet activation and the cascade of events that leads to clots in VITT.

But even if he were ready to launch a clinical trial of these therapies, there are few people in whom to test them. Since he saw the first cases in March, the United Kingdom has changed its vaccination policy, and now recommends the Oxford–AstraZeneca vaccine only for people over 40. VITT is more frequent in younger vaccine recipients, possibly because of their more-robust immune responses.

It is unclear whether other countries will have the same luxury of restricting Oxford–AstraZeneca vaccines to older people, given that it is relatively cheap and widely available compared with the mRNA vaccines, for example. Until now, VITT has primarily been reported in Europe and the United States, but researchers don't yet know whether this reflects regional differences in susceptibility to VITT, or differences in reporting systems that gather data on potential vaccine side effects. In Thailand, for instance, researchers reported in July that there had been no cases of VITT after 1.7 million doses of the Oxford–AstraZeneca vaccine were given<sup>12</sup>.

Nicolson says the number of people referred to his hospital with VITT has declined drastically: "We're not seeing it any more, it's almost stopped happening."

*doi: <https://doi.org/10.1038/d41586-021-02291-2>*

## References

1. 1.

Pavord, S. *et al.* *N. Engl. J. Med.* **384**, 2124–2130 (2021).  
<https://doi.org/10.1056/NEJMoa2109908> (2021).

2. 2.

Schultz, N. H. *et al.* *N. Engl. J. Med.* **384**, 2092–2101 (2021).

3. 3.

Greinacher, A. *et al.* *N. Engl. J. Med.* **384**, 2202–2211 (2021).

4. 4.

Scully, M. *et al.* *N. Engl. J. Med.* **384**, 2832–2839 (2021).

5. 5.

Huynh, A., Kelton, J. G., Arnold, D. M., Daka, M. & Nazy, I. *Nature* **384**, 2124–2130 (2021).  
<https://doi.org/10.1038/s41586-021-03744-4> (2021).

6. 6.

Greinacher, A. *et al.* Preprint at Research Square  
<https://doi.org/10.21203/rs.3.rs-440461/v1> (2021).

7. 7.

Othman, M., Labelle, A., Mazzetti, I., Elbatarny, H. S. & Lillicrap, D. *Blood* **109**, 2832–2839 (2007).

8. 8.

Baker, A. T. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.05.19.444882> (2021).

9. 9.

Greinacher, A. *et al.* *Blood* <https://doi.org/10.1182/blood.2021012938> (2021).

10. 10.

Kowarz, E. *et al.* Preprint at Research Square  
<https://doi.org/10.21203/rs.3.rs-558954/v1> (2021).

11. 11.

Nicolai, L. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.06.29.450356> (2021).

12. 12.

Noikongdee, P. *et al.* Preprint at medRxiv  
<https://doi.org/10.1101/2021.07.15.21260622> (2021).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02291-2>

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

# Opinion

- **[Origins of SARS-CoV-2: window is closing for key scientific studies](#)** [ 25 August 2021]  
Comment • Authors of the March WHO report into how COVID-19 emerged warn that further delay makes crucial inquiry biologically difficult.
- **[Five principles for climate-resilient cities](#)** [ 24 August 2021]  
Correspondence •
- **[First authors: is co-equal genuinely equal?](#)** [ 24 August 2021]  
Correspondence •
- **[Could Europe become the first climate-neutral continent?](#)** [ 24 August 2021]  
Correspondence •
- **[Home laboratory: interactive science from the kitchen](#)** [ 24 August 2021]  
Correspondence •

- COMMENT
- 25 August 2021

# Origins of SARS-CoV-2: window is closing for key scientific studies

Authors of the March WHO report into how COVID-19 emerged warn that further delay makes crucial inquiry biologically difficult.

- [Marion Koopmans](#) <sup>0</sup>,
- [Peter Daszak](#) <sup>1</sup>,
- [Vladimir G. Dedkov](#) <sup>2</sup>,
- [Dominic E. Dwyer](#) <sup>3</sup>,
- [Elmoubasher Farag](#) <sup>4</sup>,
- [Thea K. Fischer](#) <sup>5</sup>,
- [David T. S. Hayman](#) <sup>6</sup>,
- [Fabian Leendertz](#) <sup>7</sup>,
- [Ken Maeda](#) <sup>8</sup>,
- [Hung Nguyen-Viet](#) <sup>9</sup> &
- [John Watson](#) <sup>10</sup>



The World Health Organization assembled a team of staff and independent experts tasked with understanding the origins of SARS-CoV-2. Credit: Aly Song/Reuters/Alamy

Our group was convened by the World Health Organization (WHO) in October 2020. We have been the designated independent international members of a joint WHO–China team tasked with understanding the origins of SARS-CoV-2. Our report was published this March<sup>1</sup>. It was meant to be

the first step in a process that has stalled. Here we summarize the scientific process so far, and call for action to fast-track the follow-up scientific work required to identify how COVID-19 emerged, which we set out in this article.

The window of opportunity for conducting this crucial inquiry is closing fast: any delay will render some of the studies biologically impossible. Understanding the origins of a devastating pandemic is a global priority, grounded in science.

## The mandate

We, all the members of the international expert team, each submitted detailed, confidential statements to the WHO on potential conflicts of interest, including funding, collaborative studies, public statements and other issues around the origins of COVID-19 that could be perceived as conflicts. After the WHO had reviewed these, team members were appointed in their individual capacity, not as representatives of their employers.

So far, our mission has been guided by terms of reference agreed between the WHO and China in 2020, before our involvement<sup>1</sup>. These terms tasked us with making a detailed reconstruction of the early phase of the pandemic, beginning in Wuhan, China, where the first known cases were reported. Our mandate was to conduct a collaborative study with leading scientists in China to review data they had generated on the basis of initial questions from the WHO. We refined the generic list of questions described in the mandate into a detailed workplan described in the mission report<sup>1</sup> (see also Annex A; [go.nature.com/3k26jzx](https://go.nature.com/3k26jzx)).



## [WHO report into COVID pandemic origins zeroes in on animal markets, not labs](#)

The workplan specified eight items: specific retrospective studies detailing the profile of respiratory illness in the general community and hospitalized people in Wuhan and Hubei in the second half of 2019; a review of patient files for 76,000 cases in the same time period that had been notified by 233 Wuhan health centres; a review of death certificates and analysis of those data for possible clusters; and a detailed reconstruction of the investigation into the early outbreak, combining all data and findings from the various groups involved in human, animal and environmental studies (a One Health approach; see [go.nature.com/3jy7ekh](https://go.nature.com/3jy7ekh)). The other four items were: extensive mapping and trace-back of the supply chain of products sold at the Huanan seafood market in Wuhan; testing of a wide range of livestock, wildlife, pets and zoo animals for evidence of infection with SARS-CoV-2; analysis of published and unpublished viral genomic data and linking them with metadata for reconstruction of initial clusters; and a review of relevant literature related to the origins mission.

The possibility of a laboratory origin for the virus's introduction into the local human population — what has come to be called the lab-leak hypothesis — was not part of the WHO's original terms of reference for the team.

## The mission

This January, we undertook a 28-day mission to Wuhan to interview clinical, laboratory and public-health professionals and visit institutions involved in the early epidemic response and subsequent investigations. Our work was supported by a team of staff from the WHO China office and from WHO headquarters in Geneva, Switzerland; staff from the Food and Agriculture Organization of the United Nations (FAO) and the World Organisation for Animal Health (OIE); and a WHO-appointed team leader<sup>1</sup>. The huge burden of preparatory work was shouldered by the team in China, including more than 1,000 health-care professionals who collected, analysed, presented and discussed data and study outcomes during our joint mission.

Scientific discussions between the international and Chinese teams during this mission were lively. Large amounts of information were exchanged on the basis of the work carried out. It took days of discussion to develop recommendations on essential further work and ongoing data sharing. We drafted a model of the potential ‘pathways of emergence’ to structure our thoughts. We listed current evidence for and against these pathways (see Fig. 1 of ref. 1).

We found the laboratory origin hypothesis too important to ignore, so brought it into the discussions with our Chinese counterparts. And we included it as one of the hypotheses for SARS-CoV-2 origin in our report.



Officials collect COVID-19 test samples in a fresh market in China's Shanxi province in January.Credit: Wei Liang/China News Service/Getty

We had limited time on the ground in Wuhan and a limited mandate. So we prioritized understanding the role of labs in the early days of the epidemic, the overall lab biosafety procedures and potential staff illness or absenteeism owing to respiratory disease in the late part of 2019. We spoke to the leadership and staff at the three Wuhan labs handling coronaviruses: the Wuhan Institute of Virology, the Chinese Center for Disease Control and Prevention (CDC) in Wuhan, and the Hubei provincial CDC. We reviewed published work from these labs to assess their scientific history of working with coronaviruses related to severe acute respiratory syndrome (SARS).

The Chinese team was and still is reluctant to share raw data (for instance, on the 174 cases identified in December 2019), citing concerns over patient confidentiality. Access to data on these cases was not specified in the mandate, although the WHO had demanded it during the investigation, and has done so since . The legal and possible other barriers could not be addressed in the short time frame of our visit. Also, by then, it was clear that

the 174 cases were not likely to be the earliest ones, so we considered them less urgent for understanding origins.

It was therefore agreed that a second phase of studies would address these concerns and review these data.

## The report

In our joint report<sup>1</sup>, members of both teams concluded unanimously that there was clear evidence of widespread SARS-CoV-2 circulation in Wuhan during December 2019. We reported evidence for earlier emergence but reached no resolution on when, where and how that occurred. We concluded that the Huanan seafood market had a significant role in the early part of the pandemic, and that there were credible links to wild-animal markets to follow up. We agreed that the earliest cases of COVID-19 had probably been missed, as is common for outbreaks of new diseases<sup>2</sup>.



### [Divisive COVID ‘lab leak’ debate prompts dire warnings from researchers](#)

Our joint report summarized the evidence base that was generated during this first phase of origin tracing. It concluded that there was no definitive proof for or against any of the four proposed pathways: direct zoonotic introduction (through a spillover from wild animals) and three indirect routes of introduction (see Fig. 1 of ref. 1 ). These three are: zoonotic infection from handling infected farmed animals; zoonotic introduction

through the consumption of contaminated food or food from infected animals; or introduction through escape from a laboratory working with animal viruses. The report noted that we considered direct introduction or indirect zoonotic introduction through an intermediate host the most plausible.

As laid out in our terms of reference, this initial study was not expected to provide definitive answers to the origin of SARS-CoV-2. Rather, phase 1 was always intended to form the foundation of a longer process of scientific investigation that could last for months or years. Therefore, the report put forward recommendations for phase 2 studies that would follow the evidence and trace back further along the most likely pathways. As a joint WHO–China study report, these recommendations were agreed on by members of both the international and the Chinese team. The report also stated that this assessment could be revised if new evidence became available.

## The response

Before the report was released, formal statements to the WHO from some governments were circulated in February, with three contentions: that China had not shared data adequately; that we had paid insufficient attention to the lab-leak hypothesis; and that our scientific conclusions were influenced by China’s political stance regarding transmission through the food chain.

Since its release, our report has received extensive coverage in the popular and scientific press and on social media. Much of this has focused on how we conducted the work, and has critiqued us, our methods and results. Five months on, criticisms of the WHO–China joint study continue to emerge.

When asked, our team has emphasized that much new information was shared by the Chinese team as a result of the agreed studies, and that even more was shared as part of the iterative process between the international and Chinese teams.



A woman pushes a cart at the closed wholesale seafood market in Wuhan, China, last January. Credit: Getty

Our critics have also suggested that the report dismisses the possibility of a lab leak. A laboratory origin hypothesis is presented in the pathway model in Figure 5 on page 119 of the report; we explicitly state in the report that it is possible. We held frank discussions with key scientists in the relevant Wuhan institutions — a line of inquiry that exceeded our original mandate. When we reviewed the responses to our questions on this issue, and all other available data, we found no evidence for leads to follow up; we reported this fact.

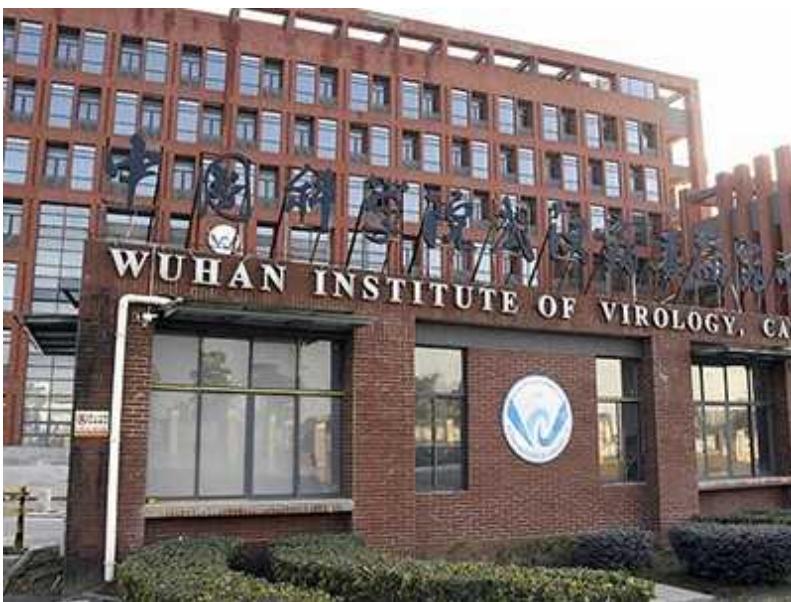
In our report, we state that if evidence supporting any of the hypotheses becomes known following publication, phase 2 studies should carefully examine this. For instance, we described that there was evidence of the presence of live animals in the market at the end of December 2019, but that the data presented to the team did not show definitive evidence of live mammals. This evidence came to light after publication<sup>3</sup> (as we discuss in more detail later in this article).

Another criticism was that the potential for introduction of SARS-CoV-2 through frozen food was included owing to pressure from China. The report addressed this hypothesis for three reasons: analysis showed that frozen food imported from all over the world was sold at the Wuhan market, including frozen wild-animal meat; foodborne viral-disease outbreaks are widely documented, including occasionally from frozen foods; and SARS-CoV-2 can remain infectious when frozen<sup>4</sup>. Therefore, the team felt it could not rule out introduction from undercooked meat from infected animals.

Some of the public discourse around the report probably originates from miscommunication and misunderstanding about the nature of the work. Although the published report correctly calls it a joint study to reflect what was laid out in the World Health Assembly resolution and terms of reference, it was publicly called an investigation by journalists, by representatives from some member states and, on occasion, by representatives of the WHO. This might have led to expectations that the report would provide watertight evidence based on formal audits of the institutes involved in the studies.

## New data

There have been calls from scientists for further investigation of the lab-leak hypothesis<sup>5</sup>. And there has been a wave of media items that give equivalence to the weight of evidence for a lab leak and for emergence through an intermediate host — an equivalence that the currently available data do not support, in our view.



### The COVID lab-leak hypothesis: what scientists do and don't know

The arguments and data for a zoonotic spillover event were summarized in a review published as a July preprint by a group of scientists who were not part of the international team<sup>6</sup>. That review includes new data released since the report, on SARS-CoV-2-related coronaviruses in bats in China's Yunnan province<sup>7,8</sup> and an inventory of live mammals for sale in Wuhan markets up until November 2019, some of which could have theoretically been able to harbour SARS-related coronaviruses<sup>3</sup>. This inventory, compiled by scientists from the United Kingdom, Canada and China, would have been welcomed by the team had it been available earlier; it needs to be taken up in the phase 2 studies.

In June, a preprint<sup>9</sup> was published analysing genomic data that had been deleted after March 2020 from the database of the US National Center for Biotechnology Information at the request of the scientists from China who generated the information (that team had published its findings based on the raw data in June 2020<sup>10</sup>). Our colleagues in China contacted the authors of the June 2020 paper, retrieved the data and added them to the SARS-CoV-2 genome phylogenetic data published in our report. The data were from people who had an onset of illness in January, so they did not contribute any new information to the origins question.

In the report, and since, we have publicly called for any data supporting the lab-leak hypothesis to be published and submitted to the WHO. None has, so far.

## Six priorities

To keep up the momentum for phase 2 studies, our team has met weekly since the publication of the joint report. We have continued collaboration with our Chinese co-authors, including work on a list of corrections to the phase 1 report. Both the international team and the Chinese team have now put forward to the WHO priorities for phase 2 studies, developed from the recommendations in the joint report.

The international team listed the following priorities:

**Further trace-back studies.** On the basis of disease reporting, look for early COVID-19 cases in all regions inside and outside China that have the earliest evidence for SARS-CoV-2 circulation.

**Antibody surveys.** Use standardized methods in the regions that have the earliest evidence for SARS-CoV-2 circulation (inside and outside China) to identify any places where infections occurred that were not observed through disease reporting.

**Trace-back and community surveys.** These will need to be conducted at sites of wildlife farms that supplied animals to markets in Wuhan in the months before human cases were recognized (inside and outside China, depending on supply-chain analysis).

**Risk-targeted surveys of possible hosts.** Assess wild bats and other potential reservoirs or intermediate hosts in China and neighbouring countries, and selected high-risk farmed animals (including those farmed for fur), for evidence of exposure.

**Detailed risk-factor analysis.** Analyse pockets of earlier cases evidenced from the antibody surveys or other studies, and conduct an assessment of all possible exposures.

**Follow-up.** Investigate any credible new leads.

## Time's up

The search for the origins of SARS-CoV-2 is at a critical juncture. There is willingness to move forward from both the WHO international team and the Chinese team.

Crucially, the window is rapidly closing on the biological feasibility of conducting the critical trace-back of people and animals inside and outside China. SARS-CoV-2 antibodies wane, so collecting further samples and testing people who might have been exposed before December 2019 will yield diminishing returns. Chinese wildlife farms employ millions of people (14 million, according to a 2016 census<sup>11</sup>) and supplied live mammals to cities across China, including Wuhan<sup>3</sup>. In response to the SARS-CoV-2 pandemic, many of these farms are now closed and the animals have been culled, making any evidence of early coronavirus spillover increasingly difficult to find.

In July, four months after the full report and five months after our debriefing, the WHO informed member states of plans to create a committee that will oversee future origins studies. We are pleased to see both this and its implication that outbreak investigations will be conducted routinely, rather than in an ad hoc manner that could be perceived as politically motivated or with potentially punitive goals.

However, applying this new process to the continuing SARS-CoV-2 origins mission runs the risk of adding several months of delay. Member-state representatives would need to negotiate detailed terms around the sensitive issue of investigating laboratory practices, then nominate and select team members, who would then have to develop a work plan.

Therefore, we call on the scientific community and country leaders to join forces to expedite the phase 2 studies detailed here, while there is still time.

*Nature* **596**, 482-485 (2021)

doi: <https://doi.org/10.1038/d41586-021-02263-6>

# References

1. 1.  
World Health Organization. *WHO-Convened Global Study of Origins of SARS-CoV-2: China Part. Joint WHO–China Study* (WHO, 2021).
2. 2.  
Anderson, R. M. *et al.* *Phil. Trans. R. Soc. Lond. B* **359**, 1091–1105 (2004).
3. 3.  
Xiao, X., Newman, C., Buesching, C. D., Macdonald, D. W. & Zhou, Z.-M. *Sci. Rep.* **11**, 11898 (2021).
4. 4.  
Huang, S. Y. *et al.* *mSphere* **6**, e00104-21 (2021).
5. 5.  
Bloom, J. D. *et al.* *Science* **372**, 694–694 (2021).
6. 6.  
Holmes, E. C. *et al.* Preprint at Zenodo  
<https://doi.org/10.5281/zenodo.5075888> (2021).
7. 7.  
Li, L.-L. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.03.17.435823> (2021).
8. 8.  
Zhou, H. *et al.* *Cell* <https://doi.org/10.1016/j.cell.2021.06.008> (2021).

9. 9.

Bloom, J. D. Preprint at bioRxiv  
<https://doi.org/10.1101/2021.06.18.449051> (2021).

10. 10.

Wang, M. *et al.* *Small* **16**, 2002169 (2020).

11. 11.

Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. *IPBES Workshop on Biodiversity and Pandemics: Workshop Report*. (IPBES, 2020).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02263-6>

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

- CORRESPONDENCE
- 24 August 2021

# Five principles for climate-resilient cities

- [Christian Albert](#)<sup>0</sup>,
- [Samuel Rufat](#)<sup>1</sup> &
- [Christian Kuhlicke](#)<sup>2</sup>

The recent catastrophic floods in China, Belgium and Germany underscore the importance of building climate-resilient cities. We represent a consortium that has proposed a comprehensive climate-adaptation programme to help guide urban decision-making and governance to better prepare for future climate-related events (see [go.nature.com/3ctpwau](https://go.nature.com/3ctpwau)).

Such efforts hinge on five principles learnt from earlier disasters. The world must improve early-warning systems and strengthen flood barriers and civil protection, in particular for smaller watersheds; develop ‘sponge’ cities and landscapes that harness nature-based flood-risk mitigation; carry out climate-risk assessments of crucial infrastructure such as hospitals, transport and freshwater supply; climate-proof exposed buildings; and reinforce action with strong political cooperation and solidarity, particularly with the most vulnerable groups of people.

We have the scientific knowledge to develop climate-resilient cities (see, for example, [X. Bai et al. \*Nature\* 555, 23–25; 2018](#)). However, political will and bold decisions will be essential for implementing the solutions and the necessary societal transformations.

*Nature* **596**, 486 (2021)

doi: <https://doi.org/10.1038/d41586-021-02309-9>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02309-9>

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

- CORRESPONDENCE
- 24 August 2021

# First authors: is co-equal genuinely equal?

- [Jonathan Kipnis](#) 0

‘Equal’ distribution of co-first authors on research papers should be a win-win concept — not just for those authors, but also for multi-disciplinary science. Yet some seek to reshuffle their respective positions on CVs for career purposes. How can we ensure that co-equal means genuinely equal?

Principal investigators and trainees must use the term responsibly, with endorsement by all of the project’s participants. Besides trainees who work side-by-side on a shared project, co-first authorship might be justified if one trainee supplies information that strengthens a crucial conclusion or spends months revising a paper abandoned by another trainee, for example.

My recent tweet (see [go.nature.com/3sek93o](https://go.nature.com/3sek93o)) prompted suggestions for improving recognition of co-first authorship, such as by using an expanded citation format of ‘X, Y, Z *et al.*’, and by highlighting each as a first author with EndNotes in papers and on PubMed or Google Scholar. This would make it easier for faculty members to recognize participants’ equal contributions when evaluating them for promotion and tenure, irrespective of name order.

*Nature* **596**, 486 (2021)

doi: <https://doi.org/10.1038/d41586-021-02310-2>

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

- CORRESPONDENCE
- 24 August 2021

# Could Europe become the first climate-neutral continent?

- [Nebojsa Nakicenovic](#) <sup>0</sup> &
- [Peter D. Lund](#) <sup>1</sup>

Two scientific reports released in June indicate what it would take for Europe to become the first climate-neutral continent. As the lead author of the Group of Chief Scientific Advisors to the European Commission and chair of the energy project of Science Advice for Policy by European Academies, respectively, we contributed to these reports, advising policymakers on how to accelerate the transition from fossil fuels to renewable energy (see [go.nature.com/3k5f6i7](https://go.nature.com/3k5f6i7) and [go.nature.com/3yw4pjg](https://go.nature.com/3yw4pjg)).

Central to the transition is immediate adoption of innovative technologies that are backed by regulatory and market measures, along with social and behavioural changes to incentivize and support low-carbon energy choices. Investment must be stepped up to hasten development of flexible, efficient and resilient energy systems that rely on electrification and hydrogen.

A coordinated combination of policies, measures and instruments, including carbon pricing as a driving force, will shape an effective, consistent and just regulatory system. For example, this could extend current emissions-trading arrangements and introduce a border carbon-tax adjustment.

*Nature* **596**, 486 (2021)

doi: <https://doi.org/10.1038/d41586-021-02311-1>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02311-1>

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

- CORRESPONDENCE
- 24 August 2021

# Home laboratory: interactive science from the kitchen

- [Poorti Kathpalia](#)<sup>0</sup>,
- [Arpita Konar](#)<sup>1</sup> &
- [Beena Pillai](#)<sup>2</sup>

The COVID-19 pandemic has compromised laboratory training and scientists' access to the latest experimental tools. Digital technology has helped hugely, but it cannot offer the essential hands-on learning experience of a lab. This shortfall in experimental know-how is hindering scientists from generating and using new data, particularly in disadvantaged countries.

Our own training in the latest advances in RNA biology by international experts at a 2020 workshop was brought to a halt as India hurtled towards two successive COVID-19 waves. Our monthly outreach events for schoolchildren also stalled.

To provide these children with an immersive research experience, even though they were confined to their homes, we transformed our lab into a live online classroom. Teaching staff were masked up and safely distanced. The children watched on their digital devices as we isolated DNA, and they reproduced the protocol in their kitchens using reagents from a kit that we mailed to them in advance. They discussed the methods and scientific principles with our graduate students at the same time.

The challenges of these interactive classes offered these schoolchildren a stimulating and fulfilling learning experience.

*Nature* **596**, 486 (2021)

doi: <https://doi.org/10.1038/d41586-021-02278-z>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02278-z>

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

# Work

- **[Three questions to address rigour and reproducibility concerns in your grant proposal](#)** [ 19 August 2021]  
Career Column • Addressing weaknesses and limitations in your science will reassure potential funders, say grant-writing coaches Jennifer L. Wilson and Crystal M. Botham.
- **[Innovative tools take aim at antibiotic-resistant microbes](#)** [ 23 August 2021]  
Technology Feature • Diagnostics that rely on bacterial movements, genomics and machine learning could help to address a global crisis.
- **[Making sense of quantum-level chaos](#)** [ 23 August 2021]  
Where I Work • Fabio Deelan Cunden draws inspiration for his studies of randomness from ancient books and artefacts in a mathematics museum.

- CAREER COLUMN
- 19 August 2021

# Three questions to address rigour and reproducibility concerns in your grant proposal

Addressing weaknesses and limitations in your science will reassure potential funders, say grant-writing coaches Jennifer L. Wilson and Crystal M. Botham.

- [Jennifer L. Wilson](#)<sup>0</sup> &
- [Crystal M. Botham](#)<sup>1</sup>



Credit: Getty

Since 2018, the US National Institutes of Health (NIH) has [required that research proposals explicitly describe scientific rigour](#). They want to know scientists' approaches to ensuring the fidelity of their data, minimizing bias and maximizing new knowledge. The NIH did this to address transparency and reproducibility challenges in research. Other funders have also signalled a long-term commitment to better practices in science. In 2020, for example, the European Commission's Directorate-General for Research and Innovation, the body responsible for the European Union's research and innovation policy, issued guidance to improve research reproducibility. Writing about rigorous choices can be simple. As grant coaches at the Grant Writing Academy at Stanford University in California, we work with scientists who have already thought carefully about the quality of their science and want to address rigour in their grants. Our process focuses on asking simple questions to help them sufficiently justify the rigour in their research proposals.



### [Collection: Careers toolkit](#)

For example, one recent proposal included a sentence like this: “We will use the stroke mouse model and treat the mice with our novel compound daily, for three weeks.”

As grant coaches without specific knowledge of the model, we probed the writer with questions looking for more detail, such as “Why this mouse model?” and “Why daily dosing?” As it turns out, the mouse model and dosing regimen were standard in that laboratory, and the lab had published a study that showed the model was relevant to the stroke outcome of interest.

We pushed the writer to be more specific in their writing, to communicate how their approaches minimize bias due to the choice of model organism: ‘We will use the stroke model mouse, a model system we and others have shown to be relevant to understanding stroke outcomes [references], to test the effect of novel therapeutic compounds.’

This sentence is more successful because it clearly shows how the writer’s choices are adequate for addressing their research question, and how this choice will help them to derive knowledge about the disease condition.

## **Three-question framework**

Our three-question framework helps to frame scientific choices when writing research proposals. Our process guides writers to explain their experimental choices, by asking questions that address limitations in their proposal:

1. What are the essential weaknesses or limitations in your science? Every scientific method has limits. Often, these limits are methodological or field-specific.
2. Which methods will you use, or are you already using, to address these limitations? These may be standard methods in your research group, but it’s important to highlight them as such for a reviewer.
3. What makes these methods adequate? Justify your choices by confirming how the field has tested or accepted these methods for overcoming limitations.

Revisiting our mouse-model example, one limitation of the approach is whether testing the compounds in the stroke model sufficiently replicated human stroke progression (this addresses question 1). Using the stroke-

appropriate mouse model addressed this weakness (and answers question 2), and referencing peer-reviewed publications justified the writer's selection of this model (which answers question 3). The revision justifies the animal model's relevance to the proposed research and references published data that support that statement.

Through our grant coaching, we have found that addressing scientific rigour often requires careful and specific wording: instead of, "We will use our new method to anticipate drug effects," we would guide a writer to "We will calibrate our new method using a landmark dataset, a gold-standard comparison in our field, to benchmark against known effects before anticipating new drug effects." A new method could bias results, but benchmarking the method against a well-regarded dataset of known effects justifies the method's adequacy for understanding effects of a new drug.



Jennifer Wilson and Crystal Botham developed a framework for writing grant proposals. Credit: Jennifer Brophy

In another scenario, the sentence, “We will assess treatment effects by comparing wound healing of the untreated left leg and treated right leg,” needs an introductory clause: “We have previously shown that wound-healing rates are different between individual animals, making it difficult to compare between them [references]. Thus, we will generate two wounds on the same animal and apply treatment to only one wounded area.” This explanation helps a reviewer to appreciate the writer’s ability to assess the treatment’s effect, and to prevent the effect from bias due to differences in wound-healing ability of individual animals.

Addressing rigour in research proposals is often less about changing scientific choices or overall project design, and more about justifying how experimental-design choices address limitations that could prevent the researcher from answering their question or diminish the knowledge they derive from their experiment.



### Collection: Funding

Justifying scientific choices requires deliberate practice to achieve strong, persuasive writing. The grant writer must be aware of and unafraid to share the limitations to their science. The selection of the mouse model is a decision that supports rigour, but someone not familiar with the lab’s research, such as a grant reviewer, might not understand this unless the writer specifically justifies it. In the case of the stroke model, the selection

of the mouse model is a decision that improves quality, but that might have been taken for granted by the grant writers because everyone in the lab in this example uses this model. Understanding why the lab uses this stroke model would have enabled this writer to better justify their choice to a reviewer who isn't familiar with the lab's best practices.

Often, it's difficult to decide which choices require justification. We recommend examining published work and talks to understand limitations and how they were addressed. It might feel daunting to address all possible limitations to a research approach, so start by investigating journal publication requirements for reproducibility and transparency. Many journals have specific requirements about the reporting of protocols, the use of biological samples, the availability of analysis code and other technical details. Journal requirements are designed to overcome field-specific challenges to reproducibility and transparency. Take note of how others justify their choices — you don't have to go as far as explaining whether an experiment needs a control (almost all good experiments have at least one control), but you will probably need to justify how a particular control is well-suited to your research question.

Gaining this awareness is not an overnight process. Learning the limitations of a scientific field is an ever-evolving process, and assessing how others justify their rigorous decisions will deepen your understanding of how to make them. Our framework provides one way to prioritize and address hurdles to reproducible and transparent science. We encourage writers to have peers outside their research group read their work and use the three-question framework. Often, an outsider can provide a fresh perspective on the choices that a writer has taken for granted or failed to explicitly address.

We feel that developing awareness of rigorous practices during the writing of research proposals will elevate rigorous thinking throughout the research enterprise.

*Nature* **596**, 609-610 (2021)

doi: <https://doi.org/10.1038/d41586-021-02286-z>

This is an article from the Nature Careers Community, a place for Nature readers to share their professional experiences and advice. [Guest posts are encouraged.](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02286-z>

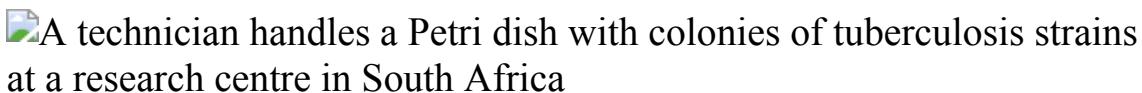
| [Section menu](#) | [Main menu](#) |

- TECHNOLOGY FEATURE
- 23 August 2021

# Innovative tools take aim at antibiotic-resistant microbes

Diagnostics that rely on bacterial movements, genomics and machine learning could help to address a global crisis.

- [Jyoti Madhusoodanan](#) <sup>0</sup>



Bacterial resistance to antibiotics is a major problem around the world. New solutions are hoping to change that. Credit: Joao Silva/NYT/Redux/eyevine

Maha Farhat spent months in 2007 tending to patients at a hospital in Durban, South Africa. Many were infected with HIV. But the infection that preyed on the then-medical-resident's mind, and her patients', was caused not by a virus, but by a bacterium: *Mycobacterium tuberculosis*, the pathogen that causes tuberculosis. In particular, she was concerned about strains that are resistant to common antibiotics.

Although immunocompromised individuals are especially susceptible to tuberculosis, the infection isn't unique to people with HIV: *M. tuberculosis* claimed 1.4 million lives worldwide in 2019, 208,000 of whom had HIV. "Tuberculosis was briefly superseded by COVID-19, but it's still the top infectious-disease killer globally," says Farhat, now a physician and bioinformatician at Harvard University in Boston, Massachusetts. Drug-resistant forms of tuberculosis are a major contributor to the problem.

Drug-resistant pathogens of all types have precipitated an antibiotic-resistance crisis that threatens public health, agriculture, animal husbandry

and more. But spotting these strains and identifying effective treatments is tricky. Labs equipped to handle especially infectious pathogens, such as *M. tuberculosis*, can be hard to come by in resource-limited countries, and instruments for testing for drug sensitivity can take days to return results. In many cases, physicians test for resistance only after one or more standard antibiotics fail. While waiting, patients might begin an unnecessary or ineffective course of antibiotics, or leave the clinic without treatment.

Farhat and other researchers are turning to tools such as atomic force microscopy, genomics and machine learning to create point-of-care diagnostic tests that they hope will provide results in minutes, minimizing the use of incorrect or unnecessary prescriptions. “An increase in rapidity is the most important advance needed,” says clinical microbiologist Evgeny Idelevich at the University Medical Center Greifswald in Germany.

## Gauging growth

The gold-standard method for assessing drug susceptibility of microbes, known as a disk diffusion test, dates to 1889. Researchers culture bacteria on an agar plate, then place tiny paper disks loaded with drugs on the growing cells; zones around the disks become transparent if a drug kills bacteria or stalls their growth, indicating that the microbes are susceptible to the medication.

Companies have automated that same principle in antimicrobial sensitivity testing instruments, such as the BD Phoenix from BD Biosciences, headquartered in New Jersey, and the VITEK 2 from bioMérieux, based in Marcy-l’Étoile, France. These systems seed bacteria in liquid cultures with antibiotics and look for optical changes that indicate bacterial growth or death. The tests typically require somewhere between 4 and 8 hours, although results can take a day or more to arrive because clinicians must send samples to clinical microbiology labs<sup>1</sup>.

But researchers are also exploiting assays that are more commonly associated with the physical sciences than with microbiology labs.

In 2018, for instance, Idelevich devised a miniaturized version of the liquid culture test that relies on MALDI–TOF, a mass-spectrometry technique that uses laser-induced ionization and a long ‘flight tube’, through which ions travel, to identify molecules on the basis of their mass and charge. Idelevich and his colleagues placed microdroplets of cultures of two pathogens — *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* — directly on a solid support matrix used for MALDI–TOF and incubated each droplet with a different drug. They then processed the sample with a system specifically designed for bacterial identification: the MALDI Biotyper from Bruker Daltonik. Intensities of characteristic spectral peaks indicated whether the cultures were susceptible or resistant to the antibiotic<sup>2</sup>.



### Tuberculosis genomes track human history

In 2013, Giovanni Longo, at the National Research Council of Italy in Rome, and his colleagues found that when they bound pathogenic *Escherichia coli* to miniature diving-board-like structures called cantilevers and exposed them to antibiotics, the cantilever bobbed up and down because of small movements of attached, living bacteria. The movements ceased if the microbes were susceptible to the antibiotics. The movement was visible under an atomic force microscope within minutes — long before microbes replicated — meaning the test can identify live bacteria far faster than is possible with an assay that looks for bacterial growth<sup>3</sup>.

Rachel McKendry, a nanotechnology researcher at University College London, and her then-graduate-student Isabel Bennett wanted to take that approach into the clinic. But attaching the bacteria to cantilevers that were 200 micrometres long in a Petri dish was easier said than done. “Only a fraction of the cantilevers would have bacteria attached, and sometimes they’d be either clumpy or too few attached,” Bennett says.

As she worked with Longo’s team to fine-tune the process, Bennett detected large differences in reflected light, which suggested that similar bacterial movements could be spotted even when the microbes were not tethered to the cantilevers. So, the team switched tactics: they altered the set-up to track bacteria as they floated across the structures’ surfaces. They made the cantilevers from a hard, reflective material, and developed software to analyse bacterial movement so that the readout was proportional to the number of bacteria in solution<sup>4</sup>. “This deceptively simple signal turned out to be really a nice way to detect resistance compared to current methods,” McKendry says.

Although not yet commercialized, the system could be adapted and scaled up for clinical use, Bennett suggests. The reflective surfaces could be turned into inserts placed in routinely used microtitre plates, and the atomic force microscope replaced with a DVD player’s optical reader to capture the signal. “It could potentially be a very easy, low-cost set-up,” she says.

Physicist Kamil Ekinci, at Boston University in Massachusetts, is pursuing another proxy for bacterial antibiotic resistance: electrical current. His team placed a urine sample spiked with *K. pneumoniae*, a common cause of urinary-tract infections, directly into a single channel of a microfluidic device with an antibiotic, and tracked electrical conductance through the channel<sup>5</sup>. “If the bacteria grow and clog the channel, they create more electrical resistance,” Ekinci says. “We’re basically transducing the bacterial growth into an electrical signal.”

The advantage, Ekinci adds, is that an electrical signal is easier to amplify and visualize than are microscopy images. “In principle, our technique can detect a single bacterial division,” he says — although he adds that the method might not work for all bacteria, particularly slow-growing pathogens such as *M. tuberculosis*.

## Measuring molecular markers

Tests based on bacterial growth are easy, cheap and non-specific: a single test works across a wide range of pathogens. But because test results depend on growth conditions and using the right concentration of antibiotics, “everything else is a disadvantage”, says Susanne Häussler, who studies medical microbiology at Rigshospitalet in Copenhagen.

As an alternative, Häussler and others are turning to genomics for clues to drug resistance. This ‘culture-independent testing’ is the next big shift in the field, says epidemiologist Sophia Koo at Harvard University.

Relying on genes that are clearly linked to antibiotic-resistance mechanisms is an ideal route to a quicker diagnostic because it doesn’t require lengthy periods of bacterial incubation, says infectious-diseases researcher Gary Schoolnik at Stanford University in California. But it’s important to know which sequences in the bacterial genome are important for drug resistance, says Thomas Grys, a clinical microbiologist at Mayo Clinic in Phoenix, Arizona. “If you don’t, you could easily miss a new mechanism or detect a fragment of a gene that’s not actually conferring resistance.”



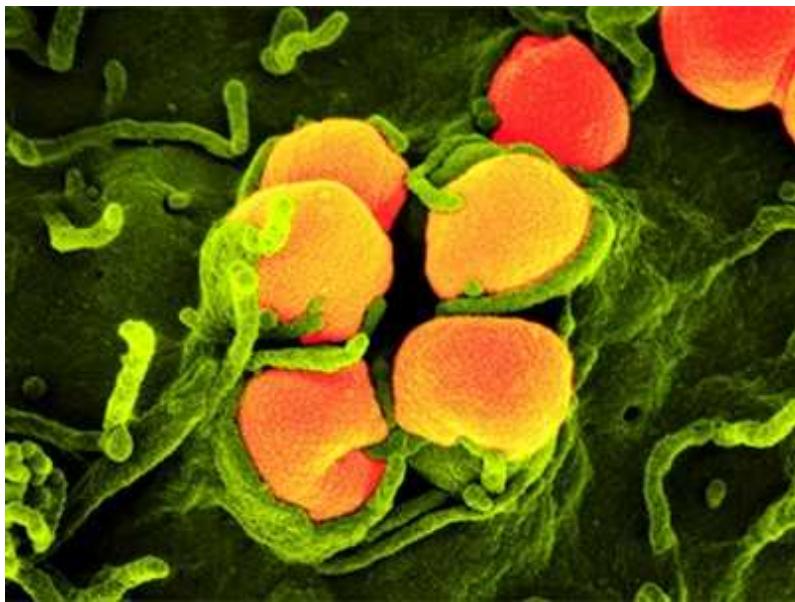
A disk diffusion test is a common way to uncover how effective an antibiotic is. Credit: Doncaster and Bassetlaw Hospitals/SPL

Schoolnik is also chief medical officer at Visby Medical, a California-based start-up that won US\$19 million in 2020 as part of an Antimicrobial Resistance Diagnostic challenge sponsored by the US National Institutes of Health and the US Department of Health and Human Services' Biomedical Advanced Research and Development Authority.

The company's test is a single-use, point-of-care diagnostic, run using a simple hand-held device, to spot drug resistance in sexually transmitted pathogens such as *Neisseria gonorrhoeae*. The assay focuses on mutations that confer resistance to ciprofloxacin, a commonly used oral antibiotic for gonococcal infections. Mutations in the gene encoding the enzyme gyrase A spell the difference between *Neisseria* strains that are resistant or susceptible to ciprofloxacin.

PCR-based tests to detect such variants are of limited use in clinics because of the need for instruments, reagents and technicians who are trained to perform reactions. Visby's diagnostic bypasses these constraints by reducing

the assay to a simple colour change. Amplified fragments flow into a chamber on the device that contains capture probes for each variant of the gene. The binding results in a colour change that reflects whether a strain is sensitive or resistant to ciprofloxacin<sup>6</sup>.



### Untreatable gonorrhoea on the rise worldwide

Others continue to explore whole-genome sequencing to capture the spectrum of variants that confer resistance. But developing low-cost, speedy tests based on such information remains a challenge. “It’s not just about the presence of a resistance gene, but also its expression,” says Nicole Wheeler, a data scientist at the University of Birmingham, UK, who studies machine-learning approaches to genomics. “The more transcriptomic and proteomic data we collect, the more we stand a chance of improving our ability to predict resistance,” Wheeler says.

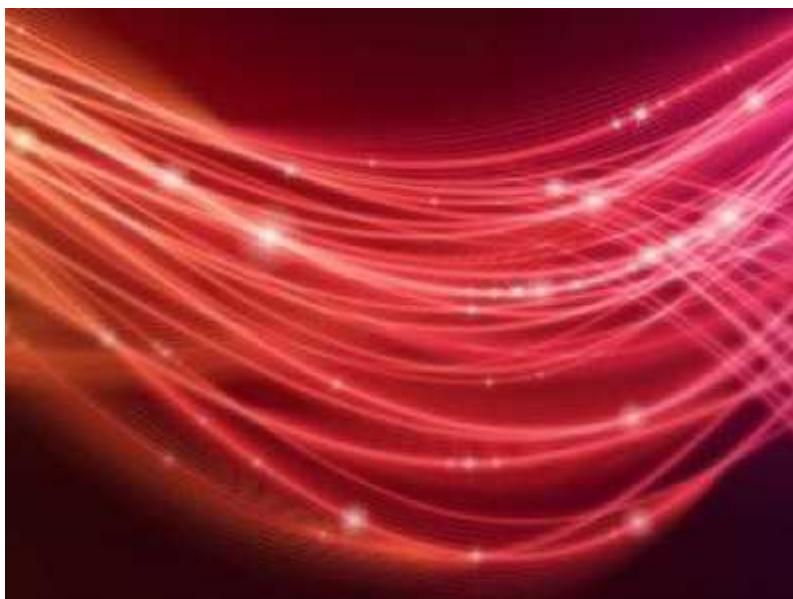
Techniques based solely on genome sequencing work well for some pathogens, such as *Salmonella enterica*, but mutations in multiple regulatory genes can alter gene-expression patterns (and thus resistance) in others, including *P. aeruginosa*. “In principle, all transcriptome data is in the genome,” Häussler says. “But it’s sometimes easier to look at the transcriptome instead of looking for all the possible mutations that alter gene expression.”

In 2014, for instance, Chikara Furusawa, a bioengineer at RIKEN in Osaka, Japan, studied lab strains of *E. coli* adapting to growth in the presence of different antibiotics, and found that he could use changes in gene expression to predict resistance more accurately than he could with genomic DNA sequences themselves<sup>7</sup>. “The correlation between gene expression and resistance was significantly higher than that between resistance and genomic markers,” Furusawa says.

## Forecasting future resistance

In their work, Häussler and her colleagues honed in on a mix of genomic and transcriptomic markers as the best ‘signature’ to predict antibiotic resistance in *P. aeruginosa*<sup>8</sup>.

But to improve their model, they turned to machine learning. Instead of simply identifying resistance-conferring mutations, they used their algorithms to identify a signature of DNA and RNA variations that predicted a strain’s resistance to antibiotics<sup>8</sup>. The algorithm helps to identify only key traits — it won’t be a part of an eventual diagnostic test, Häussler says. Still, such approaches can overcome the problem of capturing all ‘bug–drug’ combinations through the genome alone, Wheeler says.



Rather than simply informing clinicians of a pathogen's current resistance profile, these algorithms could also reveal which antibiotic-resistance mechanisms a strain might develop in response to treatments. Still, "deciding whether an algorithm should be trusted or not is challenging", Wheeler says. "They're black boxes. Even if you have all the code and all the data, you don't necessarily know what's driving the model to say sample A is resistant to azithromycin, for instance."

Another problem developers are working to overcome is overfitting, Wheeler says: an algorithm might "memorize a whole lot of unimportant features" in data, rather than learn to find true correlations. Because bacterial gene sequences can be very similar, machine-learning tools might oversimplify a problem and draw the wrong conclusions. Wheeler likens the problem to a flawed image search: an algorithm that is trained on many pictures of farm animals in fields might identify a photograph of an open space as a sheep. Bacteria frequently pass antibiotic-resistance genes around to each other on small circular chunks of DNA that aren't part of their genomes. "But because the rest of the genome is the same, the algorithm might say the strain is still sensitive," she says. "What we really want from these models is for them to learn the biology of resistance."

Given the constraints of studying and testing tuberculosis, Farhat adopted a machine-learning approach that uses whole genome sequences to make predictions. In April, she and her colleagues described a web-based tool called GenTB that can predict resistance to several tuberculosis drugs<sup>9</sup>. The tool's performance varies with the quality of input sequence data and the drug in question. Whereas one common mutation is responsible for up to 80% of resistance to the first-line drugs used for TB, several rare variants confer small boosts in resistance to second-line drugs, Farhat says. "Sometimes, you only see the resistance when several such mutations are present."

## Work in progress

Whichever approach they use, researchers face the same fundamental challenge: to design a diagnostic that improves significantly on current devices. Current tests can already return results to clinicians in less than 24

hours for a dollar or two per test, Grys says. “The question is not whether a test is good,” he says. “The question is: is it better than what we have right now? It’s important to set a trajectory that helps us meet the goals.”

Some tests in development are restricted in the kinds of sample they can process, or the bacteria or antibiotics they can test. Visby’s diagnostic is currently limited to gonococcal infections, for instance, and Ekinci’s microfluidic device requires urine samples and cannot handle infections caused by more than one species of bacterium. Others that require advanced microscopes or spectrometry, such as cantilevers, will need to be adapted before they can be used by non-specialists working in resource-poor clinics around the world. Because many of these approaches test one or two microbes against a handful of drugs, they’ve conquered only “the tip of the iceberg”, says Alex van Belkum, director of microbiology research at bioMérieux. “There’s still a big lag between these technologies and the automated antibiotic susceptibility testing systems currently in laboratories.”

As in the COVID-19 pandemic — during which rapid tests have proved crucial to detecting and stopping the spread of the virus — low-cost, point-of-care diagnostics are essential in reducing the misuse of antibiotics, says McKendry. “Antimicrobial resistance is a very complex problem, and new tests are only one part of the solution.”

*Nature* **596**, 611-613 (2021)

doi: <https://doi.org/10.1038/d41586-021-02292-1>

## References

1. 1.

van Belkum, A. *et al.* *Nature Rev. Microbiol.* **18**, 299–311 (2020).

2. 2.

Idelevich, E. A., Sparbier, K., Kostrzewska, M. & Becker, K. *Clin. Microbiol. Infect.* **24**, 738–743 (2018).

3. 3.

Longo, G. *et al.* *Nature Nanotechnol.* **8**, 522–526 (2013).

4. 4.

Bennett, I., Pyne, A. L. B. & McKendry, R. A. *ACS Sens.* **5**, 3133–3139 (2020).

5. 5.

Yang, Y., Gupta, K. & Ekinci, K. L. *Proc. Natl Acad. Sci. USA* **117**, 10639–10644 (2020).

6. 6.

Morris, S. R. *et al.* *Lancet Infect. Dis.* **21**, 668–676 (2021).

7. 7.

Suzuki, S., Horinouchi, T. & Furusawa, C. *Nature Commun.* **5**, 5792 (2014).

8. 8.

Khaledi, A. *et al.* *EMBO Mol. Med.* **12**, e10264 (2020).

9. 9.

Gröschel, M. I. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.03.27.437319> (2021).

---

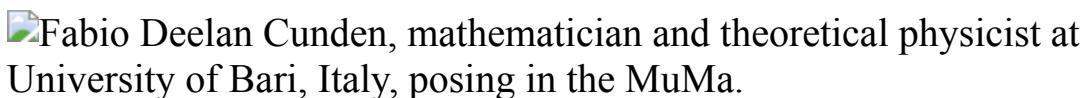
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02292-1>

- WHERE I WORK
- 23 August 2021

# Making sense of quantum-level chaos

Fabio Deelan Cunden draws inspiration for his studies of randomness from ancient books and artefacts in a mathematics museum.

- [Amber Dance](#) <sup>0</sup>



Fabio Deelan Cunden is a mathematician and theoretical physicist at the University of Bari, Italy. Credit: Cosimo Sanitate for *Nature*

The blue ‘string art’ here is an example of a ruled surface, a complex shape that you can generate by moving a simple straight line. The white object is a Clebsch surface, another complex surface that is based on simple equations. On the lower left are red dice of different shapes, which I selected because they provide a primitive example of what I study at the University of Bari, Italy — probability and randomness.

Specifically, I’m interested in chaotic systems. With one of these, if you pick two points that are very close, they will diverge once the system evolves. To understand this, first think of a billiards table. This is a regular, nonchaotic system. If you hit two balls at similar angles from similar starting points, you can predict that they’ll end up near the same pocket. But if you add an obstacle in the middle, such as a pint glass, the system becomes chaotic and you can’t predict the balls’ paths.

I study chaos at the quantum level, where atoms and subatomic particles interact. At this scale, our ‘billiards table’ is called a quantum dot. Imagine it as a sort of box, where some electrons behave like the billiard balls. I seek to understand the chaotic motion of these electrons using random matrix theory, a type of mathematics. I’d also like to apply this approach to describe aspects of the folding of proteins and of a form of artificial intelligence called machine learning.

I often find I need some social interaction for inspiration, such as conferences or coffee with other researchers. The pandemic has made this impossible, but since I started here last December I’ve found an alternative in MuMa, the university’s Museum of Mathematics, where I’m sitting, just upstairs from my office. It’s filled with mathematical objects and ancient books by giants such as the astronomer and physicist Galileo Galilei. When I’m stuck, I head to MuMa and see others’ ideas. It makes me feel part of a chain going all the way back to Galileo.

*Nature* **596**, 616 (2021)

doi: <https://doi.org/10.1038/d41586-021-02293-0>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02293-0>

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

# Research

- **[Protein-structure prediction revolutionized](#)** [ 23 August 2021]  
News & Views • The full might of a world-leading artificial-intelligence laboratory has been brought to bear on protein-structure prediction. The resulting method, AlphaFold2, promises to transform our understanding of proteins.
- **[African tropical montane forests store more carbon than was thought](#)** [ 25 August 2021]  
News & Views • The inaccessibility of African montane forests has hindered efforts to quantify the carbon stored by these ecosystems. A remarkable survey fills this knowledge gap, and highlights the need to preserve such forests.
- **[Single proton cooled by distant ions](#)** [ 25 August 2021]  
News & Views • Laser-cooled ions have been used to substantially lower the temperature of a proton located several centimetres away. This technique could be useful in ultraprecise measurements of the properties of antimatter particles.
- **[A persistent look at how tumours evade therapy](#)** [ 11 August 2021]  
News & Views • Understanding how resistance to chemotherapy occurs could lead to better anticancer treatments. Persister cells in tumours can contribute to this resistance. A method to characterize these cells in detail sheds light on their origins.
- **[A bridge across the democracy–expertise divide](#)** [ 04 August 2021]  
News & Views • An innovative algorithm provides a way of fairly selecting representative individuals for citizens' assemblies to learn about and deliberate on certain topics. Such groups hold promise for closing the gap between democracy and expertise.
- **[After the pandemic: perspectives on the future trajectory of COVID-19](#)** [ 08 July 2021]  
Perspective • This Perspective discusses possible future patterns of SARS-CoV-2 infection, the development of variants, potential changes in the patterns of spread and the implications for vaccine deployment and the potential consequences of these issues for the development of policy.
- **[Chromatic periodic activity down to 120 megahertz in a fast radio burst](#)** [ 25 August 2021]  
Article • The fast radio burst FRB 20180916B repeats with a periodicity of 16 days, and is now found to emit down to a frequency of 120 MHz, much lower than previously observed.

- [Universal pair polaritons in a strongly interacting Fermi gas](#) [ 25 August 2021]
 

Article • Directly coupling cavity photons to the photo-association resonances of pairs of atoms in a strongly interacting Fermi gas generates pair polaritons—hybrid excitations coherently mixing photons, atom pairs and molecules.
- [Sympathetic cooling of a trapped proton mediated by an LC circuit](#) [ 25 August 2021]
 

Article • A single electromagnetically trapped proton is sympathetically cooled to below ambient temperature by coupling it through a superconducting LC circuit to a laser-cooled cloud of Be<sup>+</sup> ions stored in a spatially separated trap.
- [Single-crystal, large-area, fold-free monolayer graphene](#) [ 25 August 2021]
 

Article • Restricting the initial growth temperatures used for chemical vapour deposition of graphene on metal foils produces optimum conditions for growing large areas of fold-free, single-crystal graphene.
- [Rechargeable Na/Cl<sub>2</sub> and Li/Cl<sub>2</sub> batteries](#) [ 25 August 2021]
 

Article • Rechargeable Na/Cl<sub>2</sub> and Li/Cl<sub>2</sub> batteries are produced with a microporous carbon positive electrode, aluminium chloride in thionyl chloride as the electrolyte, and either sodium or lithium as the negative electrode.
- [Direct observation of ultrafast hydrogen bond strengthening in liquid water](#) [ 25 August 2021]
 

Article • Liquid ultrafast electron scattering measures structural responses in liquid water with femtosecond temporal and atomic spatial resolution to reveal a transient hydrogen bond contraction then thermalization preceding relaxation of the OH stretch.
- [High aboveground carbon stock of African tropical montane forests](#) [ 25 August 2021]
 

Article • The aboveground carbon stock of a montane African forest network is comparable to that of a lowland African forest network and two-thirds higher than default values for these montane forests.
- [Genome of a middle Holocene hunter-gatherer from Wallacea](#) [ 25 August 2021]
 

Article • The palaeogenomic analysis of a pre-Neolithic skeleton associated with a Toalean burial context from Wallacea.
- [Fair algorithms for selecting citizens' assemblies](#) [ 04 August 2021]
 

Article • Principles from the field of fair division are used to develop selection algorithms for citizens' assemblies that produce panels that are representative of the population while simultaneously selecting individuals with near-equal probabilities.

- **Oxytocin neurons enable social transmission of maternal behaviour** [ 11 August 2021]  
Article • Behavioural studies and neural recordings in mice show that virgin mice can acquire maternal behaviour through an oxytocin-dependent mechanism.
- **Cotranslational prolyl hydroxylation is essential for flavivirus biogenesis** [ 18 August 2021]  
Article • A proteomics-based strategy is used to examine how three different RNA viruses (polio, Zika and dengue) remodel translation in the host to recruit host machineries necessary for the production of viral proteins.
- **Antibody epitopes in vaccine-induced immune thrombotic thrombocytopenia** [ 07 July 2021]  
Article • Alanine-scanning mutagenesis is used to identify the PF4 epitope that is recognized by anti-PF4 antibodies in patients with vaccine-induced immune thrombotic thrombocytopenia, revealing that the epitope corresponds to the heparin-binding site on PF4.
- **Tonic prime-boost of STING signalling mediates Niemann–Pick disease type C** [ 21 July 2021]  
Article • The lysosomal protein Niemann–Pick type C1 (NPC1) is identified as a cofactor in the trafficking of STING, and loss of NPC1 is shown to both ‘prime’ and ‘boost’ STING signalling leading to neurological disease.
- **Cycling cancer persister cells arise from lineages with distinct programs** [ 11 August 2021]  
Article • Lineage tracing by barcoding of individual cells using a lentivirus library shows that cycling and non-cycling drug-tolerant persister cells in cancer arise from different lineages with distinct transcriptional and metabolic programs.
- **Highly accurate protein structure prediction with AlphaFold** [ 15 July 2021]  
Article • AlphaFold predicts protein structures with an accuracy competitive with experimental structures in the majority of cases using a novel deep learning architecture.
- **Highly accurate protein structure prediction for the human proteome** [ 22 July 2021]  
Article • AlphaFold is used to predict the structures of almost all of the proteins in the human proteome—the availability of high-confidence predicted structures could enable new avenues of investigation from a structural perspective.
- **Molecular basis for DarT ADP-ribosylation of a DNA base**  
[ 18 August 2021]  
Article • Structural and mechanistic data of the ADP-ribosyltransferase DarT demonstrate the role of ADP-ribosylation of DNA by this enzyme in generating toxicity and regulating cellular

signalling processes in bacteria.

- **Cryo-EM structures of full-length Tetrahymena ribozyme at 3.1 Å resolution** [ 11 August 2021]

Article • Cryo-electron microscopy has been used to determine the structure of the Tetrahymena ribozyme (a catalytic RNA) at sufficiently high resolution to model side chains and metal ions.

- **Unusual width of the superconducting transition in a hydride** [ 25 August 2021]

Matters Arising •

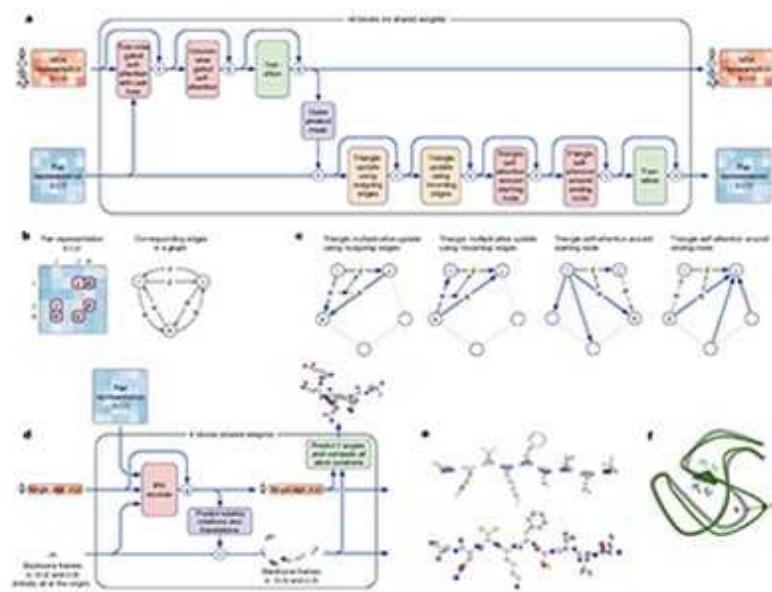
- NEWS AND VIEWS
- 23 August 2021

# Protein-structure prediction revolutionized

The full might of a world-leading artificial-intelligence laboratory has been brought to bear on protein-structure prediction. The resulting method, AlphaFold2, promises to transform our understanding of proteins.

- [Mohammed AlQuraishi](#) 0

Most proteins self-assemble into specific 3D structures that, together with other biological molecules, determine the function and behaviour of cells. Over the past five decades, biologists have experimentally determined the structures of more than 180,000 proteins and deposited them in the Protein Data Bank<sup>1</sup>, a freely available online resource. Despite this painstaking effort, the structures of hundreds of millions of proteins remain unknown, including more than two-thirds of those in the human proteome — the full set of proteins produced by our genome.

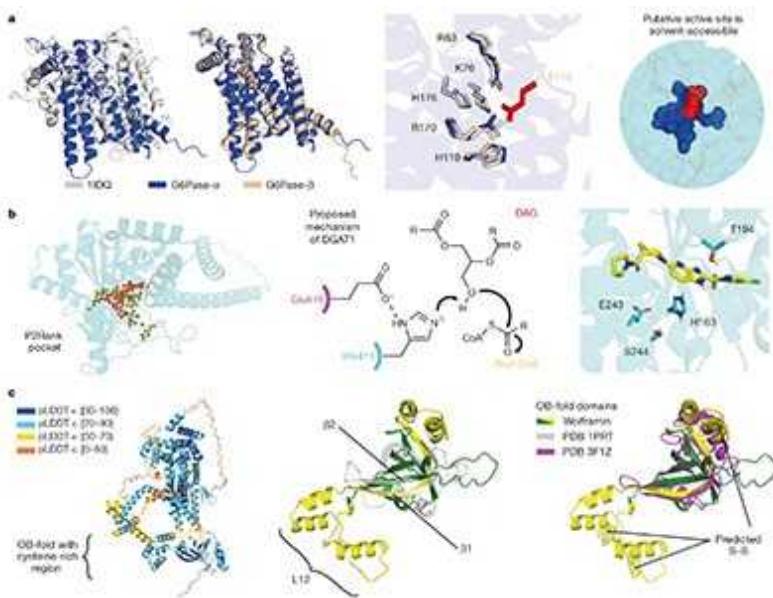


## [Read the paper: Highly accurate protein structure prediction with AlphaFold](#)

In two papers in this issue, scientists at DeepMind, Google's London-based sister company, describe a machine-learning method, AlphaFold2, that [predicts protein structures with near-experimental accuracy<sup>2</sup>](#), and report [its application to the human proteome<sup>3</sup>](#). DeepMind has also announced that it has applied AlphaFold2 to the proteomes of 20 model organisms (see [go.nature.com/2w6zhus](http://go.nature.com/2w6zhus)). AlphaFold2 is free for academics to use and, in collaboration with the European Bioinformatics Institute in Hinxton, UK, DeepMind will make the predicted structures of almost all known proteins freely available to all.

AlphaFold2 — as the name implies — is the second iteration of a system that DeepMind introduced three years ago at the Thirteenth Critical Assessment of Structure Prediction (CASP13) competition. The first version of AlphaFold was technically impressive<sup>4</sup>, and outperformed the other CASP13 entrants at the task of predicting protein structures from amino-acid sequences. However, it had a median accuracy of 6.6 ångströms for the most difficult set of proteins tested — that is, for the middle-ranked protein in the set, the atoms in the proposed structures were, on average, 6.6 Å away from their actual positions. This is much less accurate than experimental methods. Moreover, the original AlphaFold arguably represented only an incremental improvement over competing algorithms, in both design and performance.

AlphaFold2 fundamentally changes this. Its median accuracy at CASP14, which was held in 2020, was 1.5 Å — comparable to the width of an atom and approaching the accuracy of experimental methods. Moreover, its design has few parallels with existing algorithms.



[Read the paper: Highly accurate protein structure prediction for the human proteome](#)

The prediction of protein structures is difficult for many reasons: the number of plausible shapes for any given protein is huge, but an algorithm must pick just one; the number of known structures is (relatively) small, limiting the data available for training structure-predicting systems; the rules underlying protein biophysics are only approximately known, and are expensive to simulate; and the forces that determine a protein’s structure result not only from local interactions between nearby chemical groups in the protein molecule, but also from long-range interactions spanning the whole protein. Jumper *et al.*<sup>2</sup> report a multitude of ideas to address these challenges in their design of AlphaFold2.

Central to this design is a machine-learning framework — known as an artificial neural network — that considers both local and long-range interactions in protein molecules. This differs from previous algorithms, which commonly considered only local interactions to reduce the computational burden of structure prediction. AlphaFold2 does not try to

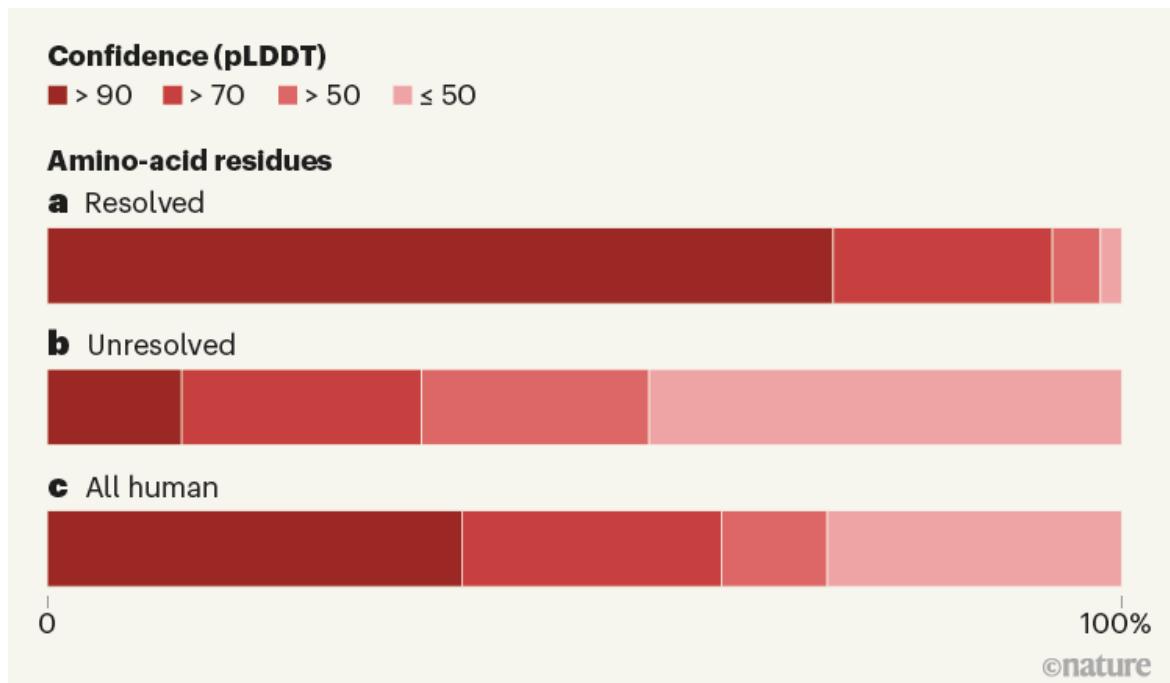
capture long-range interactions through computational brute force, which would be hopeless even with the resources available at Google. Instead, the authors introduced computational operations that efficiently capture long-range interactions on the basis of fundamental aspects of protein geometry. For example, the operations account for the fact that the coordinates of any three atoms in a protein must satisfy the triangle inequality rule (in other words, the sum of the lengths of any two sides of the triangle defined by the coordinates must be greater than or equal to the length of the remaining side).

AlphaFold2 applies these operations repeatedly (about 200 times) to gradually refine a model of a protein into its final 3D structure. Such iterative refinement, used millions of times, rather than hundreds, is a central component of physics-based approaches to protein-structure prediction<sup>5</sup>. But it is rarely used in machine-learning approaches — which instead predict structures by recognizing patterns of mutation in evolutionarily related proteins to detect co-evolving, and therefore spatially proximal, amino-acid residues<sup>6</sup>. AlphaFold2 breaks the mould by combining these two strategies. Crucially, it does not impose known rules of protein biophysics or try to mimic the physical process of protein folding, as has previously been attempted<sup>7,8</sup>. Instead, it performs purely geometric refinements learnt from its repeated attempts to predict protein structures. In this sense, it exemplifies the learning-driven revolution that has swept the field of protein modelling<sup>6,9</sup>.

In a companion paper, Tunyasuvunakool *et al.*<sup>3</sup> report the use of AlphaFold2 to predict the structures of almost all human proteins that independently acquire well-defined 3D shapes, for a total of 23,391 proteins. Predictions at this scale were previously possible, but three features of the new system provide a big leap forward.

First, the accuracy of the predictions is sufficiently high to generate biological insights and hypotheses that can be tested experimentally. Second, a calibrated self-assessment of each prediction provides a reliable estimate of correctness at the level of individual amino-acid residues (Fig. 1), enabling biologists to make inferences about confidently predicted regions. Third, AlphaFold2 is applicable to whole proteins, including large ones that

have multiple, independently self-assembling units — a common feature of mammalian proteins. The resulting resource ‘confidently’ predicts nearly 60% of all human-protein regions; most of the remaining regions might be unable to acquire well-defined structures, or be able to do so only in the presence of other biomolecules.



**Figure 1 | The confidence of protein-structure predictions by AlphaFold2.** Jumper *et al.*<sup>2</sup> report a machine-learning system, called AlphaFold2, that predicts the 3D structures of proteins from amino-acid sequences. Tunyasuvunakool *et al.*<sup>3</sup> used the same system to predict the structures of all human proteins that self-assemble into specific 3D structures. AlphaFold2 produces a confidence metric called the predicted local distance difference test (pLDDT) to estimate how well the predicted position of each amino-acid residue agrees with experimentally determined positions, on a scale of 1 to 100. The charts show the fractions of residues corresponding to different ranges of pLDDT for: **a**, residues that were previously resolved in structure-determination experiments (3,440,359 residues); **b**, residues that could not be resolved in experiments (589,079 residues); **c**, all of the residues in human proteins (10,537,122 residues). (Data from ref. 3.)

AlphaFold2 has already helped structural biologists to solve crystallographic protein structures<sup>10</sup> and refine ones derived from cryo-electron microscopy experiments. It provides biophysicists studying protein motion with starting (static) structures, and those studying protein interactions with hypotheses about how protein surfaces bind to each other. AlphaFold2 also presents opportunities to formulate new algorithms for bioinformatics based on protein structures, and might help systems biologists to understand the behaviour of cellular pathways and molecular machines on the basis of the structures that comprise them. And the study of evolution, which has long relied on genetic sequences, can now more readily be formulated in terms of the onset of new classes of protein structure (folds) and their relationship to cellular function and organismal fitness.

It is tempting to compare the scale of this advance to that of the Human Genome Project, but there are important differences. In contrast to the human genome sequence, the predicted structures have not been experimentally verified; it will take time for evidence of their correctness to emerge, so that scientists can gain confidence in the predictions. Of course, experimental measurements can also be affected by ‘noise’, bias and incompleteness — 20 years passed between the publication of the first draft of the human genome and the complete sequence<sup>11</sup> — and modern structure-determination techniques routinely involve some computational inference. As predictions improve, disagreements between protein models and experiments could become difficult to resolve, a situation familiar to physicists<sup>12</sup> but largely unprecedented in biology.

Disordered protein regions, which do not have well-defined shapes but often encode functionally crucial parts of proteins, present an ongoing and fundamental challenge to AlphaFold2 and, therefore, to our understanding of protein structure. Future methods must take this disorder into account and begin to reflect the flexibility inherent in most proteins.

Other differences between the Human Genome Project and the present advance are in AlphaFold2’s favour. Structure predictions are (relatively) cheap and will soon be available for all proteins, whereas genetic-sequencing technology took years to deploy and mature. Computational methods evolve rapidly, and it might therefore soon be possible to predict the structures of multi-protein complexes, alternative conformations of a

protein (for proteins that adopt them) and the structures of designed proteins with a level of accuracy similar to that currently achieved by AlphaFold2. Finally, protein structures provide immediate biological insights, because they fit within established conceptual frameworks that relate a protein's structure to its function — unlike genetic sequences, which were largely inscrutable at the dawn of the genomics era. The fruits of this revolution might thus be more swiftly reaped.

*Nature* **596**, 487–488 (2021)

doi: <https://doi.org/10.1038/d41586-021-02265-4>

## References

1. 1.  
Bernstein, F. C. *et al.* *J. Mol. Biol.* **112**, 535–542 (1977).
2. 2.  
Jumper, J. *et al.* *Nature* **596**, 583–589 (2021).
3. 3.  
Tunyasuvunakool, K. *et al.* *Nature* **596**, 590–596 (2021).
4. 4.  
Senior, A. W. *et al.* *Nature* **577**, 706–710 (2020).
5. 5.  
Kuhlman, B. & Bradley, P. *Nature Rev. Mol. Cell Biol.* **20**, 681–697 (2019).
6. 6.  
AlQuraishi, M. *Curr. Opin. Chem. Biol.* **65**, 1–8 (2021).

7. 7.

Jumper, J. M., Faruk, N. F., Freed, K. F. & Sosnick, T. R. *PLoS Comput. Biol.* **14**, e1006578 (2018).

8. 8.

Ingraham, J., Riesselman, A., Sander, C. & Marks, D. *Int. Conf. Learning Representations* <https://openreview.net/forum?id=Byg3y3C9Km> (2019).

9. 9.

Wu, Z., Johnston, K. E., Arnold, F. H. & Yang, K. K. *Curr. Opin. Chem. Biol.* **65**, 18–27 (2021).

10. 10.

Millán, C. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.06.21.449228> (2021).

11. 11.

Nurk, S. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.05.26.445798> (2021).

12. 12.

Brumfiel, G. *Nature* <https://doi.org/10.1038/nature.2012.10249> (2012).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02265-4>

- NEWS AND VIEWS
- 25 August 2021

# African tropical montane forests store more carbon than was thought

The inaccessibility of African montane forests has hindered efforts to quantify the carbon stored by these ecosystems. A remarkable survey fills this knowledge gap, and highlights the need to preserve such forests.

- [Nicolas Barbier](#) 0

In [a paper in \*Nature\*](#), Cuni-Sanchez *et al.*<sup>1</sup> report the assembly of a large database of tree inventories for 226 mature montane-forest plots in 12 African countries. The authors analyse the data to determine the amount of aboveground biomass and carbon stored in these highly diverse and threatened ecosystems. Their results suggest that African montane forests store more carbon than was previously thought, and the findings should help to guide efforts to conserve these ecosystems.

Cuni-Sánchez and colleagues measured trunk diameters and heights of the trees in plots, and identified the botanical species to deduce wood density — an approach that constitutes the gold standard for estimating the biomass, and thus the amount of carbon, contained per unit of forest area. This method involves the use of general statistical equations for describing tree form, called allometric models, and considers only the aboveground parts of trees. It therefore disregards several other pools of carbon, notably in the roots and soil. The overall approach might seem crude, but recognizing and measuring the many hundreds of tree species found on steep, cloud-shrouded slopes (Fig. 1), let alone the underground carbon, without visiting

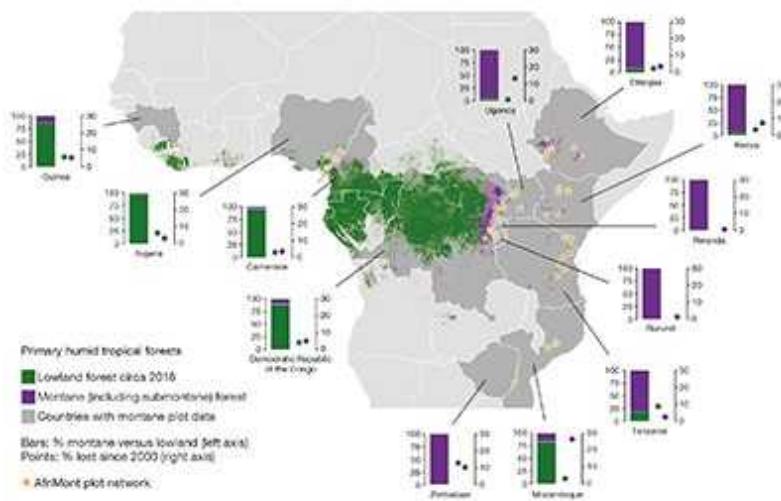
the sites, will long remain difficult, even with the best drones and satellite systems.



**Figure 1 | Montane forest in Boginda, Ethiopia.** Cuni-Sanchez *et al.*<sup>1</sup> use data from a survey of montane tropical forests in Africa to quantify the amount of carbon stored above ground in these ecosystems. Credit: Bruno D'Amicis/Nature Picture Library

Anyone who has conducted field inventories in tropical mountains knows that measuring and identifying 72,336 trees, often just a few steps away from the void, is an amazing feat. For comparison, a previously reported study<sup>2</sup> based its estimates of the carbon stored in montane African forests on as few as seven plots. The study also brings together contributions from numerous researchers and institutions, including many in Africa, to greatly increase the size of the data set, which is also a remarkable achievement. Even so, the total area of forest studied is less than 150 hectares, whereas African montane forest covers about 100,000 times that area, inevitably raising questions about how representative the inventory is.

Statisticians might raise their eyebrows at the sampling design. As is usually the case in meta-analyses, the data set was neither homogeneous (for example, there is a roughly tenfold variation in the plot sizes), nor were the sites selected at random. However, the authors did their best to rule out possible biases induced by sampling artefacts.



## [Read the paper: High aboveground carbon stock of African tropical montane forests](#)

Cuni-Sánchez *et al.* chose not to discuss one tricky aspect of surveys of this sort (extensively discussed elsewhere<sup>2</sup>): how should the land area of a steep slope be measured? The authors followed standard practice, which is to measure the extent of forest plots and of land-cover types in reference to horizontal, planimetric areas (that is, the areas that would be represented on a 2D map, as if seen from the air). This tends to overestimate aboveground carbon because the sloped surface area is greater than that of the planimetric area — which means that the tree density of the planimetric area is higher than it is on the slope. By contrast, the use of planimetric areas underestimates total montane-forest area (by about 40%; see ref. 2). These two biases should roughly cancel each other out when estimating carbon stocks, or changes to stocks, for a region or country. But care should be taken not to combine data acquired using planimetric and non-planimetric areas in future meta-analyses, because the resulting estimates could end up well off the mark.

One might expect that trees in mature African montane forests would be, on average, shorter — and therefore store less carbon — than their lowland counterparts, because of their lower environmental temperatures and shallow soils, frequent landslides and strong winds. However, this is not what Cuni-Sanchez *et al.* report. Instead, they find that average aboveground carbon stocks are not significantly different from those of mature lowland forests. This contrasts with the situation in the neotropics and southeast Asia, where montane forests store, on average, less carbon than do lowland forests.

However, the new results fit with the 2016 discovery that the tallest African trees (81.5 metres) grow on Mount Kilimanjaro<sup>3</sup>, the highest mountain in Africa. African forests, in general, tend to contain fewer but larger-statured tree stands than does, for example, Amazonia<sup>4</sup>. The current study confirms that this peculiarity applies even at high altitudes.

The authors investigate several possible drivers for the variations in biomass observed at different sites in their study, including topography, climate, landslide hazard, and even the presence of elephants or certain conifers (Podocarpaceae), but were unable to identify any clear pattern. Many environmental, historical and biological effects probably interact, with each of these effects varying greatly in ways that are poorly captured by available data sets. These effects must therefore be disentangled before a predictive model of African montane carbon distribution can be developed.



## Tropical carbon sinks are saturating at different times on different continents

Nevertheless, Cuni-Sánchez and colleagues' study underlines a crucial message: African montane forests are immensely valuable, and not only because they host the source of the River Nile, mountain gorillas and ecosystems such as mysterious lichen-covered forests. They also store vast amounts of carbon, and thereby have a key role in tackling climate change. Of course, this immense intrinsic value does not preclude intense human exploitation of these ecosystems, which can lead to rapid degradation and deforestation. For instance, on the basis of satellite monitoring, Cuni-Sánchez and colleagues report that Mozambique lost nearly one-third of its montane forests between 2000 and 2018.

There is, however, the faint hope that putting a financial value on carbon, and the establishment of economic incentives to avoid deforestation in tropical countries, might help to check the flood of damage<sup>5</sup>. The aim is to reward African countries — for which montane forest sometimes constitutes the last remaining forests — for their conservation endeavours, and for renouncing efforts to access the timber and ore in these ecosystems, even when such resources are otherwise desperately lacking. By gathering the best-available data to provide precise, country-level estimates of average aboveground carbon content in African montane forests, Cuni-Sánchez and colleagues' study will add weight to such efforts — not least because the new estimates are, on average, two-thirds higher than the values reported by the Intergovernmental Panel on Climate Change<sup>6</sup>.

The next step should be to extend measurements in these forests, particularly by continuing to support national forest-inventory efforts. These inventories target all vegetation types, rather than just the most intact forests, and all carbon pools, using standardized protocols and systematic sampling methods. Remote sensors, both in the sky and in space, should also be used to fully map the detailed spatial variation of forest diversity, structure and dynamics. But there is no excuse for delaying policymaking — we already know enough to justify immediate decisive action to preserve yet another of Earth's threatened treasures.

*doi: <https://doi.org/10.1038/d41586-021-02266-3>*

## References

1. 1.  
Cuni-Sanchez, A. *et al.* *Nature* **596**, 536–542 (2021).
2. 2.  
Spracklen, D. V. & Righelato, R. *Biogeosciences* **11**, 2741–2754 (2014).
3. 3.  
Hemp, A. *et al.* *Biodivers. Conserv.* **26**, 103–113 (2017).
4. 4.  
Lewis, S. L. *et al.* *Phil. Trans. R. Soc. B* **368**, 20120295 (2013).
5. 5.  
Venter, O. *et al.* *Science* **326**, 1368 (2009).
6. 6.  
Domke, G. *et al.* in *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* Vol. 4 (eds Calvo Buendia, E. *et al.*) Ch. 4, 4.48 (IPCC, 2019).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02266-3>

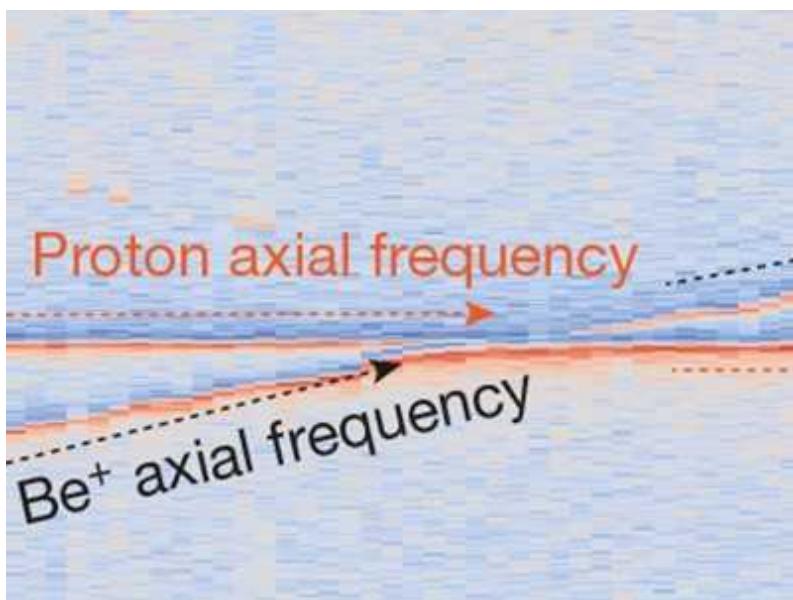
- NEWS AND VIEWS
- 25 August 2021

# Single proton cooled by distant ions

Laser-cooled ions have been used to substantially lower the temperature of a proton located several centimetres away. This technique could be useful in ultraprecise measurements of the properties of antimatter particles.

- [Manas Mukherjee](#) 

In [a paper in \*Nature\*](#), Bohman *et al.*<sup>1</sup> (the BASE Collaboration) report the cooling of a single proton by a cloud of laser-cooled beryllium ions. Remarkably, the ions were separated from the proton by a distance of about 9 centimetres — which is too far apart for the charges on the ions to have interacted with that of the proton. This means that the ions could not have exerted a direct cooling effect on the proton. Instead, the researchers used an indirect cooling process, mediated by an electric circuit that established an effective interaction. This approach has potential applications in studies of antimatter particles and in the field of quantum information.

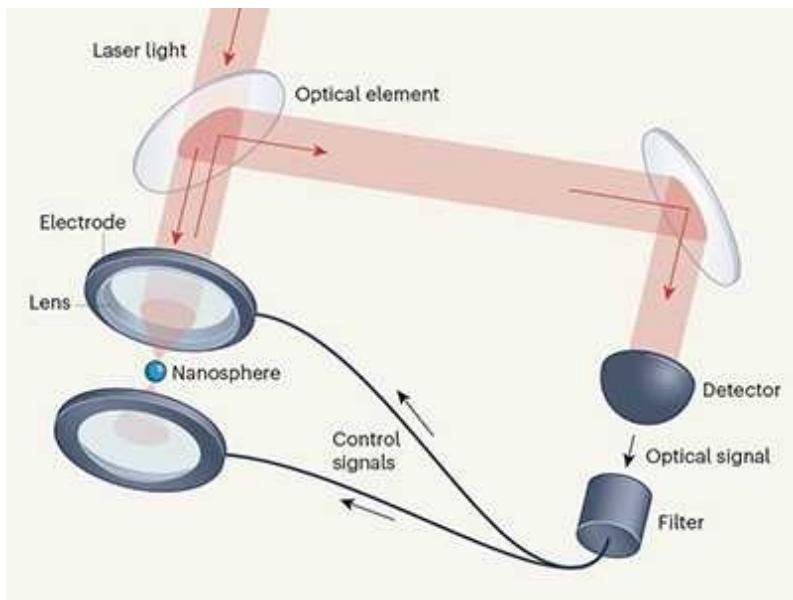


## [Read the paper: Sympathetic cooling of a trapped proton mediated by an LC circuit](#)

A wealth of knowledge about nature's inner workings comes from studies of fundamental particles, such as electrons and protons. Currently, the most accurate theoretical model of the forces of nature is the standard model of particle physics, which describes how fundamental particles interact with each other and thereby build up the macroscopic world. The standard model has passed many stringent tests using various experimental tools, at particle energies that range from  $10^{11}$  electronvolts in particle accelerators<sup>2</sup> to only about 0.0001 electronvolts in ion traps<sup>3</sup>. However, it is widely accepted that the standard model does not explain some natural phenomena, such as the fact that the Universe is made up of only matter. It also does not account for the existence of dark matter — the invisible and largely unaccounted for mass of the Universe.

High-precision measurements of fundamental particles and their corresponding antiparticles provide opportunities to verify the standard model, and maybe even to find evidence of new physics that goes beyond the currently accepted model<sup>4,5</sup>. Two conditions must be met to perform such measurements: the particles must be spatially confined; and they should be very nearly at rest (that is, the particles must be cooled to almost zero kelvin, to minimize their kinetic energy). The first of these requirements can be solved using a combination of static electric and magnetic fields in a device called a Penning trap.

Laser-cooling methods were first reported<sup>6,7</sup> in 1975, and have since been widely developed to reduce the motion of particles. This approach works well for atoms, but not for particles that do not absorb light, such as protons. Scientists have therefore invented other cooling methods, such as resistive cooling<sup>8</sup> (in which ions dissipate their energy by inducing a current in a cold electric circuit) and synchrotron cooling<sup>9</sup> (in which fast-rotating particles with low mass radiate energy by emitting electromagnetic radiation). However, the lowest particle temperatures achieved using those approaches are roughly 1,000 to one million times higher than those of laser-cooled atoms.



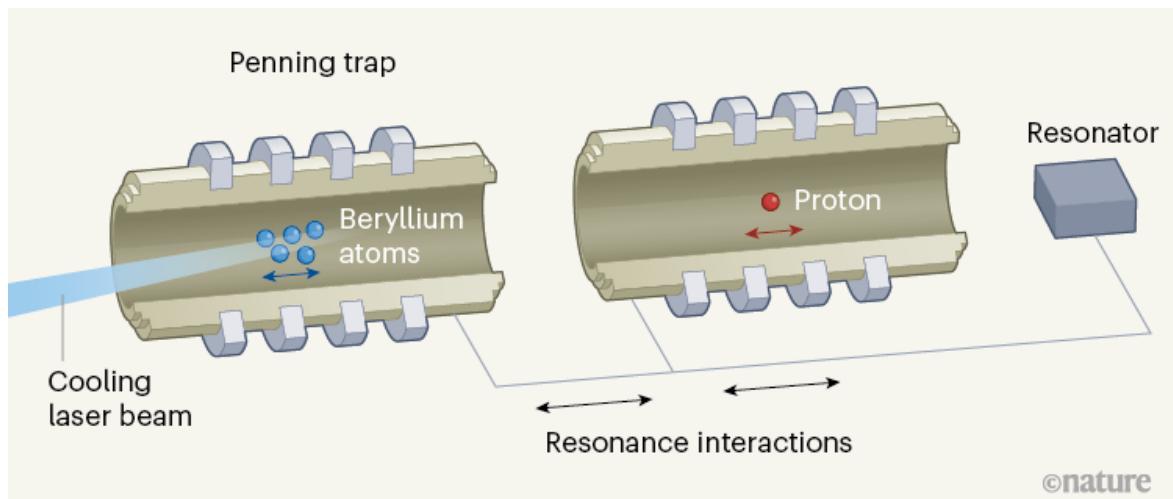
### Measurement-based system provides quantum control of nanoparticles

An interesting alternative is to cool a charged particle by bringing it close to another, colder charged particle<sup>10</sup> — an approach commonly known as sympathetic cooling. For example, consider a positively charged atomic ion that is being continuously laser-cooled to one-thousandth of a kelvin, and which is then brought close to a proton that is initially at 4 K in an ion trap. The proton and ion will repel each other within the confinement of the ion trap, effectively transferring kinetic energy from the proton to the ion. Because the ion is constantly being laser-cooled, the repulsive interactions will eventually chill the proton to the same temperature as the ion, even though the proton is not being cooled directly.

Sympathetic cooling works well, but the nearby presence of an ion would be undesirable when making ultraprecise measurements of a proton's properties. Furthermore, the method requires that the particle and the ion have charges of the same polarity, to provide the necessary repulsive interactions. Bohman and colleagues' work provides a potential solution to these issues.

The authors used separate Penning traps to confine a cloud of beryllium ions and a proton in an ultrahigh vacuum, and continuously laser-cooled the ions (Fig. 1). The proton and the ions were then set up to 'talk' to an electrical resonator circuit, which enables the two trapped-particle systems to interact

only when the natural oscillation frequencies (the resonance frequencies) of the two systems match exactly. Bohman *et al.* demonstrated the influence of the ions on the proton using an established technique, in which electrical ‘noise’ in the resonator circuit is analysed to directly determine the temperatures of the two systems.



**Figure 1 | Sympathetic cooling at a distance.** Bohman *et al.*<sup>1</sup> used devices called Penning traps to capture a proton at 17 kelvin and a cloud of beryllium ions, which was continuously cooled by a laser to a much lower temperature. The traps were connected by a wire to a circuit known as a resonator. Oscillations (coloured arrows) of the proton and of the ions generate electrical currents (not shown) in the electrodes of their respective traps; these currents oscillate at the same frequency as the particles that generate them. If the natural oscillation frequency of the proton is the same as that of the ions and of the electrical current in the resonator circuit, a phenomenon called resonance allows the currents in the system to interact. The ions therefore cool the resonator, which, in turn, cools the proton. Such indirect cooling of the proton by the ions is called sympathetic cooling through the resonator.

To further ensure that the proton cooling is indeed caused by the ions, the authors fixed the oscillation frequency of the proton, and then varied the oscillation frequency of the ions. They observed that cooling interactions occurred only when the ions’ natural oscillation frequency matched that of both the proton and the resonator circuit, as expected. Furthermore, the researchers found that numerical simulations of the cooling set-up matched

the observed experimental result, confirming the ions' proton-cooling influence.

Impressively, Bohman *et al.* show that the proton temperature can be reduced by 85%, which would be a substantial amount in an ultraprecise measurement of a fundamental particle. The authors' technique opens up the possibility of being able to cool any charged particle by 'wiring it up' to laser-cooled ions, with any distance between the particle and the ions.

The results also have implications for research in quantum information. A goal for this field is to exchange single bits of quantum information between spatially separated quantum systems. However, it is challenging to do this using a conducting wire. Bohman and colleagues' findings suggest a possible solution to this problem, but it will first be necessary to broaden our understanding of how single quanta of energy are exchanged over large distances, and to greatly improve the rate of energy exchange between the separated systems.

*Nature* **596**, 490–491 (2021)

doi: <https://doi.org/10.1038/d41586-021-02267-2>

## References

1. 1.

Bohman, M. *et al.* *Nature* **596**, 514–518 (2021).

2. 2.

CMS Collaboration. *Phys. Lett. B* **716**, 30–61 (2012).

3. 3.

Neuhauser, W., Hohenstatt, M., Toschek P. E. & Dehmelt, H. *Phys. Rev. A* **22**, 1137–1140 (1980).

4. 4.

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

5. 5.

Schneider, G. *et al.* *Science* **358**, 1081–1084 (2017).

6. 6.

Hänsch, T. W. & Schawlow, A. L. *Opt. Commun.* **13**, 68–69 (1975).

7. 7.

Wineland, D. J. & Dehmelt, H. *Bull. Am. Phys. Soc.* **20**, 637 (1975).

8. 8.

Cornell, E. A. *et al.* *Phys. Rev. Lett.* **63**, 1674 (1989).

9. 9.

Wineland, D. J. & Dehmelt, H. G. *J. Appl. Phys.* **46**, 919 (1975).

10. 10.

Drullinger, R. E., Wineland, D. J. & Bergquist, J. C. *J. Appl. Phys.* **22**, 365–368 (1980).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02267-2>

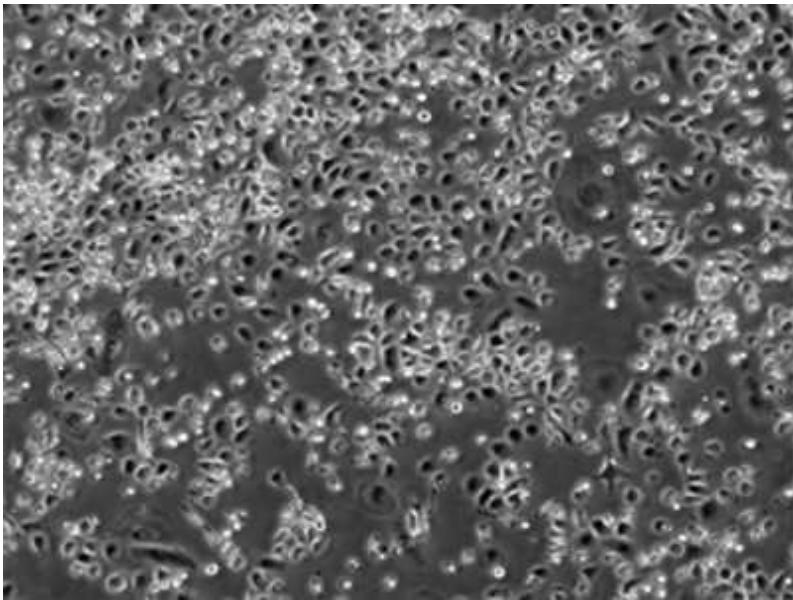
- NEWS AND VIEWS
- 11 August 2021

# A persistent look at how tumours evade therapy

Understanding how resistance to chemotherapy occurs could lead to better anticancer treatments. Persister cells in tumours can contribute to this resistance. A method to characterize these cells in detail sheds light on their origins.

- [Karen Gomez](#) <sup>0</sup> &
- [Raul Rabadan](#) <sup>1</sup>

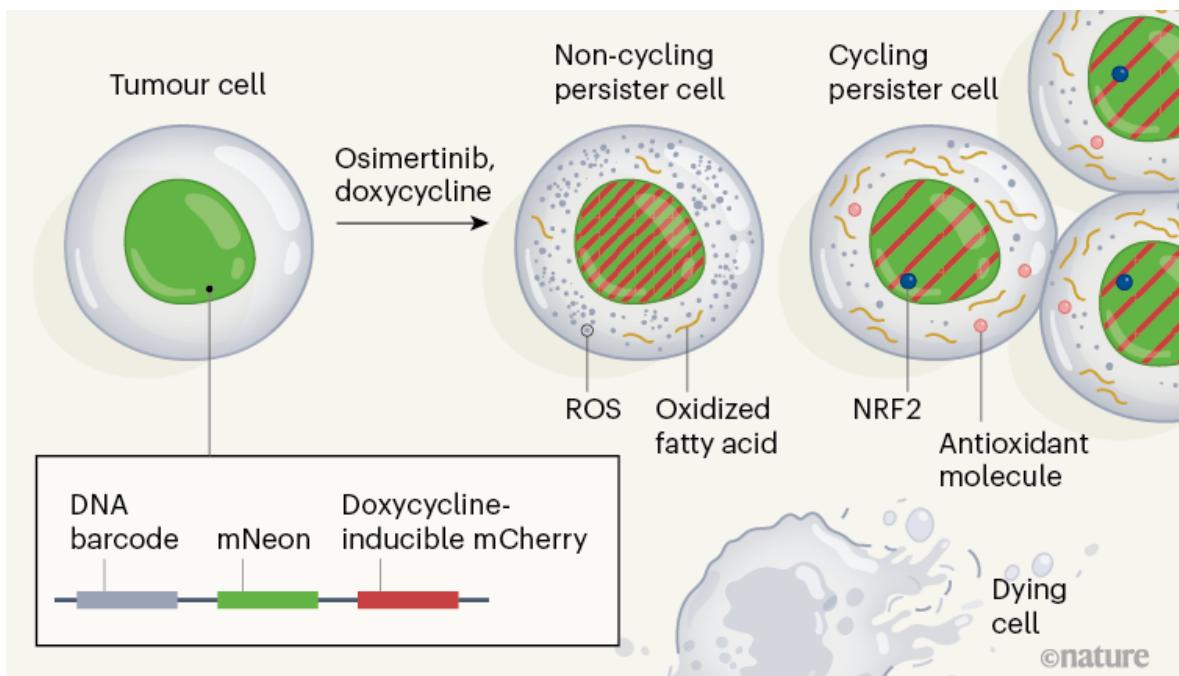
Cancer can recur when a subset of tumour cells, called persister cells, survive chemotherapy. Most of these persisters are non-dividing (quiescent) in the presence of the therapeutic drug, but a rare subpopulation can re-enter the cell cycle during treatment, which enables them to proliferate. Much research has focused on the genetic mechanisms underlying such resistance to treatment. However, emerging data suggest that non-genetic mechanisms (such as changes to the complex of DNA and protein called chromatin) might also have a role in the development of a persistent state. [Writing in \*Nature\*](#), Oren *et al.*<sup>1</sup> examine the cellular lineages and gene-expression profiles of persister cells by using a method called DNA barcoding to trace tumour cells and their descendants. Their findings illuminate the role of non-genetic, reversible mechanisms in resistance to chemotherapy for a range of tumours from different tissues.



[Read the paper: Cycling cancer persister cells arise from lineages with distinct programs](#)

The authors analysed cell divisions in human lung cancer cells grown *in vitro* that have a mutation in the gene encoding the epidermal growth factor receptor (EGFR). The cells were treated with osimertinib, an inhibitor of this receptor. Oren and colleagues tracked the outcomes for cellular lineages of the tumour cell line and found that 8% of the lineages gave rise to persister cells after 14 days, and 13% of the persisters resumed the cell cycle and proliferated to form cell colonies. These results show that these cycling and non-cycling persisters arise early during the course of treatment, and that they evolve from separate cell lineages.

To characterize the molecular mechanisms associated with cycling and non-cycling persister cells, the authors developed a system that they call Watermelon, to simultaneously trace each cell's lineage, proliferation status and transcriptional state (Fig. 1). To determine whether the persister state was due to a genetic, irreversible property of the persister cells, the authors re-exposed the persister cell population to osimertinib after a pause in treatment. They found that cells from both cycling and non-cycling populations reacquired drug sensitivity, suggesting that a non-genetic, reversible mechanism underlies persistence.

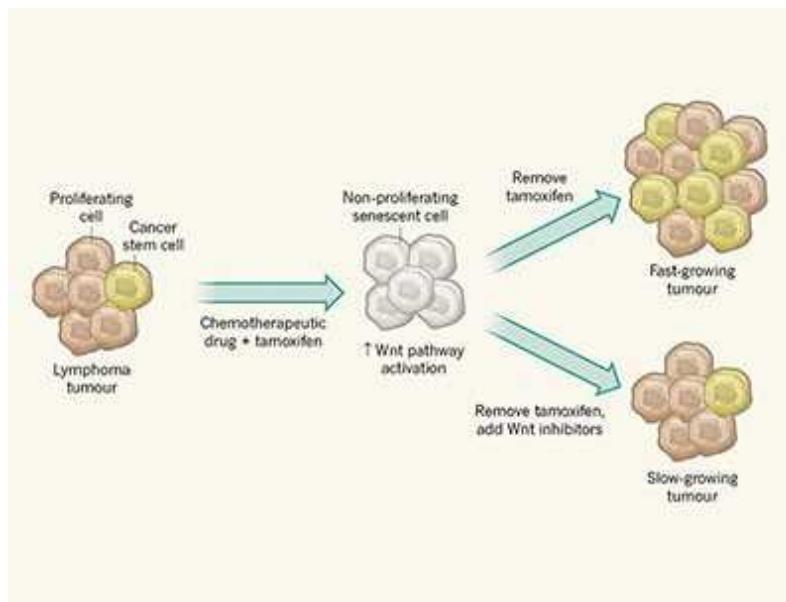


**Figure 1 | A method to analyse persister cells in tumours.** Oren *et al.* present the Watermelon technique for analysing persisters — tumour cells that evade destruction by chemotherapeutic drugs such as osimertinib. The authors introduced an engineered DNA sequence into human tumour cells grown *in vitro*. The sequence contains a unique DNA ‘barcode’, which identifies the cells in that lineage. Also included are genes encoding a green fluorescent protein (mNeon), and a red fluorescent protein, mCherry, that requires the molecule doxycycline for its expression. The authors treated the cells with osimertinib and doxycycline, and analysed the surviving persister cells using single-cell analysis. The non-cycling (non-dividing) persisters had a higher level of mCherry compared with the cycling (dividing) persisters. The persister cells contain reactive oxygen species (ROS), which cause oxidative damage. Cycling persister cells had lower levels of ROS and higher levels of oxidized fatty acids compared with the non-cycling persisters. Cycling persister cells display hallmarks of antioxidant defences, including the expression of antioxidant molecules and the transcription factor NRF2.

The authors assessed gene expression using the method of single-cell RNA sequencing at different time points during a two-week treatment, and compared these signatures in cycling and non-cycling persisters. The cycling persistent state was uniquely characterized by the upregulation of defence

programs that produce antioxidant molecules — including expression signatures characteristic of the metabolism of the antioxidant glutathione, as well as production of the protein NRF2, which is a transcription factor induced in response to oxidative stress. Moreover, the expression of several genes that are NRF2 targets correlated with lineages that had a large number of descendant persister cells, and the genetic engineering of cells to deplete a negative regulator of NRF2 resulted in an increase in the fraction of persisters that were cycling.

Osimertinib treatment induced the formation of reactive oxygen species (ROS), which can cause oxidative stress. At the end of treatment, cycling persisters had significantly lower levels of ROS compared with the non-cycling persisters. When the authors decreased ROS levels in cells through the addition of ROS scavenger molecules, the fraction of persister cells that were cycling increased. These analyses therefore suggest that the redox state of cells has a role in the regulation of cycling persisters.



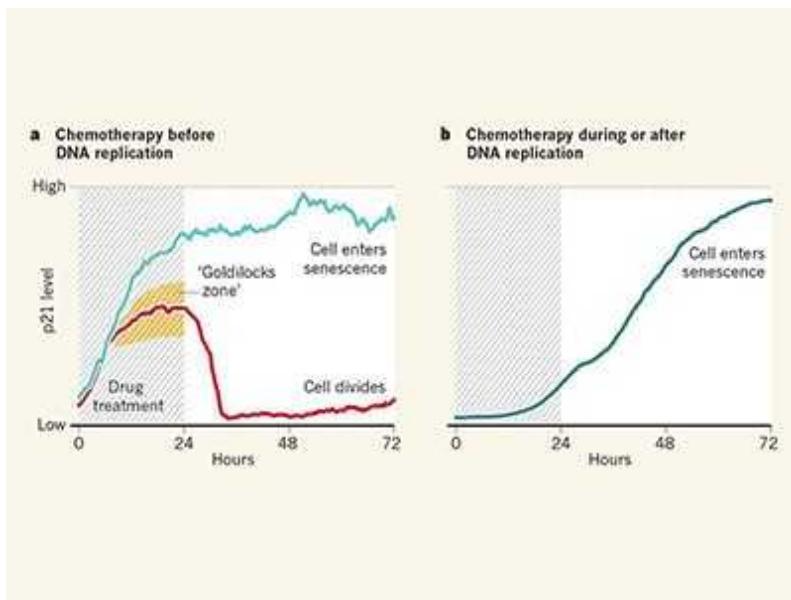
### Escape from senescence boosts tumour growth

Recognizing that redox balance is linked to metabolism, the authors profiled the products of metabolism in the cycling and non-cycling persisters, and identified 56 products that differed in their abundance between these two cell populations. The authors found a greater abundance of fatty acids linked to the molecule carnitine (a result of a preliminary step in the oxidation of

fatty acids) in cycling states than in non-cycling states. The authors also noted an increase in the oxidation of fatty acids as a consequence of osimertinib treatment. Modulation of the pathway affecting fatty-acid oxidation revealed that increasing or decreasing fatty-acid oxidation leads to an increase or decrease in the fraction of cycling persisters, respectively. These results support the idea that a metabolic shift in fatty-acid oxidation affects the proliferative capacity of persisters.

To test whether their observations extended beyond the model system of lung cancer, Oren *et al.* generated Watermelon models of further types of human cancer, using melanoma, lung, breast and colorectal tumours. They treated the cells with suitable inhibitors, characteristic of chemotherapies, depending on the genetics underlying the particular cancer. In most of these models, the cycling persisters showed elevated fatty-acid metabolism, antioxidant responses and NRF2 signatures compared with the non-cycling persisters, showing that the authors' findings extend to cancer types other than lung cancer.

These *in vitro* findings were validated using an engineered mouse model in which the animals had an inducible version of a mutant EGFR in lung tumours. After osimertinib treatment, the persister cells had a higher level of ROS and gene-expression signatures characteristic of fatty-acid metabolism compared with the cells in mice that had not received treatment. The authors also assessed gene-expression changes before and after chemotherapy in samples of cells from people with EGFR-driven lung adenocarcinoma, with melanoma driven by a mutant version of the enzyme BRAF (treated with inhibitors of BRAF and the enzyme MEK), and with breast cancer driven by a mutant version of the HER2 protein (treated with lapatinib). In all these scenarios, signatures of ROS production and fatty-acid metabolism were increased in the persister cells after treatment compared with samples of untreated tumour cells, and were higher in cycling than in non-cycling persisters.



### A dynamic view of chemotherapy effectiveness

Oren and colleagues' study fits into the wider context of current work highlighting the importance of non-genetic mechanisms in persister-cell survival and proliferation<sup>2–4</sup>. One major problem when studying persisters is that they are a small fraction of the initial population of tumour cells, making it difficult to characterize them by sequencing cells in bulk. The value of the authors' Watermelon method is that it enables the detailed characterization of persisters at the resolution of single cells. One future direction might be to apply similar single-cell approaches to study non-genetic mechanisms of resistance in other types of cancer, such as pancreatic<sup>5</sup> or prostate<sup>6</sup> tumours, which are fields where such research is emerging.

Understanding the dynamics of persister cells is crucial to the development of more-effective chemotherapies for cancer treatment. Previous studies found that the response pathway to the hormone oestrogen<sup>7</sup>, which has a role in breast cancer, and the pathway related to the cell-death process termed ferroptosis<sup>8,9</sup> are associated with the persister state. Oren *et al.* found that, although inhibiting these pathways did decrease the amount of persister cells, there was an increase in the fraction of persisters that were cycling, suggesting that these would not be optimal chemotherapy targets.

By contrast, the authors report that inhibiting the pathway for fatty-acid oxidation using the inhibitor drug etomoxir resulted in a decrease in both the fraction of cells that were persisters and the fraction of the persisters that were cycling. This promising result indicates that modulation of this pathway, and genes that have functions related to this pathway, might be worth considering in the development of new treatment strategies.

*Nature* **596**, 491-493 (2021)

doi: <https://doi.org/10.1038/d41586-021-02117-1>

## References

1. 1.

Oren, Y. *et al.* *Nature* <http://doi.org/10.1038/s41586-021-03796-6> (2021).

2. 2.

Salgia, R. & Kulkarni, P. *Trends Cancer* **4**, 110–118 (2018).

3. 3.

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

4. 4.

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

5. 5.

Chen, P.-Y. *et al.* *Cancer Res.* **78**, 985–1002 (2018).

6. 6.

Jolly, M. K., Kulkarni, P., Weninger, K., Orban, J. & Levine, H. *Front. Oncol.* **8**, 50 (2018).

7. 7.

Garon, E. B. *et al.* *J. Thorac. Oncol.* **8**, 270–278 (2013).

8. 8.

Angeli, J. P. F., Shah, R., Pratt, D. A. & Conrad, M. *Trends Pharmacol. Sci.* **38**, 489–498 (2017).

9. 9.

Hangauer, M. J. *et al.* *Nature* **551**, 247–250 (2017).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02117-1>

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

- NEWS AND VIEWS
- 04 August 2021

# A bridge across the democracy–expertise divide

An innovative algorithm provides a way of fairly selecting representative individuals for citizens' assemblies to learn about and deliberate on certain topics. Such groups hold promise for closing the gap between democracy and expertise.

- [Mark E. Warren](#) 9

There is a growing gulf between experts and citizens. Distrust in science is on the increase, as are conspiracy theories that challenge evidence-based decision-making. Populist attacks on institutions that provide expertise for democratic societies and processes — administrative agencies, universities and research organizations — are on the increase, facilitated by social media. Can we strengthen democracy while also ensuring that governance benefits from expertise? [Writing in Nature](#), Flanigan *et al.*<sup>1</sup> present a way of fairly and democratically selecting representative groups of citizens tasked with advising on issues that often combine politics and expertise. As supplements to the conventional institutions of electoral democracy, these bodies show promise as a means of bridging the democracy–expertise gulf.

Citizens' assemblies, the term used by Flanigan and co-authors, are a form of deliberative minipublic, the term I use here: bodies of 20–500 ordinary citizens selected near-randomly, through a process often known as sortition, and convened to learn, deliberate and make recommendations to decision makers and sometimes to the broader public (Fig. 1). They achieve three things that more-familiar institutions of democratic government do not<sup>2</sup>.

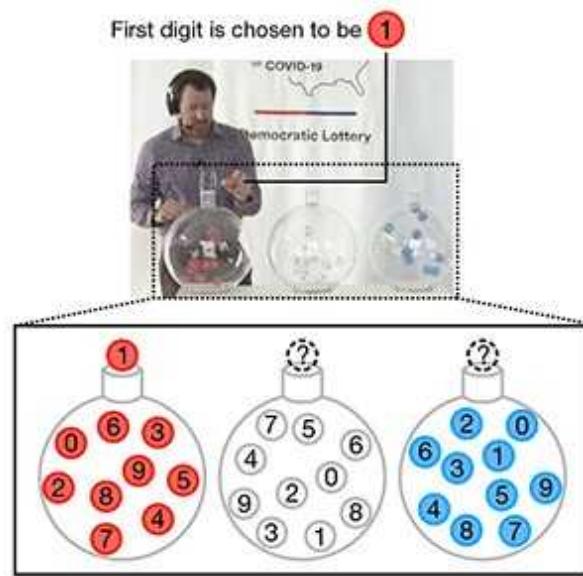


**Figure 1 | Members of Climate Assembly UK, a citizens' assembly on climate change.** Citizens' assemblies are a form of deliberative minipublic, in which 20–500 individuals who are representative of the demographics of a broader public are near-randomly selected to learn about, deliberate on and make recommendations on certain topics. Flanigan *et al.*<sup>1</sup> devised an algorithm to select individuals for citizens' assemblies fairly and in a way that ensures assemblies are demographically representative. Credit: Fabio De Paola/Alamy

First, because members of deliberative minipublics are selected to mirror a relevant public (they are descriptively representative of the public), they do a better job of representing groups that tend to be under-represented in elected bodies (such as legislatures), or in processes for which participants self-select (such as public hearings, petitioning and lobbying). Second, because a few ordinary citizens are acting as representatives of other citizens, the public tends to like and trust these bodies, probably because they are non-elitist, and not invested in professional politics. Third, deliberative

minipublics integrate expertise, because members are tasked with learning about an issue, hearing from experts and advocates, and then deliberating over recommendations.

There is abundant evidence<sup>3,4</sup> that, when supported in well-designed processes, ordinary citizens can integrate expertise with moral, value-based and political considerations. On ‘hot’ issues (such as abortion, climate change and Brexit), deliberative minipublics’ demographically representative samples of citizens tend to be less polarized than are advocates and elected representatives. This is in part because the selection process does not over-represent what are known as motivated reasoners — people who select information to support a pre-conceived position.



[Read the paper: Fair algorithms for selecting citizens' assemblies](#)

Flanigan and colleagues focus on the composition of deliberative minipublics. They propose and test an algorithm that maximizes fairness in selecting members by equalizing the probability of selection. Why is this important? As political entities, deliberative minipublics must be viewed as legitimate representative bodies by the broader public, if they are to bridge the democracy–expertise divide. Although research remains patchy, evidence<sup>5</sup> suggests that their legitimacy, as perceived by the broader public, is driven by their being representative of people who are ‘like us’ — underscoring the political value of such descriptive representation.

Our current understanding is that people like deliberative minipublics in part because they represent ordinary citizens, and not elites with political agendas. As such, people are more likely to trust the results obtained<sup>5</sup>. Furthermore, recommendations delivered by these citizens' assemblies often have greater impact on the public than does the same information delivered by experts<sup>4</sup>. To achieve this kind of legitimacy, those who organize deliberative minipublics must ensure the credibility of the selection processes.

But this is easier said than done. Deliberative minipublics are often constituted in a two-step process: invitations are sent to a certain number of people — usually more or less randomly selected from, say, voter lists. As Flanigan and colleagues note, those who respond positively, however, are not representative of the broader population: they are usually older, more educated, more likely speak the dominant language or belong to a dominant ethnicity, and more likely to have stable residence. A random selection from this group would thus not produce a descriptively representative body.

Because of this, organizers typically impose demographic quotas for certain categories when they select a minipublic from a pool of volunteers, so that they can produce a body that looks like the broader public. This is where the algorithm designed by Flanigan and colleagues does its work: using ideas from the field of 'fair division', an aspect of game theory, the algorithm equalizes the chances that someone will be selected to serve, even given the quota constraints necessary to correct for bias in the volunteer pool. Doing this maximizes the chances that the body will look like the broader public, without unfairness in the selection process.



### Encounters with inequality lead to demands for taxes on the rich

Challenges remain. Broader publics, especially those with a high proportion of distrustful citizens, still need to be persuaded that the selection processes used are fair, and few are likely to understand the selection algorithm itself. Therefore, much depends on the credibility of the organizers, and on their ability to translate the selection algorithm into a visible, transparent process that intuitively makes sense — showing, for example, that the algorithm performs like a lottery. Organizers must avoid perceptions that quotas to remove bias ‘rig’ the process. And it is not only citizens who must have confidence in deliberative minipublics, but also political elites, whose views of citizens’ capacity to understand a topic might be affected by the successes of populist politicians who mobilize ignorance.

Despite these challenges, deliberative minipublics are one of the most promising ways of reducing the widening gulf between democracy and expertise. And they are gaining in use and effectiveness<sup>6</sup>. Although there is no authoritative census of such initiatives, a search on the crowdsourcing website *Participedia.net* (<https://participedia.net>) combining the search terms “random sampling” or “stratified random sampling” with a method involving a “deliberative and dialogic process” returns more than 520 events from around the world in which randomly selected groups learn and deliberate. A report<sup>6</sup> published last year from the Organisation for Economic Co-operation and Development (OECD) lists 290 “representative

“deliberative” processes in OECD countries, including some that started back in the 1990s, although most were organized within the past decade.

As deliberative minipublics become more widely used, it is crucial that citizens and political elites view them as credible and legitimate. Flanigan and colleagues’ selection algorithm is a key step forwards, and is likely to provide a global benchmark for boosting this promising democratic innovation.

*Nature* **596**, 493-494 (2021)

doi: <https://doi.org/10.1038/d41586-021-02006-7>

## References

1. 1.

Flanigan, B., Gölz, P., Gupta, A., Hennig, B. & Procaccia, A. D. *Nature* <https://doi.org/10.1038/s41586-021-03788-6> (2021).

2. 2.

Setälä, M. & Smith, G. in *The Oxford Handbook of Deliberative Democracy* (eds Bächtiger, A., Dryzek, J. S., Mansbridge, J. & Warren, M. E.) 300–314 (Oxford Univ. Press, 2018).

3. 3.

Fishkin, J. S. *Democracy When the People are Thinking: Revitalizing Our Politics Through Public Deliberation* (Oxford Univ. Press, 2018) .

4. 4.

Gastil, J. & Knobloch, K. *Hope for Democracy: How Citizens Can Bring Reason Back into Politics* (Oxford Univ. Press, 2020).

5. 5.

Warren, M. E., & Gastil, J. *J. Polit.* **77**, 562–574 (2015).

## 6. 6.

OECD. *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave* (OECD, 2020).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02006-7>

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

- Perspective
- [Published: 08 July 2021](#)

# After the pandemic: perspectives on the future trajectory of COVID-19

- [Amilio Telenti](#) ORCID: [orcid.org/0000-0001-6290-7677](https://orcid.org/0000-0001-6290-7677)<sup>1,2</sup>,
- [Ann Arvin](#) ORCID: [orcid.org/0000-0001-5404-1498](https://orcid.org/0000-0001-5404-1498)<sup>1</sup>,
- [Lawrence Corey](#) ORCID: [orcid.org/0000-0002-2179-2436](https://orcid.org/0000-0002-2179-2436)<sup>3</sup>,
- [Davide Corti](#) ORCID: [orcid.org/0000-0002-5797-1364](https://orcid.org/0000-0002-5797-1364)<sup>4</sup>,
- [Michael S. Diamond](#) ORCID: [orcid.org/0000-0002-8791-3165](https://orcid.org/0000-0002-8791-3165)<sup>5,6,7</sup>,
- [Adolfo García-Sastre](#) ORCID: [orcid.org/0000-0002-6551-1827](https://orcid.org/0000-0002-6551-1827)<sup>8,9,10,11,16</sup>,
- [Robert F. Garry](#).<sup>12</sup>,
- [Edward C. Holmes](#) ORCID: [orcid.org/0000-0001-9596-3552](https://orcid.org/0000-0001-9596-3552)<sup>13</sup>,
- [Phillip S. Pang](#) ORCID: [orcid.org/0000-0002-8811-9801](https://orcid.org/0000-0002-8811-9801)<sup>1</sup> &
- [Herbert W. Virgin](#) ORCID: [orcid.org/0000-0001-8580-7628](https://orcid.org/0000-0001-8580-7628)<sup>1,14,15</sup>

*Nature* volume 596, pages 495–504 (2021)

- 45k Accesses
- 564 Altmetric
- [Metrics details](#)

## Subjects

- [SARS-CoV-2](#)
- [Viral infection](#)

## Abstract

There is a realistic expectation that the global effort in vaccination will bring the pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) under control. Nonetheless, uncertainties remain about the type of long-term association that the virus will establish with the human population and, in particular, whether coronavirus disease 2019 (COVID-19) will become an endemic disease.

Although the trajectory is difficult to predict, the conditions, concepts and variables that influence this transition can be anticipated. Persistence of SARS-CoV-2 as an endemic virus, perhaps with seasonal epidemic peaks, may be fuelled by pockets of susceptible individuals and waning immunity after infection or vaccination, changes in the virus through antigenic drift that diminish protection and re-entries from zoonotic reservoirs. Here we review relevant observations from previous epidemics and discuss the potential evolution of SARS-CoV-2 as it adapts during persistent transmission in the presence of a level of population immunity. Lack of effective surveillance or adequate response could enable the emergence of new epidemic or pandemic patterns from an endemic infection of SARS-CoV-2. There are key pieces of data that are urgently needed in order to make good decisions; we outline these and propose a way forward.

[Download PDF](#)

## Main

Early in 2020, the world observed a sharp increase in the reported number of SARS-CoV-2 infections. The rapid accumulation of cases contrasted not only with the historical numbers of the SARS-CoV outbreak in 2003, but also with the numbers from the pandemic in 2009 caused by influenza H1N1, with the caveat that perhaps cases of an H1N1 infections were underdiagnosed (Fig. 1). The pattern and impact of the pandemic revealed flaws in the worldwide response to the infection—some local of which were in nature, whereas others were more systematic across many different countries.

**Fig. 1: Ominous signs in the early days of the pandemic.**

---

 **figure1**

---

Towards the end of January 2020, there was an alarming increase in the reported number of SARS-CoV-2 infections (red) that contrasted not only with the growth of historical cases of SARS-CoV in 2003 (purple; counts starting from 19 March 2003), but also with growth curves for the 2009 swine H1N1 infection (blue; counts starting from 24 April 2009). However, each pandemic was probably broader than currently estimated. Data are from Github (<https://github.com/CSSEGISandData/COVID-19>) and the World Health Organization.

With the ongoing deployment of several highly effective SARS-CoV-2 vaccines in many countries, there is an expectation that this virus will disappear. However, two reasons temper our hope in reaching this conclusion: patchy vaccine coverage due to disparities in global access to vaccines and vaccine hesitancy, and vaccines may not always block virus transmission (despite reducing the burden of disease). In addition, although mass vaccine deployment may signal the end of the pandemic, the end of the pandemic does not necessarily equate to the end of SARS-CoV-2. Thus, it is critical to consider what the new equilibrium between humans and this virus and its evolutionary descendants might be. The goal of this Perspective is to discuss the probable transition to a new phase of SARS-CoV-2 infection in humans as an endemic pathogen, perhaps with intermittent epidemic peaks (Box 1). We base our assessment on ongoing data from the COVID-19 pandemic and observations from previous epidemics. We highlight the role of the dynamic interactions between changes in population immunity and ongoing viral evolution and immune escape in shaping the future association of

SARS-CoV-2 with humans. We also discuss the possibility that the virus will retain considerable virulence long term. We believe that a thorough understanding of this transition period, and informed guesses about the future of the pandemic are necessary to inform the next steps for public health. It is with that goal in mind that we identify key gaps in our current knowledge and tools with the hope of refining our response as well as guiding scientific initiatives.

## Box 1 Definition of terms

Definitions are adapted in part from  
<https://www.cdc.gov/mmwr/preview/mmwrhtml/su48a7.htm>.

**Endemic disease** A disease that is constantly present with an incidence that waxes or wanes over a relatively prolonged period (often years of decades).

**Epidemic** Occurrence of a disease in a pattern that is clearly in excess of normal expectations. Can also refer to a new disease occurring regionally without evolution into a pandemic.

**Pandemic** An epidemic in which a disease spreads worldwide, crossing international boundaries and spreading between continents.

**Transmissibility** The likelihood that a pathogen will spread from an infected individual to an uninfected individual.

**Virulence** The capacity to cause severe illness once the pathogen infects a host.

**Fitness** Reproductive success, in this context the capacity of a virus to produce infectious progeny in a given environment.

**Control** An acceptable reduction in a disease in the setting of ongoing epidemic or endemic transmission.

**Elimination of disease** Diminution to zero of a disease in the setting of ongoing epidemic or endemic transmission.

**Elimination of infection** Diminution to zero of an infection in the human population. This goal is particularly difficult to attain when there are reservoirs of zoonotic transmission in contact with humans or vector species, as is observed for bird species and the transmission of influenza.

**Eradication** Permanent reduction to zero of the worldwide incidence of an infection caused by a specific agent as a result of deliberate efforts; intervention measures are

therefore no longer needed.

**Extinction** Absence of a pathogen in humans, animal reservoirs or laboratory sources. Extinction has not been attained for any pathogen as stocks of smallpox and rinderpest are still held in some laboratories.

## Observations from previous pandemics

We believe that it is pertinent to use observations from past infectious disease epidemics to help to predict what the evolutionary future of this pathogen could look like. Key questions include whether COVID-19 will become a familiar but high impact seasonal disease similar to influenza and whether SARS-CoV-2 will become more or less virulent than it is currently. Such comparisons are not as straightforward as they might initially seem. In contrast to the coronaviruses that cause the common cold (HCoV-229E, HCoV-HKU1, HCoV-NL63 and HCoV-OC43), SARS-CoV-2 has a higher virulence, yet it also differs from the even-more-serious coronaviruses severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) in that asymptomatic transmission is frequent. Thus, comparisons to other coronaviruses do not enable a definitive prediction of the future behaviour of SARS-CoV-2. At a first glance, SARS-CoV-2 seemingly has a capacity to evolve that outstrips that seen in the other human coronaviruses. We do not know whether this reflects a lack of comparable data for the other viruses that have entered the human population long ago, a recent zoonotic origin that has resulted in a strong selection pressure for adaptation to transmission and/or immune evasion in the human host. Indeed, recent studies have shown that seasonal coronaviruses (such as HCoV-229E) have also experienced antigenic evolution in recent decades<sup>1</sup>. The overall uncertainty of these parameters makes it difficult to accurately predict the future post-pandemic equilibrium between SARS-CoV-2 and the human population.

A more meaningful comparison can be made with the emergence of new human influenza viruses, particularly the H1N1 influenza A virus that caused the global pandemic of 1918/1919. In the Northern Hemisphere, the 1918/1919 pandemic was associated with a relatively mild wave in the spring followed by a much more severe wave in the autumn. Although an infection fatality rate of around 2% is commonly cited for the devastating autumn wave, the accuracy of this number is difficult to assess<sup>2</sup>. However, both of these waves fell outside of the usual human influenza season (beginning in March and September, respectively), suggesting that the climatic factors that probably drive seasonality are less important when a new virus enters a very large population of susceptible hosts with little or no pre-existing immunity to the pathogen<sup>3,4</sup>. The same appears to be true of SARS-CoV-2. Although there have been suggestions of emerging seasonality, the reality is that the current changing patterns of the incidence of COVID-19 could better reflect the differing extent and timing of non-

pharmaceutical interventions such as social distancing. Major ‘summer’ (or tropical climate) outbreaks have been reported in such locations as Brazil, India and South Africa.

Although it is possible that SARS-CoV-2 may eventually evolve into a winter seasonal virus such as influenza and the common-cold coronaviruses, this may not occur until there is more widespread population immunity and fewer susceptible hosts in regions that have the optimal climatic conditions for transmission<sup>3</sup>. If SARS-CoV-2 does not become mostly seasonal, implications for the timing of vaccination and vaccine booster campaigns will be considerable.

## Viral evolution, transmission and disease severity

In general, transmission of respiratory viruses is mediated by replication in, and shedding from, the upper respiratory tract, whereas severe disease is associated with the invasion of and replication in the lower respiratory tract. Mutations that increase virus replication in both respiratory sites could be selectively favoured if they not only increase the transmissibility but may also result in higher virulence, causing more-severe disease. Notably, mutations that increase replication only in the upper respiratory tract could be selected based on higher transmissibility but may decrease virulence. Indeed, experiments in ferrets with avian influenza H5N1 virus (which is highly virulent, but poorly transmissible in humans) gave rise to viruses with increased transmissibility and decreased virulence. This is due to changes in the receptor specificity that favour replication in the upper respiratory tract to the detriment of replication in the lower respiratory tract<sup>5</sup>. However, in the case of SARS-CoV-2, mutations that further optimize the use of the human angiotensin-converting enzyme 2 (ACE2) as a virus receptor (which is present in both the upper and lower respiratory tract) or alter the capacity of co-receptors to influence tropism and infection are likely to increase both transmission and virulence<sup>6,7</sup>. By contrast, mutations that increase replication at 33 °C, the temperature of the human upper respiratory tract, while decreasing replication at 37 °C, the temperature of the lower respiratory tract, are expected to increase transmission but decrease virulence. A ‘wild card’ mutation that—for example—enables the evasion of innate immunity could have profound effects on both transmission and virulence or even the nature of disease. The likelihood that one or all of these changes may occur, or have already occurred, is not possible to predict with certainty given the paucity of data on the status of the virus and disease worldwide.

What about the evolution of SARS-CoV-2 during the pandemic? A reasonable expectation early in the pandemic was that the virus could evolve to develop increased transmissibility, reflecting adaptations to propagation in the new human host. Such a process probably occurred during the large outbreak of Ebola virus in western Africa,

resulting in the fixation of mutations that increased affinity of the virus for the human cellular receptor<sup>8</sup>. It is now clear that the D614G substitution in SARS-CoV-2 increased transmission, leading to its emergence as a dominant strain; in addition, the mutations in the recent B.1.1.7 (Alpha) and B.1.617.2 (Delta) variants have further increased transmission in humans, enabling these variants to successively become dominant in every region in which they have been introduced<sup>9</sup>. Compared with influenza virus, SARS-CoV-2 has shown an unprecedented capacity to evolve global variants that outcompete regional variants in extremely short time windows and before considerable selective pressure owing to pre-existing immunity. However, whether the SARS-CoV-2 virus will eventually evolve into a more virulent virus is less predictable, as virulence is not necessarily a selectable phenotypic trait that increases the fitness of the virus.

Lessons for understanding the evolution of SARS-CoV-2 can again be tentatively drawn from the 1918/1919 influenza pandemic. The autumn wave of the influenza pandemic was associated with far higher virulence than the spring wave. Thereafter, the 1918/1919 virus continued to cause epidemics until the 1950s when it was replaced by a novel zoonotic H2N2 influenza A virus. Importantly, some of these later seasonal epidemics of H1N1 were also associated with relatively high virulence due to ongoing antigenic drift: two of the worst outbreaks of influenza in the twentieth century in terms of excess deaths occurred in 1928–1929 and 1934–1936, respectively, and were due to descendants of the 1918/1919 H1N1 pandemic influenza virus<sup>10</sup>. Moreover, it is well-documented that bacterial co-infections increase the severity of the disease caused in humans by influenza virus. In this respect, we still do not know well the consequences of co-infections of SARS-CoV-2 with other human pathogens, including influenza virus, the circulation of which was markedly decreased during the SARS-CoV-2 pandemic, but that is likely to be a prevalent human respiratory pathogen when many of the measures adopted to mitigate the spread of SARS-CoV-2 in humans are lifted.

The severity of disease caused by SARS-CoV-2 is bound to decrease with increasing population immunity. Even in individuals who are not fully protected from infection by vaccination or previous infection, pre-existing immunity is likely to reduce the severity of symptoms after infection, and to prevent future severe pandemics arising from antigenically related coronaviruses that are circulating in bats and other possible animal reservoirs. Nevertheless, the evolution of the virus to the low level of virulence seen in common-cold coronaviruses may not occur or may take several decades to manifest. More broadly, many years of data and theory have told us that it is probably naive to make strong predictions about the evolution of virulence in any complex system<sup>11</sup>.

## SARS-CoV-2 will evolve and evade immunity

The emergence of new virus variants that imperil the control of the pandemic is a prominent theme in public discourse. These new variants are defined by the US Centers for Disease Control and Prevention as variants of interest, variants of concern or variants of high consequence (<https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html#Interest>). At present, the B.1.1.7 (Alpha), B.1.351 (Beta), P.1 (Gamma) and B.1.617.2 (Delta) variants that are circulating in the USA and worldwide are variants of concern. In the context of the evolving pandemic, there is a need to review the expectation of how SARS-CoV-2 could evolve to a form that might derail the control of the pandemic or alter the nature of an ensuing endemic or combined endemic/epidemic phase.

The *Coronaviridae* family is characterized by relatively high replication fidelity compared with other RNA viruses, defined by the processivity of their polymerases as required by their exceptionally large genomes<sup>12</sup>. On this basis, there had been opinions that SARS-CoV-2 evolution would be limited, in turn ensuring the durability of vaccines and therapeutics and supporting optimism that population immunity can end the pandemic. However, although—on average—SARS-CoV-2 evolves (perhaps 3–4 times) more slowly than influenza virus<sup>13</sup>, the virus is accumulating mutations more rapidly than might be expected given its relative replication fidelity with approximately two mutations fixed per month<sup>14</sup> and far higher rates of change are seen in some of the variants of concern. Coronaviruses also have a high rate of viral RNA recombination<sup>15</sup>; thus, humans who are infected with two variants of SARS-CoV-2 may acquire multiple mutations from both variants at the same time. We also cannot exclude potential recombination events in the future between SARS-CoV-2 and other human coronaviruses. In addition, prolonged infections of immunosuppressed individuals who cannot efficiently clear the infection may provide an opportunity for the accumulation of multiple mutations. Furthermore, vaccination may not be effective in individuals with compromised immunity<sup>16</sup>. Thus, there may be stochastic events in the emergence of future variants based on the infection of a limited number of immunosuppressed or vaccine-unresponsive persons. In the USA alone there are an estimated 10 million individuals with potential limitations to their immune response.

Why are we witnessing the emergence of variants despite the relatively fastidious replication machinery of coronaviruses? The speed of evolution of a viral pathogen is not only dependent on the background mutation rate, but also on the virus generation time, the duration of infection, the number of variants that develop during the infection of an individual, the structural and functional constraints in specific regions of viral proteins, and the extent and strength of natural selection acting on the virus. In addition, the greater the number of individuals infected with the virus, the larger the pool and diversity of mutant viruses that is generated. Although transmission events between two hosts routinely generate bottlenecks that purify away most of the low-frequency mutant viruses, large numbers of transmission events may enable the

transmission of a more fit virus, with the global spread of B.1.1.7 (Alpha) and now the spread of B.1.617 (Delta) in India serving as important examples<sup>[17,18,19,20](#)</sup>.

Functional domains in viral proteins that can accept mutations without losing their overall structure and function are sites of potentially selectable mutations. The region of the spike protein of SARS-CoV-2 that interacts with the human ACE2 receptor exhibits particular structural and functional plasticity<sup>[21,22](#)</sup>. With new selection pressures generated by vaccines or immunity to natural infection—or by the use of antivirals—the possibility of viral adaptations to overcome immune and/or antiviral pressure will probably be a continuing reality. There is a risk of viral diversification in the currently uncontrolled or incompletely controlled pandemic in many regions of the world. In this regard, the level of knowledge of the effect of mutations outside of the viral spike protein is in its infancy, which limits the ability to predict the evolutionary pathways that the virus will follow in the future.

A rapid transition to an endemic phase may decrease the number of circulating virus variants by limiting the extensive exploration of the fitness landscape that takes place during the pandemic phase. Hence, the nature of the future equilibrium between SARS-CoV-2 and humans relies on both the speed and inclusivity of responses to the pandemic across diverse geographies and cultures, as this directly influences the speed of the emergence of problematic variants.

Understanding the transition to an endemic phase, with potential seasonal peaks, would benefit from new tools that can forecast what virus variants may emerge and spread. Spreading variants can be predicted to some extent from epidemiological and biological data, including ACE2 binding measured in the context of deep-scanning mutagenesis of the viral spike protein<sup>[23](#)</sup>. Immune escape is beginning to drive the spread of virus variants at a time in the pandemic when high levels of vaccine- and infection-induced immunity have not yet been achieved worldwide<sup>[24](#)</sup>. Full containment of the pandemic minimizes the likelihood of SARS-CoV-2 adapting to the host by reducing the length of transmission chains<sup>[25](#)</sup>. This appears unlikely to be the case unless very high levels of vaccination can be accomplished worldwide.

## Interspecies spread

SARS-CoV-2 does not only infect humans; it also has pantropic properties<sup>[26](#)</sup>. SARS-CoV-2 infections have been established in a range of animal species, including bats, cats, dogs, ferrets, hamsters, deer mice, otters, white-tailed deer and various nonhuman primates<sup>[27,28,29](#)</sup>. Zoonotic transmission from humans to animals has been documented in farmed mink, dogs and cats<sup>[29,30,31,32](#)</sup> as well as in lions and tigers in zoos<sup>[33](#)</sup>. Thus, the host range of SARS-CoV-2 extends to a variety of mammalian species, including

those maintaining large populations in the wild. Virus evolution can occur in animal hosts, generating a suite of genomic changes in addition to those seen during human-to-human transfer (Fig. 2). As expected, mutations related to species specificity occur in the receptor-binding domain (RBD) of the spike protein and are important, because changes to this region may enable immune escape and/or confer a transmission or fitness advantage. Variation in the N-terminal domain or in or near the furin-cleavage site represent other mutational hotspots in the spike protein that are common to variants of concern and after interspecies transfers (Fig. 2). However, the effect of mutations outside of the spike protein in these interspecies adaptations has not yet been examined and is a wild card that may limit the predictability of the course of the pandemic.

**Fig. 2: Mutations arising in the spike protein of SARS-CoV-2 upon sustained transmission between humans and between humans and animals.**

 figure2

Black thick lines show mutations shared between infection in humans and animals. Thin lines indicate mutations limited to infection in humans or animals. Homology modelling of the SARS-CoV-2 spike used reference sequence QHD43416.1 and a closed pre-fusion configuration of the spike trimer (Protein Data Bank (PDB) 6VXX)<sup>81</sup> as a template. NTD, N-terminal domain; VOC, variant of concern. The figure is adapted from <https://virological.org/t/mutations-arising-in-sars-cov-2-spike-on-sustained-human-to-human-transmission-and-human-to-animal-passage/578>.

Infected animals can be the source of two evolutionarily related problems. First, upon animal infection, the human virus can undergo evolution that could introduce adaptive mutations. An example of such an event occurred in mink in Denmark and the Netherlands<sup>30,34</sup>. Human-to-animal transfer resulted in the introduction of an adaptive substitution, Y453F, and the subsequent outbreak of this variant in humans (referred to as the ‘mink variant’, B.1.1.298). The Y453F substitution is in the RBD of the spike protein and increases the affinity of the spike protein for human ACE2 compared with the original SARS-CoV-2 strain, suggesting an avenue for the enhanced transmission or pathogenicity of SARS-CoV-2<sup>35</sup>. Second, an animal coronavirus infection in animals carrying SARS-CoV-2 may pose a serious risk for the generation of hybrid viruses through recombination between viral genomes. These hybrid viruses could have new properties related to immune evasion or virulence. Genomic recombination—which is frequently observed in coronaviruses<sup>15</sup>—may have played a part in the evolution of SARS-CoV-2 (<https://virological.org/t/recombinant-sars-cov-2-genomes-involving-lineage-b-1-1-7-in-the-uk/658>), including potential recombination with diverse coronaviruses that are present in a variety of animal species. New variants that can be transmitted back to humans in an interspecies ‘ping pong’ of infections could contribute to further SARS-CoV-2 diversification, as it is the case for influenza A viruses<sup>36</sup>. Infection and propagation of SARS-CoV-2 in nonhuman species could lead to sequence alterations, interspecies transmission and adaptations that could compromise human immunity or affect virulence and that could diminish binding to monoclonal antibodies that are in clinical use<sup>37,38</sup>. As an example, one RBD substitution—N501Y—that occurs in B.1.1.7 (Alpha), B.1.351 (Beta), P.1 (Gamma) and other emerging variants of concern enables the productive infection of laboratory mice and possible expansion of the host range to wild mice<sup>39,40,41</sup>. This mutation also diminished neutralization by a monoclonal antibody in clinical trials<sup>42</sup>.

Establishment of SARS-CoV-2 in other species could provide a refuge for the virus to re-emerge in human populations in an evolutionarily distinct form—for example, upon waning of vaccine coverage or diminished natural or vaccine-induced immunity that occurs over time. It is also possible that after decades of separate circulation of SARS-CoV-2 in humans and animals, the human viruses will have diverged antigenically owing to immune pressure, but the animal viruses may not have. This could lead to a population of young individuals born in post-pandemic years with no pre-existing immunity against the old SARS-CoV-2 strains who are therefore susceptible to infection with the animal SARS-CoV-2 viruses that are antigenically related to the original SARS-CoV-2 pandemic strain. In fact, this is the most likely explanation of the pandemic in 2009 caused by a swine influenza virus, which is a descendant of the human H1N1 viruses from the 1918/1919 pandemic virus, but which is antigenically related to the H1N1 viruses that circulated in humans in the beginning of the twentieth century<sup>43</sup>. The potential for such events demands active research into possible susceptible secondary reservoir hosts, and the development of therapeutic and

prophylactic interventions that are agnostic to variations in the virus sequence. It is important to recognize that such solutions need to be on the shelf against the possible emergence of a highly problematic strain of SARS-CoV-2, as the speed with which the virus has spread during this pandemic shows the limitations of even an exceptionally fast response in the development of vaccines or therapeutics<sup>44</sup>.

## The role of vaccines and the correlates of protection

The remarkably rapid development of safe and highly effective vaccines that mitigate the burden of COVID-19 is an historic achievement. Nevertheless, fundamental questions remain as to the mechanism(s) of protection against the disease, the extent of protection against asymptomatic infection and the duration of vaccine-induced humoral and cellular immunity. The effects of potential differences between the immunity induced by the vaccines compared with natural infection and between different COVID-19 vaccines also remain unclear.

Policies to guide vaccine campaigns in the fight against any virus benefit when a test of immunity that correlates with vaccine efficacy can be identified. Antibody assays that measure the neutralization of antigen binding are typically used to determine the rates of seroconversion after vaccine administration, but these may not be fully useful as correlates of protection in an individual because antibodies also restrict viral infection according to their effector functions<sup>45,46,47</sup> and CD4<sup>+</sup> and CD8<sup>+</sup> T cell responses are critical for antiviral immunity<sup>48</sup>. Recently, CD4<sup>+</sup> T cells have been reported to shape the development of humoral and CD8<sup>+</sup> T cell responses to the spike protein after vaccination with an mRNA vaccine<sup>49</sup>. Assays for neutralizing antibodies elicited by the spike protein have been the primary measure of immunity induced by the SARS-CoV-2 vaccines. Most neutralization assays are variable across assay systems and cell lines<sup>50,51</sup>, which limits their use to define a correlate that spans clinical studies of different vaccines, and they do not distinguish between responses to different epitopes on the spike protein, some of which are highly immunodominant<sup>52</sup>. The establishment of an international standard by the World Health Organization to enable the normalization of data from different assays is an important step towards addressing this issue ([www.who.int/teams/blueprint/covid-19](http://www.who.int/teams/blueprint/covid-19)). A recent model that relates the efficacy of immunization with mRNA, adenoviral vector and other SARS-CoV-2 vaccines to neutralizing titres supports that these titres have predictive value<sup>53</sup>.

Notably, as observed for influenza A virus, regions of the spike protein of SARS-CoV-2 that elicit the most potent neutralization responses are also the most variable in emerging variants of concern. Indeed, vaccine-induced and natural immunity—as measured by neutralization assays—is reduced against some variants. Although vaccines are effective against current variants, the impact of such changes on the

prevention of COVID-19, especially severe illness, requires continuous assessment. At a minimum, the longevity of protection is likely to be affected, assuming that antibodies are the primary defence mechanism. The level of total antibody in serum after natural infection decreases with a half-life of about 50–100 days<sup>52,54,55</sup> and vaccine-induced antibodies also peak shortly after immunization. However, waning antibody levels after vaccination cannot be equated to renewed susceptibility to disease because the immune system has been primed to rapidly mount memory B and T cell responses that mitigate the consequences of repeated infection.

More information about the mechanisms of immune protection against COVID-19 that are poorly understood, such as mucosal immunity and innate immune barriers, as well as the impact of immune protection on the transmission of SARS-CoV-2 is essential to inform vaccine policies to control the pandemic and for the need for boosters or a next-generation of vaccines as the pandemic transitions to an endemic or epidemic pattern.

## Lessons from vaccination against respiratory viruses

The challenges for SARS-CoV-2 vaccine programmes have instructive similarities and potential differences from experiences with other viral vaccines. Vaccines against 17 viral pathogens are approved in the USA. The effects of vaccines can be divided into control, elimination of disease, elimination of infection, and eradication and extinction (Box 1). The first aim of SARS-CoV-2 vaccine programmes is the rapid control of new infections in as many geographical regions as possible, an outcome that depends on the widespread—both within counties and worldwide—uptake of vaccines that effectively reduce transmission. Control also requires broad access to rapid diagnostic methods and surveillance to detect ongoing transmission. Of note, highly effective vaccines can achieve the elimination of disease even if the infection is not eliminated. Smallpox is the only example of vaccine-mediated eradication of an infection in humans, which required a global initiative that combined high levels of vaccine coverage, active surveillance and rapid and targeted vaccination efforts in regions in which outbreaks occurred. The importance of combining sustained immunization campaigns with effective surveillance and rapid molecular diagnosis are hard-learned lessons, as illustrated by the failure to eliminate polio despite the lack of an animal reservoir and the availability of two effective vaccines.

Experiences with viruses that are transmitted by respiratory droplets and/or aerosols but that have persisted in the human population despite the availability of effective vaccines point to obstacles and inform about SARS-CoV-2 vaccine strategies.

Considerable achievements in the elimination both of the disease and infections caused by such viruses—including measles, mumps and rubella—have been made in many, but often not in all, geographical regions. Even in optimal circumstances,

communities that are under-vaccinated owing to poor access or that resist vaccination serve as reservoirs for the reintroduction of pathogens and disease outbreaks when vaccination coverage falls at the population level.

A lesson from experience with measles, mumps, rubella and varicella vaccines is the importance of two-dose regimens. Because respiratory transmission is efficient, even a low incidence of primary vaccine failure, defined as no seroconversion to the first dose, leaves enough susceptible individuals in the population to support outbreaks. Secondary vaccine failure, defined as disease despite seroconversion, occurs with single-dose regimens that elicited antibodies after one dose, as was observed with the varicella vaccine and also occurs in some cases with two doses of varicella or mumps vaccines<sup>56</sup>. Under these circumstances, the vaccinated individuals benefit from protection against severe illness, but breakthrough infections remain a source of transmission. Limited information about SARS-CoV-2 vaccines suggests that this pattern of breakthroughs, albeit with markedly reduced severity, may occur but with a frequency that is as yet undetermined at the population level.

In the case of measles, even though the viral fusion protein is genetically stable—in contrast to the spike protein of SARS-CoV-2—virus entry is highly efficient and population immunity of 92–95% is required to eliminate transmission, as confirmed by recent outbreaks of measles despite high vaccine coverage<sup>57</sup>. Measles has a very high basic reproductive number ( $R_0$ ) with transmission from one case to fifteen susceptible individuals, whereas the  $R_0$  value for SARS-CoV-2 has been modelled at 2.2–5.7 for the Wuhan reference strain<sup>58</sup>. As long as estimates of transmission of variants of concern remain below measles, control may occur with lower levels of population immunity. Notably—and in contrast to measles—the occurrence of asymptomatic SARS-CoV-2 infections will interfere with rapid outbreak recognition and provide an avenue for spread within populations and across geographical regions. In this regard, SARS-CoV-2 is more similar to polio, or rubella, for which a strategy of universal and repeated vaccination campaigns rather than outbreak control has been necessary to eliminate congenital disease<sup>59</sup>.

Childhood vaccines also demonstrate the difficulty of defining immune correlates of protection against breakthrough infections that re-introduce the virus into the community. For example, receiving two doses of measles vaccine correlates with protection even though neutralizing antibody titres may be low and not be boosted by additional doses<sup>57</sup>. Occurrence of varicella in vaccinated adults appears to reflect lower cellular immune responses, which are not measured by neutralization assays<sup>60</sup>. The occurrence of mumps in highly vaccinated groups is attributed to waning immunity. Administration of a third dose of vaccine appears to reduce the spread in outbreaks, but an antibody-based immune correlate that would allow targeted revaccination of those at risk has not been established despite extensive study<sup>61</sup>. These

experiences predict that maintaining the benefits of SARS-CoV-2 vaccine programmes will require not only the monitoring of the duration of immunity as determined by serological assays, but also the ongoing local surveillance for infections in vaccinated populations coupled with the tracking of vaccine coverage rates within the community so that gaps can be rapidly addressed. Notably, the only way to identify correlates of protection will be the consistent application of robust and reproducible assays of both T- and B-cell-mediated immunity in vaccine recipients under conditions in which exposures can be documented. Such an effort will be technically challenging for vaccine producers and studies of one vaccine may not inform immune correlates for a different type of vaccine. This presents a major unaddressed challenge to understanding the mechanisms of vaccine protection and the new equilibrium between humans and SARS-CoV-2 that is currently evolving as vaccine coverage is extended. A large study on college campuses is designed to address this challenge (ClinicalTrials.gov NCT04811664). In addition, elimination of SARS-CoV-2 is unlikely without immunization of children who may be vectors for asymptomatic transmission. Decades of childhood vaccine experience documents that even with high coverage—sufficient to eliminate disease and infection in almost all individuals—protection of unvaccinated people by herd immunity is not guaranteed when reintroductions occur. Reintroductions of SARS-CoV-2 have the added challenge that these may result from either human or zoonotic sources or both, whereas endemic childhood viruses do not have animal reservoirs.

Large-scale SARS-CoV-2 vaccine programmes have already had an enormous impact on the burden of COVID-19 disease. Although SARS-CoV-2 may become endemic without being associated with severe disease<sup>62</sup>, viruses with pathogenic potential may continue to circulate, causing local outbreaks or more widespread epidemics, and requiring vaccine campaigns and possibly ring prophylaxis<sup>63</sup> using monoclonal antibodies to eliminate the disease.

## The paradigm of antigenic drift and shift in influenza

The experience with influenza A and B vaccines offers a perspective for COVID-19 vaccines when transmission is associated with the capacity of the target virus to undergo seasonal antigenic drift and periodic antigenic shift, which may be caused by recombination in the case of SARS-CoV-2<sup>64</sup> (as opposed to segment re-assortment in the influenza viruses). Antigenic drift is common to all four antigenically distinct circulating influenza A and B viruses. Most of the key mutations that lead to antigenic drift are located in the globular head of the haemagglutinin (HA), which comprises the receptor-binding motif—which, similar to SARS-CoV-2, appears to be the most structurally and functionally plastic region<sup>21</sup>—possibly enhancing the efficacy of antibody selection of variants that may evade natural or vaccine-induced immunity.

The degree of antigenic variation or ‘antigenic distance’ between the HA and neuraminidase proteins of influenza is the basis for needing to update the composition of influenza vaccines frequently. This distance is typically measured in an haemagglutination inhibition assay using ferret or human antisera generated against the influenza vaccine and circulating strains<sup>65</sup>. Whenever the fold change in titres of antisera in the haemagglutination inhibition assay generated against vaccine strains and tested against circulating strains exceeds 8–10-fold, it typically signals the need to ‘upgrade’ the vaccine composition (Fig. 3a, b). This exercise is performed for all four viruses that are part of the influenza vaccine mixture—currently H1N1, H3N2, influenza B Yamagata lineage and influenza B Victoria lineage. The need to update vaccines results from the fact that the most variable region of HA (and neuraminidase) is also the immunodominant region and, reciprocally, because the response to more conserved regions elicits antibodies endowed with poor neutralizing activity. It is worth noting that for SARS-CoV-2, the introduction of more than ten mutations in some of the variants of concern such as B.1.351 (Beta)—which has led to a reduced neutralizing titre of antisera from vaccinated donors of approximately tenfold (Fig. 3d, f)—is comparable to the extent of antigenic drift in influenza A and B viruses that typically requires a change in the viruses selected for vaccine production. Indeed, several vaccines were shown to provide modest efficacy against the B.1.351 (Beta) variant<sup>66,67</sup>. As a consequence of the reduced efficacy, the composition of SARS-CoV-2 mRNA vaccines—such as the one developed by Moderna—was recently adapted to match the B.1.351 (Beta) variant and clinical trials testing the immunogenicity of such vaccines are underway (Clinical Trials.gov NCT04785144). Antigenic drift was also shown to occur in human endemic coronaviruses such as HCoV-229E<sup>1</sup> (Fig. 3e, f).

**Fig. 3: The antigenic drift paradigm of influenza viruses and coronaviruses.**

---

 **figure3**

**a**, Structural models of the HA protein of influenza A virus H1N1 from 1999 and 2006 isolates (PDBs 5C0S and 6CF7, respectively), in which mutated residues are highlighted in red and are shown on a single monomer of each trimer. Loss of neutralization titres against parental and drifted strains of H1N1, as previously described<sup>65</sup>. **b**, Protein distance of major glycoproteins of influenza viruses. Dendograms show the protein distance of HA amino acid sequences that are representative of the variability of each virus family. Shown are HAs from all 18 influenza A virus subtypes (multiple strains shown for H1, H3 and H5) and of HAs from influenza B virus from the ancestral virus from 1940 and the two Victoria and Yamagata lineages. Scale bar, 0.100. **c**, Timeline of influenza pandemics. **d**, **e**, The spike ectodomain of SARS-CoV-2 (**d**; PDB 6VYb; spike in the open conformation) and the spike ectodomain of HCoV-229E (**e**; PDB 6U7H), in which mutated residues are highlighted in red and are shown on a single monomer of each trimer. Loss of neutralization titres against parental and drifted strains of SARS-CoV-2 and HCoV-229E by serum antibodies elicited against the parental strains, as determined previously<sup>1,82</sup>. **f**, Protein distance of spike glycoproteins of human and animal

coronaviruses. Sarbecoviruses are shown according to the phylogenetic definition of RBD clades<sup>83</sup>. Scale bar, 0.100. **g**, Timeline of the emergence of human coronaviruses. The emergence of common-cold coronaviruses is approximate and based on molecular clock dating<sup>84</sup>. Highlighted in red in **b** and **f** are viral strains used in **a**, **d** and **e**.

It remains to be established whether the evolution of SARS-CoV-2 will be accelerated by increasing immunity or whether—by contrast—the reduced circulation of the virus that is expected to occur as a consequence of widespread vaccination may slow down the accumulation of mutations<sup>68</sup>. In the case in which SARS-CoV-2 will become endemic and will continue to evolve, revaccination scenarios need to be envisioned. These may require the revaccination with the same vaccine or boosting/vaccination with a vaccine based on the most prevalent circulating variants. More difficult to predict is how often revaccinations will be needed and recommended for specific risk groups or for the general population.

As it is the case for influenza, it will be important to assess the possible influence of original antigenic sin in trapping the antibody response by the first response made to the parent antigen<sup>69,70</sup>. This is a phenomenon in which the immune response to subsequent infection or vaccination is biased towards responses imprinted in an individual's immune system by the persistence of memory B and T cells elicited by previous infections with related viruses. If this holds true for SARS-CoV-2, it might reduce the immunogenicity of vaccines against the variable sites of the spike protein, possibly boosting the response to the most conserved regions, which could be a potential beneficial outcome if cross-reactive antibodies have a protective role<sup>71</sup>. In this respect, even previous immunity against other human betacoronaviruses could shape subsequent immunity to SARS-CoV-2 infection or from vaccination<sup>72</sup>.

Another lesson from the evolution of influenza viruses is that multiple lineages of the same virus can co-exist and co-circulate. This is the case for the two co-circulating lineages of influenza B virus that originated from a common progenitor in the 1970s, which has led to the independently evolving Victoria and Yamagata lineages and resulted in a decision to convert the traditional trivalent influenza vaccine into a quadrivalent vaccine<sup>73</sup>. This decision was driven by the difficulty of predicting which of the two lineages would prevail during each season. The co-circulation of different SARS-CoV-2 lineages in the same or different geographical areas may complicate decisions on which lineages merit incorporation into new vaccines and whether these vaccines will need to evolve into multivalent formats that target several variants, like influenza vaccines. It is as yet undetermined how many strains of SARS-CoV-2 will need to be considered when planning an effective long-term vaccine strategy. It is worth noting that until now, a relatively limited set of mutations have independently emerged in multiple variants, pointing to a convergent and potentially constrained evolution of SARS-CoV-2 for immune escape.

A second, but no less important, aspect of the evolution of influenza viruses that may occur with SARS-CoV-2 is antigenic shift—that is, the introduction, through recombination, of antigenically novel forms for viral antigens. In the case of influenza, this involves the acquisition of new genome segments from zoonotic (particularly avian) viruses and has occurred at least four times in the past century: in 1918 (H1N1), in 1957 (H2N2), in 1968 (H3N2) and in 2009 (H1N1) (Fig. 3c). As a parallel, animal betacoronaviruses have already entered the human population five times, including in: 2003 (SARS-CoV), 2012 (MERS-CoV) and 2019 (SARS-CoV-2), and at some earlier time in the case of NL63-CoV and HKU1-CoV, with both MERS-CoV and SARS-CoV associated with severe disease (Fig. 3g). MERS-CoV has caused spill-over events from camels to humans since 2012 but has not evolved into a form associated with high levels of human-to-human transmission. There were exceptions, such as in South Korea, where a single imported case resulted in almost 200 infections in a hospital setting<sup>74</sup>. Therefore, the risk for MERS-CoV evolving into a more transmissible virus should not be underestimated. Notably, the level of sequence similarity between the spike proteins of SARS-CoV and SARS-CoV-2 is 76%, which is close to the 80% similarity between the pre-pandemic H1N1 strain (A/Solomon Island/03/06) and the pandemic H1N1 swine influenza strain (A/California/04/09). The risk for new sarbecoviruses to cause future zoonotic infections is considerable as this has already occurred twice in the past 20 years. Coronaviruses isolated from bats can efficiently multiply in human lung tissue<sup>75</sup>. This calls for aggressive development of countermeasures based on pan-reactive vaccines or therapeutics that can be stockpiled and be ready for deployment to avoid the health and economic devastation seen in this pandemic.

In summary, the very recent emergence of SARS-CoV-2 has not given much time to understand the role and consequences of antigenic drift and shift. These are critical analyses to establish the future needs and requirements for revaccination.

## Three possible scenarios of the future of COVID-19

The first—and most worrisome—scenario is that we will not gain rapid control of this pandemic and thus will face a future with ongoing manifestations of severe disease combined with high levels of infection that, in turn, could foster further evolution of the virus. Vaccinations and previous infection could achieve long-term herd immunity, but we will need a very broad application of vaccines worldwide combined with comprehensive disease surveillance by accurate and readily available diagnostic assays or devices<sup>76</sup>.

A second and more likely scenario is the transition to an epidemic seasonal disease such as influenza. Effective therapies that prevent progression of COVID-19 disease (for example, monoclonal antibodies that reduce hospitalization and death by 70–85%)

may bring the burden of SARS-CoV-2 infection to levels that are equivalent or even lower than influenza. However, we should remember that the annual mortality burden of influenza, in non-pandemic years, is estimated to be between 250,000 and 500,000 deaths, with up to 650,000 all-cause deaths globally, comprising around 2% of all annual respiratory deaths (two thirds among people who are 65 years and older)<sup>[77](#)</sup>. This is an extremely important health burden and equates to a relatively ‘optimistic’ view of the future of the COVID-19 pandemic.

A third scenario is the transition to an endemic disease similar to other human coronavirus infections that have a much lower disease impact than influenza or SARS-CoV-2. There is, however, limited data on the global burden of disease by common human coronaviruses<sup>[78](#)</sup> and as noted in above, it is not possible to predict with confidence whether further adaptations of SARS-CoV-2 to humans will increase or decrease its intrinsic virulence.

To better predict which scenario is likely to emerge and to better equip the world with an appropriate response, we propose several key questions that need to be answered and critical tools that need to be developed (Box 2). These comprise gaps in our knowledge in terms of epidemiology, immunology and virology, and missing surveillance, prophylactic and therapeutic tools.

This pandemic has shown both the importance of initiatives in individual countries and the interdependence of the world, and the necessity of global cooperation for pandemic control. It is the investment by a limited number of countries that has led to the biomedical discoveries that have brought forward the tools to interrupt the spread of the pandemic<sup>[79](#)</sup>. Yet, the lack of international structures for the implementation of these tools has brought into focus the disparities between advantaged and disadvantaged groups both within countries and between countries. This highlights the current inadequacies in healthcare delivery systems and access to new biomedical interventions<sup>[80](#)</sup>. Global health leaders will need to be vigilant with respect to the trajectory of SARS-CoV-2 in the near future while assessing the strategies and approaches used in the pandemic to develop more effective structures and processes to ensure a more effective and equitable response for the future.

## **Box 2 Current key gaps in developing an effective global response**

### **Research questions**

#### *Epidemiology*

- What are the effects of geographical and socioeconomic variations in vaccine coverage and disease on the ability to convert the pandemic to an endemic or

epidemic disease?

- What is the contribution of immunosuppressed populations to the rapid evolution of SARS-CoV-2?

### *Virology*

- What are the mechanisms by which viruses adapt to different hosts, thereby crossing species barriers?
- Is viral sequence evolution effectively reduced by vaccination?

### *Immunology*

- What are the correlates of protection for vaccines and natural immunity? The assessment of protection will require the coherent application of reproducible immunologic assays in populations to follow disease incidence and severity.
- What is the impact of antigenic drift?
- What are the criteria for the renewal or boosting of vaccines?
- What is the role of mucosal immunity in limiting viral shedding and preventing severe disease?

## **Tools and technologies**

### *Surveillance*

- Globally accessible diagnostics and deep-sequencing tools to establish continuous and sustained global surveillance of disease and variants.

### *Vaccines*

- Pan-sarbecovirus vaccines and monoclonal antibodies that will address both SARS-CoV-2 variants and the future introduction of pandemic coronaviruses into the human population.

### *Therapeutics*

- Next-generation therapeutics in the form of cheap oral antiviral agents.
- Long-acting monoclonal antibody prophylaxis for persons not likely to achieve effective vaccination.

- Addressing inequalities in pandemic healthcare and access worldwide to the most effective vaccines and therapeutics.

## References

1. 1.

Eguia, R. T. et al. A human coronavirus evolves antigenically to escape antibody immunity. *PLoS Pathog.* **17**, e1009453 (2021). **The historical evolution of human coronavirus 229E demonstrates its antigenic evolution and decreased neutralization potential.**

2. 2.

He, D. et al. Comparing COVID-19 and the 1918–19 influenza pandemics in the United Kingdom. *Int. J. Infect. Dis.* **98**, 67–70 (2020).

3. 3.

Baker, R. E., Yang, W., Vecchi, G. A., Metcalf, C. J. E. & Grenfell, B. T. Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. *Science* **369**, 315–319 (2020).

4. 4.

Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020). **A model of multiyear interactions between existing coronaviruses to project that recurrent wintertime outbreaks of SARS-CoV-2 will probably occur after the pandemic waves.**

5. 5.

Herfst, S. et al. Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* **336**, 1534–1541 (2012).

6. 6.

Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).

7. 7.

Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* **5**, 562–569 (2020).

8. 8.

Urbanowicz, R. A. et al. Human adaptation of Ebola virus during the west African outbreak. *Cell* **167**, 1079–1087.e5 (2016).

9. 9.

Davies, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, eabg3055 (2021).

10. 10.

Morens, D. M., Taubenberger, J. K. & Fauci, A. S. The persistent legacy of the 1918 influenza virus. *N. Engl. J. Med.* **361**, 225–229 (2009). **Descendants of the H1N1 influenza A virus that caused the pandemic of 1918/1919 have persisted in humans for more than 90 years and have continued to contribute their genes to new viruses, causing new pandemics, epidemics and epizootics.**

11. 11.

Geoghegan, J. L. & Holmes, E. C. The phylogenomics of evolving virus virulence. *Nat. Rev. Genet.* **19**, 756–769 (2018).

12. 12.

Smith, E. C., Sexton, N. R. & Denison, M. R. Thinking outside the triangle: replication fidelity of the largest RNA viruses. *Annu. Rev. Virol.* **1**, 111–132 (2014).

13. 13.

Bar-On, Y. M., Flamholz, A., Phillips, R. & Milo, R. SARS-CoV-2 (COVID-19) by the numbers. *eLife* **9**, e57309 (2020).

14. 14.

Duchene, S. et al. Temporal signal and the phylodynamic threshold of SARS-CoV-2. *Virus Evol.* **6**, veaa061 (2020).

15. 15.

Goldstein, S. A., Brown, J., Pedersen, B. S., Quinlan, A. R. & Elde, N. C. Extensive recombination-driven coronavirus diversification expands the pool of potential pandemic pathogens. Preprint at <https://doi.org/10.1101/2021.02.03.429646> (2021).

16. 16.

Grupper, A. et al. Humoral response to the Pfizer BNT162b2 vaccine in patients undergoing maintenance hemodialysis. *Clin. J. Am. Soc. Nephrol.* **16**, 1037–1042 (2021).

17. 17.

Zhou, B. et al. SARS-CoV-2 spike D614G change enhances replication and transmission. *Nature* **592**, 122–127 (2021).

18. 18.

Volz, E. et al. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell* **184**, 64–75.e11 (2021).

19. 19.

Frampton, D. et al. Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect. Dis.* [https://doi.org/10.1016/S1473-3099\(21\)00170-5](https://doi.org/10.1016/S1473-3099(21)00170-5) (2021).

20. 20.

Graham, M. S. et al. Changes in symptomatology, reinfection, and transmissibility associated with the SARS-CoV-2 variant B.1.1.7: an ecological study. *Lancet Public Health* **6**, e335–e345 (2021).

21. 21.

Thomson, E. C. et al. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**, 1171–1187.e20 (2021).

22. 22.

McCormick, K. D., Jacobs, J. L. & Mellors, J. W. The emerging plasticity of SARS-CoV-2. *Science* **371**, 1306–1308 (2021).

23. 23.

Starr, T. N. et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182**, 1295–1310.e20 (2020). **The critical value of early, comprehensive identification of constrained regions of proteins that are targets for vaccines and antibody-based therapeutics.**

24. 24.

Faria, N. R. et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815–821 (2021).

25. 25.

Lythgoe, K. A. et al. SARS-CoV-2 within-host diversity and transmission. *Science* **372**, eabg0821 (2021).

26. 26.

MacLean, O. A. et al. Natural selection in the evolution of SARS-CoV-2 in bats created a generalist virus and highly capable human pathogen. *PLoS Biol.* **19**, e3001115 (2021).

27. 27.

Richard, M. et al. SARS-CoV-2 is transmitted via contact and via the air between ferrets. *Nat. Commun.* **11**, 3496 (2020).

28. 28.

Hassan, A. O. et al. A single intranasal dose of chimpanzee adenovirus-vectored vaccine protects against SARS-CoV-2 infection in rhesus macaques. *Cell Rep. Med.* **2**, 100230 (2021).

29. 29.

Sit, T. H. C. et al. Infection of dogs with SARS-CoV-2. *Nature* **586**, 776–778 (2020).

30. 30.

Oreshkova, N. et al. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. *Euro Surveill.* **25**, (2020).

31. 31.

Patterson, E. I. et al. Evidence of exposure to SARS-CoV-2 in cats and dogs from households in Italy. *Nat. Commun.* **11**, 6231 (2020).

32. 32.

Neira, V. et al. A household case evidences shorter shedding of SARS-CoV-2 in naturally infected cats compared to their human owners. *Emerg. Microbes Infect.* **10**, 376–383 (2021).

33. 33.

McAloose, D. et al. From people to *Panthera*: natural SARS-CoV-2 infection in tigers and lions at the Bronx Zoo. *mBio* **11**, e02220 (2020).

34. 34.

Oude Munnink, B. B. et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* **371**, 172–177 (2021).  
**Human-to-animal-to-human adaptation and transmission of SARS-CoV-2 within mink farms.**

35. 35.

Bayarri-Olmos, R. et al. The SARS-CoV-2 Y453F mink variant displays a pronounced increase in ACE-2 affinity but does not challenge antibody neutralization. *J. Biol. Chem.* **296**, 100536 (2021).

36. 36.

Nelson, M. I., Gramer, M. R., Vincent, A. L. & Holmes, E. C. Global transmission of influenza viruses from humans to swine. *J. Gen. Virol.* **93**, 2195–2203 (2012).

37. 37.

Baum, A. et al. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014–1018 (2020).

38. 38.

Hoffmann, M. et al. SARS-CoV-2 mutations acquired in mink reduce antibody-mediated neutralization. *Cell Rep.* **35**, 109017 (2021).

39. 39.

Rathnasinghe, R. et al. The N501Y mutation in SARS-CoV-2 spike leads to morbidity in obese and aged mice and is neutralized by convalescent and post-vaccination human sera. Preprint at <https://doi.org/10.1101/2021.01.19.21249592> (2021).

40. 40.

Gu, H. et al. Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science* **369**, 1603–1607 (2020).

41. 41.

Montagutelli, X. et al. The B1.351 and P.1 variants extend SARS-CoV-2 host range to mice. Preprint at <https://doi.org/10.1101/2021.03.18.436013> (2021).

42. 42.

Dejnirattisai, W. et al. Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell* **184**, 2939–2954.e9 (2021).

43. 43.

Manicassamy, B. et al. Protection of mice against lethal challenge with 2009 H1N1 influenza A virus by 1918-like and classical swine H1N1 based vaccines. *PLoS Pathog.* **6**, e1000745 (2010).

44. 44.

Kelley, B., Renshaw, T. & Kamarck, M. Process and operations strategies to enable global access to antibody therapies. *Biotechnol. Prog.* **37**, e3139 (2021).

45. 45.

Bartsch, Y. C. et al. Humoral signatures of protective and pathological SARS-CoV-2 infection in children. *Nat. Med.* **27**, 454–462 (2021).

46. 46.

Winkler, E. S. et al. Human neutralizing antibodies against SARS-CoV-2 require intact Fc effector functions for optimal therapeutic protection. *Cell* **184**, 1804–1820 (2021).

47. 47.

Schäfer, A. et al. Antibody potency, effector function, and combinations in protection and therapy for SARS-CoV-2 infection in vivo. *J. Exp. Med.* **218**, e20201993 (2021).

48. 48.

Rosendahl Huber, S., van Beek, J., de Jonge, J., Luytjes, W. & van Baarle, D. T cell responses to viral infections — opportunities for peptide vaccination. *Front. Immunol.* **5**, 171 (2014).

49. 49.

Painter, M. M. et al. Rapid induction of antigen-specific CD4<sup>+</sup> T cells guides coordinated humoral and cellular immune responses to SARS-CoV-2 mRNA vaccination. Preprint at <https://doi.org/10.1101/2021.04.21.440862> (2021).

50. 50.

Chen, R. E. et al. Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* **27**, 717–726 (2021).

51. 51.

Lempp, F. A. et al. Membrane lectins enhance SARS-CoV-2 infection and influence the neutralizing activity of different classes of antibodies. Preprint at <https://doi.org/10.1101/2021.04.03.438258> (2021).

52. 52.

Piccoli, L. et al. Mapping neutralizing and immunodominant sites on the SARS-CoV-2 spike receptor-binding domain by structure-guided high-resolution serology. *Cell* **183**, 1024–1042.e21 (2020). **The critical value of early, comprehensive identification of monoclonal antibodies to define high-resolution antigenic maps of a pandemic pathogen.**

53. 53.

Khoury, D. S. et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* <https://doi.org/10.1038/s41591-021-01377-8> (2021).

54. 54.

Dan, J. M. et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).

55. 55.

Steenhuis, M. et al. Dynamics of antibodies to SARS-CoV-2 in convalescent plasma donors. *Clin. Transl. Immunol.* **10**, e1285 (2021).

56. 56.

Kuter, B. et al. Ten year follow-up of healthy children who received one or two injections of varicella vaccine. *Pediatr. Infect. Dis. J.* **23**, 132–137 (2004).

57. 57.

Griffin, D. E. Measles vaccine. *Viral Immunol.* **31**, 86–95 (2018).

58. 58.

Sanche, S. et al. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerg. Infect. Dis.* **26**, 1470–1477 (2020).

59. 59.

Plotkin, S. A. The history of rubella and rubella vaccination leading to elimination. *Clin. Infect. Dis.* **43** (Suppl 3), S164–S168 (2006).

60. 60.

Nader, S., Bergen, R., Sharp, M. & Arvin, A. M. Age-related differences in cell-mediated immunity to varicella-zoster virus among children and adults immunized with live attenuated varicella vaccine. *J. Infect. Dis.* **171**, 13–17 (1995).

61. 61.

Cardemil, C. V. et al. Effectiveness of a third dose of MMR vaccine for mumps outbreak control. *N. Engl. J. Med.* **377**, 947–956 (2017).

62. 62.

Lavine, J. S., Bjornstad, O. N. & Antia, R. Immunological characteristics govern the transition of COVID-19 to endemicity. *Science* **371**, 741–745 (2021).

63. 63.

Lee, V. J. et al. Oseltamivir ring prophylaxis for containment of 2009 H1N1 influenza outbreaks. *N. Engl. J. Med.* **362**, 2166–2174 (2010).

64. 64.

Houser, K. & Subbarao, K. Influenza vaccines: challenges and solutions. *Cell Host Microbe* **17**, 295–300 (2015).

65. 65.

Sandbulte, M. R. et al. Discordant antigenic drift of neuraminidase and hemagglutinin in H1N1 and H3N2 influenza viruses. *Proc. Natl Acad. Sci. USA* **108**, 20748–20753 (2011).

66. 66.

Madhi, S. A. et al. Efficacy of the ChAdOx1 nCoV-19 COVID-19 vaccine against the B.1.351 variant. *N. Engl. J. Med.* **384**, 1885–1898 (2021). **Mutations in the spike protein of SARS-CoV-2 in circulating variants such as B.1.351 (Beta) can reduce the efficacy of current vaccines, highlighting how progressive immune escape may contribute to endemicity.**

67. 67.

Kustin, T. et al. Evidence for increased breakthrough rates of SARS-CoV-2 variants of concern in BNT162b2-mRNA-vaccinated individuals. *Nat. Med.* <https://doi.org/10.1038/s41591-021-01413-7> (2021).

68. 68.

Cobey, S., Larremore, D. B., Grad, Y. H. & Lipsitch, M. Concerns about SARS-CoV-2 evolution should not hold back efforts to expand vaccination. *Nat. Rev. Immunol.* **21**, 330–335 (2021).

69. 69.

Lee, J. et al. Persistent antibody clonotypes dominate the serum response to influenza over multiple years and repeated vaccinations. *Cell Host Microbe* **25**, 367–376.e5 (2019).

70. 70.

Devarajan, P. & Swain, S. L. Original antigenic sin: friend or foe in developing a broadly cross-reactive vaccine to influenza? *Cell Host Microbe* **25**, 354–355 (2019).

71. 71.

Brown, E. L. & Essigmann, H. T. Original antigenic sin: the downside of immunological memory and implications for COVID-19. *mSphere* **6**, e00056-21 (2021).

72. 72.

Anderson, E. M. et al. Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *Cell* **184**, 1858–1864 (2021).

73. 73.

Ambrose, C. S. & Levin, M. J. The rationale for quadrivalent influenza vaccines. *Hum. Vaccin. Immunother.* **8**, 81–88 (2012).

74. 74.

Oh, M. D. et al. Middle East respiratory syndrome: what we learned from the 2015 outbreak in the Republic of Korea. *Korean J. Intern. Med.* **33**, 233–246 (2018).

75. 75.

Wahl, A. et al. SARS-CoV-2 infection is effectively treated and prevented by EIDD-2801. *Nature* **591**, 451–457 (2021).

76. 76.

Aschwanden, C. Five reasons why COVID herd immunity is probably impossible. *Nature* **591**, 520–522 (2021).

77. 77.

Paget, J. et al. Global mortality associated with seasonal influenza epidemics: new burden estimates and predictors from the GLaMOR Project. *J. Glob. Health* **9**, 020421 (2019).

78. 78.

Gilca, R., Carazo, S., Amini, R., Charest, H. & De Serres, G. Relative severity of common human coronaviruses and influenza in patients hospitalized with acute respiratory infection: results from 8-year hospital-based surveillance in Quebec, Canada. *J. Infect. Dis.* **223**, 1078–1087 (2021).

79. 79.

Gupta, R. Advancing new tools for infectious diseases. *Science* **370**, 913–914 (2020).

80. 80.

Gupta, R. The need for global access to biomedical innovations during pandemics. *Nat. Biotechnol.* **39**, 664–666 (2021).

81. 81.

Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **183**, 1735 (2020).

82. 82.

Wang, P. et al. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130–135 (2021).

83. 83.

Starr, T. N. et al. SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature* <https://doi.org/10.1038/s41586-021-03807-6> (2021).

84. 84.

Forni, D., Cagliani, R., Clerici, M. & Sironi, M. Molecular evolution of human coronavirus genomes. *Trends Microbiol.* **25**, 35–48 (2017).

## Acknowledgements

We thank D. Stoss, R. Gupta and B. Hubby for insights and S. Bianchi with help with the figures.

## Author information

### Affiliations

1. Vir Biotechnology, San Francisco, CA, USA

Amilio Telenti, Ann Arvin, Phillip S. Pang & Herbert W. Virgin

2. Department of Integrative Structural and Computational Biology, Scripps Research, La Jolla, CA, USA

Amilio Telenti

3. Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Lawrence Corey

4. Humabs Biomed SA, a subsidiary of Vir Biotechnology, Bellinzona, Switzerland

Davide Corti

5. Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO, USA

Michael S. Diamond

6. Department of Medicine, Washington University School of Medicine, St Louis, MO, USA

Michael S. Diamond

7. Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO, USA

Michael S. Diamond

8. Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre

9. Department of Medicine, Division of Infectious Diseases, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre

10. Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre

11. The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre

12. Department of Microbiology and Immunology, Tulane University, New Orleans, LA, USA

Robert F. Garry

13. Marie Bashir Institute for Infectious Diseases and Biosecurity, School of Life and Environmental Sciences and School of Medical Sciences, The University of Sydney, Sydney, New South Wales, Australia

Edward C. Holmes

14. Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO, USA

Herbert W. Virgin

15. Department of Internal Medicine, UT Southwestern Medical Center, Dallas, TX, USA

Herbert W. Virgin

16. Department of Pathology, Molecular and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre

## **Contributions**

Conception and design: A.T., P.P. and H.W.V.; drafting of sections: all authors.

## Corresponding authors

Correspondence to [Amalio Telenti](#) or [Ann Arvin](#) or [Lawrence Corey](#) or [Davide Corti](#) or [Michael S. Diamond](#) or [Adolfo García-Sastre](#) or [Robert F. Garry](#) or [Edward C. Holmes](#) or [Phillip S. Pang](#) or [Herbert W. Virgin](#).

## Ethics declarations

### Competing interests

A.T., A.A., D.C., P.P. and H.W.V. are employees of Vir Biotechnology Inc. and may hold shares in Vir Biotechnology Inc. L.C., M.S.D. and E.C.H. are co-founders or consultants of Vir Biotechnology Inc. The Diamond laboratory at Washington University School of Medicine has received sponsored research agreements from Moderna and Vir Biotechnology. The García-Sastre laboratory has received research support from Pfizer, Senhwa Biosciences and 7Hills Pharma. A.G.-S. has consulting agreements for the following companies involving cash and/or stock: Vivaldi Biosciences, Contraflect, 7Hills Pharma, Avimex, Vaxalto, Accurius and Esperovax. R.F.G. is co-founder of Zalgen Labs, a biotechnology company that develops countermeasures to emerging viruses.

## Additional information

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

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

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Telenti, A., Arvin, A., Corey, L. *et al.* After the pandemic: perspectives on the future trajectory of COVID-19. *Nature* **596**, 495–504 (2021). <https://doi.org/10.1038/s41586-021-03792-w>

- Received: 12 May 2021
  - Accepted: 01 July 2021
  - Published: 08 July 2021
  - Issue Date: 26 August 2021
  - DOI: <https://doi.org/10.1038/s41586-021-03792-w>
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03792-w>

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

- Article
- [Published: 25 August 2021](#)

# Chromatic periodic activity down to 120 megahertz in a fast radio burst

- [Inés Pastor-Marazuela](#) ORCID: [orcid.org/0000-0002-4357-8027<sup>1,2</sup>](https://orcid.org/0000-0002-4357-8027),
- [Liam Connor](#) ORCID: [orcid.org/0000-0002-7587-6352<sup>1,2,3</sup>](https://orcid.org/0000-0002-7587-6352),
- [Joeri van Leeuwen](#) ORCID: [orcid.org/0000-0001-8503-6958<sup>1,2</sup>](https://orcid.org/0000-0001-8503-6958),
- [Yogesh Maan](#) ORCID: [orcid.org/0000-0002-0862-6062<sup>2</sup>](https://orcid.org/0000-0002-0862-6062),
- [Sander ter Veen<sup>2</sup>](#),
- [Anna Bilous<sup>2</sup>](#),
- [Leon Oostrum](#) ORCID: [orcid.org/0000-0001-8724-8372<sup>1,2,4</sup>](https://orcid.org/0000-0001-8724-8372),
- [Emily Petroff](#) ORCID: [orcid.org/0000-0002-9822-8008<sup>1</sup>](https://orcid.org/0000-0002-9822-8008),
- [Samayra Straal<sup>5,6</sup>](#),
- [Dany Vohl](#) ORCID: [orcid.org/0000-0003-1779-4532<sup>2</sup>](https://orcid.org/0000-0003-1779-4532),
- [Jisk Attema<sup>4</sup>](#),
- [Oliver M. Boersma<sup>1,2</sup>](#),
- [Eric Kooistra<sup>2</sup>](#),
- [Daniel van der Schuur<sup>2</sup>](#),
- [Alessio Sclocco<sup>4</sup>](#),
- [Roy Smits<sup>2</sup>](#),
- [Elizabeth A. K. Adams](#) ORCID: [orcid.org/0000-0002-9798-5111<sup>2,7</sup>](https://orcid.org/0000-0002-9798-5111),
- [Björn Adebarh<sup>8</sup>](#),
- [W. J. G. de Blok<sup>2,7,9</sup>](#),
- [Arthur H. W. M. Coolen<sup>2</sup>](#),
- [Sieds Damstra<sup>2</sup>](#),
- [Helga Dénes](#) ORCID: [orcid.org/0000-0002-9214-8613<sup>2</sup>](https://orcid.org/0000-0002-9214-8613),
- [Kelley M. Hess<sup>2,7</sup>](#),

- [Thijs van der Hulst<sup>7</sup>](#),
- [Boudewijn Hut<sup>2</sup>](#),
- [V. Marianna Ivashina<sup>10</sup>](#),
- [Alexander Kutkin<sup>2,11</sup>](#),
- [G. Marcel Loose](#) [ORCID: orcid.org/0000-0003-4721-747X<sup>2</sup>](#),
- [Danielle M. Lucero<sup>12</sup>](#),
- [Ágnes Mika<sup>2</sup>](#),
- [Vanessa A. Moss<sup>2,13,14</sup>](#),
- [Henk Mulder<sup>2</sup>](#),
- [Menno J. Norden<sup>2</sup>](#),
- [Tom Oosterloo<sup>2,7</sup>](#),
- [Emanuela Orrú<sup>2</sup>](#),
- [Mark Ruiter<sup>2</sup>](#) &
- [Stefan J. Wijnholds](#) [ORCID: orcid.org/0000-0002-0463-3011<sup>2</sup>](#)

[Nature](#) volume **596**, pages 505–508 (2021)

- 699 Accesses
- 116 Altmetric
- [Metrics details](#)

## Subjects

- [Compact astrophysical objects](#)
- [High-energy astrophysics](#)
- [Time-domain astronomy](#)
- [Transient astrophysical phenomena](#)

## Abstract

Fast radio bursts (FRBs) are extragalactic astrophysical transients<sup>1</sup> whose brightness requires emitters that are highly energetic yet compact enough to produce the short, millisecond-duration bursts. FRBs have thus far been detected at frequencies from 8 gigahertz (ref. <sup>2</sup>) down to 300 megahertz

(ref. 3), but lower-frequency emission has remained elusive. Some FRBs repeat<sup>4,5,6</sup>, and one of the most frequently detected, FRB 20180916B<sup>7</sup>, has a periodicity cycle of 16.35 days (ref. 8). Using simultaneous radio data spanning a wide range of wavelengths (a factor of more than 10), here we show that FRB 20180916B emits down to 120 megahertz, and that its activity window is frequency dependent (that is, chromatic). The window is both narrower and earlier at higher frequencies. Binary wind interaction models predict a wider window at higher frequencies, the opposite of our observations. Our full-cycle coverage shows that the 16.3-day periodicity is not aliased. We establish that low-frequency FRB emission can escape the local medium. For bursts of the same fluence, FRB 20180916B is more active below 200 megahertz than at 1.4 gigahertz. Combining our results with previous upper limits on the all-sky FRB rate at 150 megahertz, we find there are 3–450 FRBs in the sky per day above 50 Jy ms. Our chromatic results strongly disfavour scenarios in which absorption from strong stellar winds causes FRB periodicity. We demonstrate that some FRBs are found in ‘clean’ environments that do not absorb or scatter low-frequency radiation.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

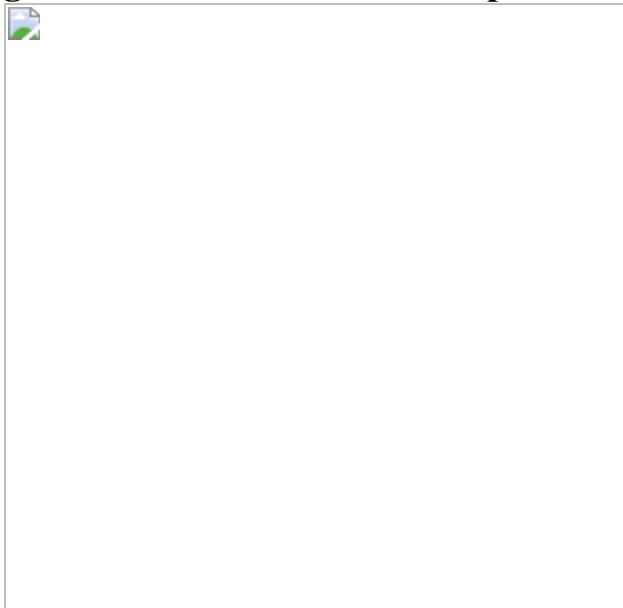
[Rent or Buy](#)

All prices are NET prices.

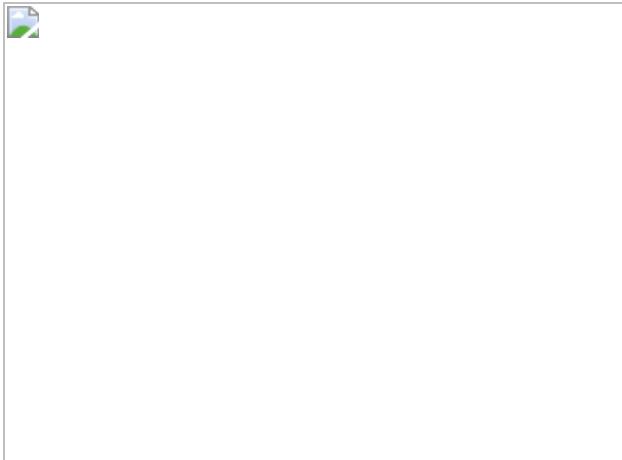
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

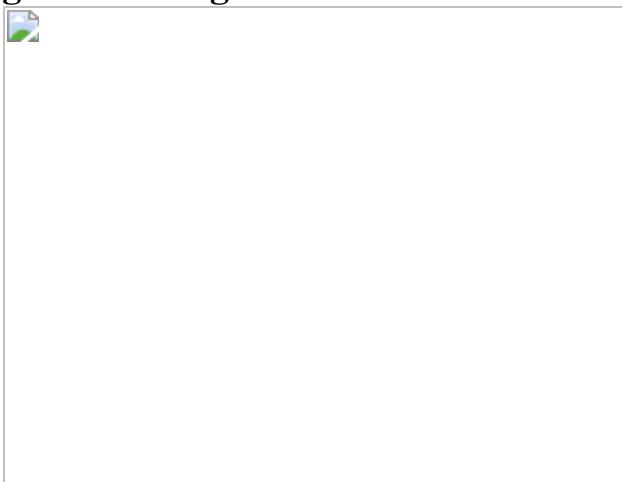
**Fig. 1: Two bursts at different phases.**



**Fig. 2: The nine LOFAR bursts.**



**Fig. 3: Coverage and detection times.**



**Fig. 4: Frequency dependence of the FRB 20180916B activity windows.**

## Data availability

Raw data were generated by the Apertif system on the Westerbork Synthesis Radio Telescope and by the International LOFAR Telescope. The Apertif data that support the findings of this study are available through the

ALERT archive, <http://www.alert.eu/FRB20180916B>. The LOFAR data are available through the LOFAR Long Term Archive, <https://lta.lofar.eu/>, by searching for ‘Observations’ at J2000 coordinates RA = 01:57:43.2000, Dec. = +65:42:01.020, or by selecting COM\_ALERT in ‘Other projects’ and downloading data which includes R3 in the ‘Observation description’.

## Code availability

The custom code used to generate these results is publicly available at <https://doi.org/10.5281/zenodo.4559593> (ref. [80](#)).

## References

1. 1.  
Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J. & Crawford, F. A bright millisecond radio burst of extragalactic origin. *Science* **318**, 777–780 (2007).
2. 2.  
Gajjar, V. et al. Highest-frequency detection of FRB 121102 at 4–8 GHz using the breakthrough listen digital backend at the Green Bank Telescope. *Astrophys. J.* **863**, 2 (2018).
3. 3.  
Chawla, P. et al. Detection of repeating FRB 180916.J0158+65 down to frequencies of 300 MHz. *Astrophys. J.* **896**, L41 (2020).
4. 4.  
Spitler, L. G. et al. A repeating fast radio burst. *Nature* **531**, 202–205 (2016).
5. 5.

The CHIME/FRB Collaboration. A second source of repeating fast radio bursts. *Nature* **566**, 235–238 (2019).

6. 6.

Fonseca, E. et al. Nine new repeating fast radio burst sources from CHIME/FRB. *Astrophys. J.* **891**, L6 (2020).

7. 7.

The CHIME/FRB Collaboration et al. CHIME/FRB detection of eight new repeating fast radio burst sources. *Astrophys. J.* **885**, L24 (2019).

8. 8.

The CHIME/FRB Collaboration. Periodic activity from a fast radio burst source. *Nature* **582**, 351–355 (2020).

9. 9.

Maan, Y. & van Leeuwen, J. Real-time searches for fast transients with Apertif and LOFAR. In *IEEE Proc. URSI GASS* <https://doi.org/10.23919/URSIGASS.2017.8105320> (IEEE, 2017).

10. 10.

Stappers, B. W. et al. Observing pulsars and fast transients with LOFAR. *Astron. Astrophys.* **530**, A80 (2011).

11. 11.

Pleunis, Z. et al. LOFAR detection of 110–188 MHz emission and frequency-dependent activity from FRB 20180916B. *Astrophys. J. Lett.* **911**, L3 (2021).

12. 12.

Coenen, T. et al. The LOFAR pilot surveys for pulsars and fast radio transients. *Astron. Astrophys.* **570**, A60 (2014).

13. 13.

Karastergiou, A. et al. Limits on fast radio bursts at 145 MHz with ARTEMIS, a real-time software backend. *Mon. Not. R. Astron. Soc.* **452**, 1254–1262 (2015).

14. 14.

Marcote, B. et al. A repeating fast radio burst source localized to a nearby spiral galaxy. *Nature* **577**, 190–194 (2020).

15. 15.

Hessels, J. W. T. et al. FRB 121102 bursts show complex time-frequency structure. *Astrophys. J.* **876**, L23 (2019).

16. 16.

Cordes, J. M. & Lazio, T. J. W. NE2001.I. A new model for the Galactic distribution of free electrons and its fluctuations. Preprint at <http://arXiv.org/abs/astro-ph/0207156> (2003).

17. 17.

McQuinn, M. Locating the “missing” baryons with extragalactic dispersion measure estimates. *Astrophys. J.* **780**, L33 (2014).

18. 18.

Chamma, M. A., Rajabi, F., Wyenberg, C. M., Mathews, A. & Houde, M. Evidence of a shared law between sources of repeating fast radio bursts. Preprint at <http://arXiv.org/abs/2010.14041> (2020).

19. 19.

Josephy, A. et al. CHIME/FRB detection of the original repeating fast radio burst source FRB 121102. *Astrophys. J.* **882**, L18 (2019).

20. 20.

Aggarwal, K. et al. VLA/realfast detection of burst from FRB180916.J0158+65 and tests for periodic activity. *Res. Notes AAS* **4**, 94 (2020).

21. 21.

Rajwade, K. M. et al. Possible periodic activity in the repeating FRB 121102. *Mon. Not. R. Astron. Soc.* **495**, 3551–3558 (2020).

22. 22.

Lyutikov, M., Barkov, M. & Giannios, D. FRB-periodicity: mild pulsar in tight O/B-star binary. *Astrophys. J.* **893**, L39 (2020).

23. 23.

Ioka, K. & Zhang, B. A binary comb model for periodic fast radio bursts. *Astrophys. J.* **893**, L26 (2020).

24. 24.

Levin, Y., Beloborodov, A. M. & Bransgrove, A. Precessing flaring magnetar as a source of repeating FRB 180916.J0158+65. *Astrophys. J.* **895**, L30 (2020).

25. 25.

Zanazzi, J. J. & Lai, D. Periodic fast radio bursts with neutron star free/radiative precession. *Astrophys. J.* **892**, L15 (2020).

26. 26.

Tong, H., Wang, W. & Wang, H. G. Periodicity in fast radio bursts due to forced precession by a fallback disk. *Res. Astron. Astrophys.* **20**, 142 (2020).

27. 27.

Metzger, B. D., Margalit, B. & Sironi, L. Fast radio bursts as synchrotron maser emission from decelerating relativistic blast waves. *Mon. Not. R. Astron. Soc.* **485**, 4091–4106 (2019).

28. 28.

Beniamini, P., Wadiasingh, Z. & Metzger, B. D. Periodicity in recurrent fast radio bursts and the origin of ultra long period magnetars. *Mon. Not. R. Astron. Soc.* **496**, 3390–3401 (2020).

29. 29.

Sand, K. R. et al. Low-frequency detection of FRB180916 with the uGMRT. *Astron. Telegr.* **13781** (2020).

30. 30.

Marthi, V. R. et al. Detection of 15 bursts from FRB 180916.J0158+65 with the uGMRT. *Mon. Not. R. Astron. Soc. Lett.* **499**, L16–L20 (2020).

31. 31.

Oosterloo, T., Verheijen, M. & van Cappellen, W. The latest on Apertif. In *Proc. ISKAF2010 Science Meeting* 043 (Proceedings of Science, Vol. 112, 2010).

32. 32.

Adams, E. A. K. & van Leeuwen, J. Radio surveys now both deep and wide. *Nature Astron.* **3**, 188 (2019).

33. 33.

Connor, L. et al. A bright, high rotation-measure FRB that skewers the M33 halo. *Mon. Not. R. Astron. Soc.* **499**, 4716–4724 (2020).

34. 34.

Oostrum, L. C. et al. Repeating fast radio bursts with WSRT/Apertif. *Astron. Astrophys.* **635**, A61 (2020).

35. 35.

van Leeuwen, J. ARTS – the Apertif Radio Transient System. In *Proc. The Third Hot-wiring the Transient Universe Workshop* <https://www.slac.stanford.edu/econf/C131113.1/proceedings.html> (eds Wozniak, P. R., Graham, M. J., Mahabal, A. A. & Seaman, R.) 79 (LANL, 2014).

36. 36.

Oostrum, L. C. *Fast Radio Bursts with Apertif*. PhD thesis, Univ. Amsterdam (2020).

37. 37.

Sclocco, A., van Leeuwen, J., Bal, H. E. & van Nieuwpoort, R. V. Real-time dedispersion for fast radio transient surveys, using auto tuning on many-core accelerators. *Astronomy and Computing* **14**, 1–7 (2016).

38. 38.

Sclocco, A., Heldens, S. & van Werkhoven, B. AMBER: a real-time pipeline for the detection of single pulse astronomical transients. *SoftwareX* **12**, 100549 (2020).

39. 39.

Sclocco, A., Vohl, D. & van Nieuwpoort, R. V. Real-time rfi mitigation for the apertif radio transient system. In *Proc. 2019 RFI Workshop — Coexisting with Radio Frequency Interference (RFI)* 1–8 (IEEE, 2019).

40. 40.

Oostrum, L. C. loostrum/darc: version 2.1. <https://doi.org/10.5281/zenodo.3784870> (2020).

41. 41.

Connor, L. & van Leeuwen, J. Applying deep learning to fast radio burst classification. *Astron. J.* **156**, 256 (2018).

42. 42.

Yao, J. M., Manchester, R. N. & Wang, N. A new electron density model for estimation of pulsar and FRB distances. *Astrophys. J.* **835**, 29 (2017).

43. 43.

Ransom, S. M. *New Search Techniques For Binary Pulsars*. Ph.D. thesis, Harvard Univ. (2001).

44. 44.

van Haarlem, M. P. et al. LOFAR: The LOw-Frequency ARray. *Astron. Astrophys.* **556**, A2 (2013).

45. 45.

Maan, Y., van Leeuwen, J. & Vohl, D. Fourier domain excision of periodic radio frequency interference. *Astron. Astrophys.* **650**, A80 (2021).

46. 46.

Perley, R. A. & Butler, B. J. An accurate flux density scale from 50 MHz to 50 GHz. *Astrophys. J. Suppl. Ser.* **230**, 7 (2017).

47. 47.

Cordes, J. M. & McLaughlin, M. A. Searches for fast radio transients. *Astrophys. J.* **596**, 1142–1154 (2003).

48. 48.

Maan, Y. & Aswathappa, H. A. Deep searches for decametre-wavelength pulsed emission from radio-quiet gamma-ray pulsars. *Mon. Not. R. Astron. Soc.* **445**, 3221–3228 (2014).

49. 49.

Houben, L. J. M. et al. Constraints on the low frequency spectrum of FRB 121102. *Astron. Astrophys.* **623**, A42 (2019).

50. 50.

Sanidas, S. et al. The LOFAR Tied-Array All-Sky Survey (LOTAAS): survey overview and initial pulsar discoveries. *Astron. Astrophys.* **626**, A104 (2019).

51. 51.

Kondratiev, V. I. et al. A LOFAR census of millisecond pulsars. *Astron. Astrophys.* **585**, A128 (2016).

52. 52.

Bilous, A. V. et al. A LOFAR census of non-recycled pulsars: average profiles, dispersion measures, flux densities, and spectra. *Astron. Astrophys.* **591**, A134 (2016).

53. 53.

Pilia, M. et al. The lowest frequency fast radio bursts: Sardinia Radio Telescope detection of the periodic FRB 180916 at 328 MHz. *Astrophys. J.* **896**, L40 (2020).

54. 54.

Huijse, P. et al. Robust period estimation using mutual information for multiband light curves in the synoptic survey era. *Astrophys. J. Suppl. Ser.* **236**, 12 (2018).

55. 55.

Scholz, P. et al. Simultaneous X-ray and radio observations of the repeating fast radio burst FRB 180916.J0158+65. *Astrophys. J.* **901**, 165 (2020).

56. 56.

Pearlman, A. B. et al. Multiwavelength radio observations of two repeating fast radio burst sources: FRB 121102 and FRB 180916.J0158+65. *Astrophys. J.* **905**, L27 (2020).

57. 57.

Scott, D. W. *Multivariate Density Estimation: Theory, Practice, and Visualization* (Wiley & Sons, 2015).

58. 58.

Lyutikov, M. Radius-to-frequency mapping and FRB frequency drifts. *Astrophys. J.* **889**, 135 (2020).

59. 59.

Purcell, C. R., Van Eck, C. L., West, J., Sun, X. H. & Gaensler, B. M. RM-Tools: rotation measure (RM) synthesis and Stokes QU-fitting. (Astrophysics Source Code Library, ascl:2005.003, 2020).

60. 60.

Ordog, A., Booth, R. A., Van Eck, C. L., Brown, J.-A. C. & Landecker, T. L. Faraday rotation of extended emission as a probe of the large-scale galactic magnetic field. *Galaxies* **7**, 43 (2019).

61. 61.

Nimmo, K. et al. Highly polarized microstructure from the repeating FRB 20180916B. *Nature Astron.* **5**, 594–603 (2021).

62. 62.

Blaskiewicz, M., Cordes, J. M. & Wasserman, I. A relativistic model of pulsar polarization. *Astrophys. J.* **370**, 643 (1991).

63. 63.

Cho, H. et al. Spectropolarimetric analysis of FRB 181112 at microsecond resolution: implications for fast radio burst emission mechanism. *Astrophys. J. Lett.* **891**, L38 (2021).

64. 64.

Beloborodov, A. M. A flaring magnetar in FRB 121102? *Astrophys. J.* **843**, L26 (2017).

65. 65.

Hotan, A. W., van Straten, W. & Manchester, R. N. Psrchive and Psrfits: an open approach to radio pulsar data storage and analysis. *Publ. Astron. Soc. Aust.* **21**, 302–309 (2004).

66. 66.

Caleb, M. et al. Simultaneous multi-telescope observations of FRB 121102. *Mon. Not. R. Astron. Soc.* **496**, 4565–4573 (2020).

67. 67.

Wang, W., Zhang, B., Chen, X. & Xu, R. On the time-frequency downward drifting of repeating fast radio bursts. *Astrophys. J.* **876**, L15 (2019).

68. 68.

Rajabi, F., Chamma, M. A., Wyenberg, C. M., Mathews, A. & Houde, M. A simple relationship for the spectro-temporal structure of bursts from FRB 121102. *Mon. Not. R. Astron. Soc.* **498**, 4936–4942 (2020).

69. 69.

Krishnakumar, M. A., Joshi, B. C. & Manoharan, P. K. Multi-frequency scatter broadening evolution of pulsars. I. *Astrophys. J.* **846**, 104 (2017).

70. 70.

Maan, Y., Joshi, B. C., Surnis, M. P., Bagchi, M. & Manoharan, P. K. Distinct properties of the radio burst emission from the magnetar XTE J1810–197. *Astrophys. J.* **882**, L9 (2019).

71. 71.

Rane, A. et al. A search for rotating radio transients and fast radio bursts in the Parkes high-latitude pulsar survey. *Mon. Not. R. Astron. Soc.* **455**, 2207–2215 (2016).

72. 72.

Lawrence, E., Wiel, S. V., Law, C. J., Spolaor, S. B. & Bower, G. C. The non-homogeneous Poisson process for fast radio burst rates. *Astron. J.* **154**, 117 (2017).

73. 73.

Vedantham, H. K., Ravi, V., Hallinan, G. & Shannon, R. The fluence and distance distributions of fast radio bursts. *Astrophys. J.* **830**, 75 (2016).

74. 74.

Keane, E. F. & Petroff, E. Fast radio bursts: search sensitivities and completeness. *Mon. Not. R. Astron. Soc.* **447**, 2852–2856 (2015).

75. 75.

Connor, L. Interpreting the distributions of FRB observables. *Mon. Not. R. Astron. Soc.* **487**, 5753–5763 (2019).

76. 76.

ter Veen, S. et al. The FRATS project: real-time searches for fast radio bursts and other fast transients with LOFAR at 135 MHz. *Astron. Astrophys.* **621**, A57 (2019).

77. 77.

Tingay, S. J. et al. A search for fast radio bursts at low frequencies with Murchison Widefield Array high time resolution imaging. *Astron. J.* **150**, 199 (2015).

78. 78.

Rowlinson, A. et al. Limits on fast radio bursts and other transient sources at 182 MHz using the Murchison Widefield Array. *Mon. Not. R. Astron. Soc.* **458**, 3506–3522 (2016).

79. 79.

Sokolowski, M. et al. No low-frequency emission from extremely bright fast radio bursts. *Astrophys. J.* **867**, L12 (2018).

80. 80.

Pastor-Marazuela, I. Reproduction package for “Chromatic periodic activity down to 120 megahertz in a fast radio burst”.  
<https://doi.org/10.5281/zenodo.4559593> (2021).

## Acknowledgements

This research was supported by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC grant agreement no. 617199 (‘ALERT’), and by Vici research programme ‘ARGO’ with project number 639.043.815, financed by the Dutch Research Council (NWO). Instrumentation development was supported by NWO (grant 614.061.613 ‘ARTS’) and the Netherlands Research School for Astronomy (‘NOVA4-ARTS’, ‘NOVA-NW3’ and ‘NOVA5-NW3-10.3.5.14’). PI of aforementioned grants is J.v.L. We further acknowledge funding from an NWO Veni Fellowship to E.P.; from Netherlands eScience

Center (NLeSC) grant ASDI.15.406 to D.V. and A.S.; from National Aeronautics and Space Administration (NASA) grant number NNX17AL74G issued through the NNH16ZDA001N Astrophysics Data Analysis Program (ADAP) to S.S.; by the WISE research programme, financed by NWO, to E.A.K.A.; from FP/2007-2013 ERC grant agreement no. 291531 ('HISTORYNU') to T.v.d.H.; and from VINNOVA Vinnmer grant 2009-01175 to V.M.I. I.P.-M. and Y.M. thank M. A. Krishnakumar for providing a software module that was useful in estimating the scatter-broadening timescale. This work makes use of data from the Apertif system installed at the Westerbork Synthesis Radio Telescope owned by ASTRON. ASTRON, the Netherlands Institute for Radio Astronomy, is an institute of NWO. This paper is based (in part) on data obtained with the International LOFAR Telescope (ILT) under project code COM\_ALERT. These data are accessible through the LOFAR Long Term Archive, <https://lta.lofar.eu/>. LOFAR (Methods) is the low frequency array designed and constructed by ASTRON. It has observing, data processing and data storage facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the ILT foundation under a joint scientific policy. The ILT resources have benefitted from the following recent major funding sources: CNRS-INSU, Observatoire de Paris and Université d'Orléans, France; BMBF, MIWF-NRW, MPG, Germany; Science Foundation Ireland (SFI), Department of Business, Enterprise and Innovation (DBEI), Ireland; NWO, The Netherlands; The Science and Technology Facilities Council, UK; Ministry of Science and Higher Education, Poland. We acknowledge use of the CHIME/FRB Public Database, provided at <https://www.chime-frb.ca/> by the CHIME/FRB Collaboration.

## Author information

### Affiliations

1. Anton Pannekoek Institute, University of Amsterdam, Amsterdam, The Netherlands

Inés Pastor-Marazuela, Liam Connor, Joeri van Leeuwen, Leon Oostrum, Emily Petroff & Oliver M. Boersma

2. ASTRON, the Netherlands Institute for Radio Astronomy, Dwingeloo,  
The Netherlands

Inés Pastor-Marazuela, Liam Connor, Joeri van Leeuwen, Yogesh  
Maan, Sander ter Veen, Anna Bilous, Leon Oostrum, Dany  
Vohl, Oliver M. Boersma, Eric Kooistra, Daniel van der Schuur, Roy  
Smits, Elizabeth A. K. Adams, W. J. G. de Blok, Arthur H. W. M.  
Coolen, Sieds Damstra, Helga Dénes, Kelley M. Hess, Boudewijn  
Hut, Alexander Kutkin, G. Marcel Loose, Ágnes Mika, Vanessa A.  
Moss, Henk Mulder, Menno J. Norden, Tom Oosterloo, Emanuela  
Orrú, Mark Ruiter & Stefan J. Wijnholds

3. Cahill Center for Astronomy, California Institute of Technology,  
Pasadena, CA, USA

Liam Connor

4. Netherlands eScience Center, Amsterdam, The Netherlands

Leon Oostrum, Jisk Attema & Alessio Sclocco

5. NYU Abu Dhabi, Abu Dhabi, United Arab Emirates

Samayra Straal

6. Center for Astro, Particle, and Planetary Physics (CAP<sup>3</sup>), NYU Abu  
Dhabi, Abu Dhabi, United Arab Emirates

Samayra Straal

7. Kapteyn Astronomical Institute, Groningen, The Netherlands

Elizabeth A. K. Adams, W. J. G. de Blok, Kelley M. Hess, Thijs van  
der Hulst & Tom Oosterloo

8. Astronomisches Institut der Ruhr-Universität Bochum (AIRUB),  
Bochum, Germany

Björn Adebahr

9. Department of Astronomy, University of Cape Town, Rondebosch,  
South Africa

W. J. G. de Blok

10. Department of Electrical Engineering, Chalmers University of  
Technology, Gothenburg, Sweden

V. Marianna Ivashina

11. Astro Space Center of Lebedev Physical Institute, Moscow, Russia

Alexander Kutkin

12. Department of Physics, Virginia Polytechnic Institute and State  
University, Blacksburg, VA, USA

Danielle M. Lucero

13. CSIRO Astronomy and Space Science, Australia Telescope National  
Facility, Epping, New South Wales, Australia

Vanessa A. Moss

14. Sydney Institute for Astronomy, School of Physics, University of  
Sydney, Sydney, New South Wales, Australia

Vanessa A. Moss

## Contributions

I.P.-M., L.C., J.v.L., Y.M., S.t.V., A.B., L.O., E.P., S.S. and D.V. analysed and interpreted the data. I.P.-M., L.C., J.v.L., Y.M. and S.t.V. contributed to the LOFAR data acquisition, and to the conception, design and creation of LOFAR analysis software. I.P.-M., L.C. and J.v.L. conceived and drafted the work, and Y.M., S.t.V., A.B., L.O., E.P., S.S. and D.V. contributed substantial revisions. L.O., J.A., O.M.B., E.K., D.v.d.S., A.S., R.S., E.A.K.A., B.A., W.J.G.d.B., A.H.W.M.C., S.D., H.D., K.M.H., T.v.d.H.,

B.H., V.M.I., A.K., G.M.L., D.M.L., A.M., V.A.M., H.M., M.J.N., T.O., E.O., M.R. and S.J.W. contributed to the conception, design and creation of the Apertif hardware, software and firmware used in this work, and to the Apertif data acquisition.

## Corresponding author

Correspondence to [Joeri van Leeuwen](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Scott Ransom and Matthew Bailes for their contribution to the peer review of this work.

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

## Extended data figures and tables

### [Extended Data Fig. 1 The FRB 20180916B fluence distribution at Apertif and at LOFAR.](#)

For each fluence  $F$  we plot how many brighter bursts are detected per hour, Rate ( $>F$ ) ( $\text{h}^{-1}$ ). The light green data points show the cumulative distribution function (CDF) of all Apertif bursts, with dash-dotted, dotted and dashed lines giving the power-law fit respectively to bursts with fluences lower than 3.2 Jy ms, between 3.2 Jy ms and 7.8 Jy ms, and above 7.8 Jy ms. The coloured solid lines correspond to different phase ranges within the active window, with no discernible difference between them

other than the rate scaling. The LOFAR fluence distribution is shown in crimson. The fit to a broken power law ('broken pl') with a fluence turnover at 104 Jy ms is shown as a grey dotted line. For the same fluence, FRB 20180916B is more active at 150 MHz than 1,370 MHz, even at the peak activity phases observed by Apertif.

### Extended Data Fig. 2 Dynamic spectra of Apertif bursts A01–A27.

We display PA (top panel), Stokes parameters I, L and V (central panel) and dynamic spectra (bottom panel), for bursts with full Stokes data (for example, panel A01 at top left). Bursts with only intensity data, such as A02, are limited to the total intensity profile. Burst identifiers are given in the top left corners, and activity cycle number in the top-right corners. Data have been dedispersed to  $DM = 348.75 \text{ pc cm}^{-3}$ , and downsampled  $2\times$  in time and  $8\times$  in frequency.

### Extended Data Fig. 3 Dynamic spectra of Apertif bursts A28–A54.

As in Extended Data Fig. 2.

### Extended Data Fig. 4 Observations and detections as a function of phase.

**a, b**, Shown are histograms of burst detections ('N. Bursts'; **a**) and of observation duration ('Obs. Duration'; **b**), both as a function of phase for the best period fitted to Apertif and CHIME/FRB data (16.29 days). In both panels, instruments are colour-coded by central frequency, with blue for high frequencies and red for low frequencies. This figure was generated using an adaptation of the frbpa package<sup>20</sup>.

### Extended Data Fig. 5 Comparison of simulated and observed activity window P values.

**a–c**, Each panel compares the  $P$  value obtained through the Kolmogorov–Smirnov statistic on two instrument burst samples. The vertical black lines give the observed  $P$  value, whereas the histograms correspond to  $10^5$  simulations of the  $P$  value that would be obtained if both instrument burst samples were drawn from the same distribution.  $N$  is the number of resulting simulations per  $P$  value. Shown are comparisons of burst samples from Apertif and LOFAR (**a**), Apertif and CHIME/FRB (**b**), and CHIME/FRB and LOFAR (**c**). In all panels, the vertical grey dotted, dash-dotted and dashed lines show respectively the  $P$  value where 68.27% ( $1\sigma$ ), 95.45% ( $2\sigma$ ) and 99.73% ( $3\sigma$ ) of the simulations give a larger  $P$  value.

### Extended Data Fig. 6 Stacked LOFAR bursts.

After dedispersion to the S/N-maximizing DM of  $349.00 \text{ pc cm}^{-3}$ , the individual bursts were co-added. **a**, The pulse profiles in eight different frequency bands of the co-added total, and fits to the scattering tail. The central frequency of the band is indicated on the vertical-axis labels. **b**, The dynamic spectrum of the stacked bursts.

### Extended Data Fig. 7 Apertif burst properties against phase.

**a**, The structure-optimized DM, with the  $348.75 \text{ pc cm}^{-3}$  average as a reference. **b**, The drift rate of bursts with multiple components. **c**, **d**, The fluence (**c**) and the average polarization position angle (PA) (**d**) of each burst. In all panels, bursts are colour-coded by activity cycle. Each colour corresponds to a different activity cycle (see key at bottom left), and the data points with a black edge represent bursts with  $\text{S}/\text{N} > 20$ . The error bars represent  $1\sigma$  (s.d.) errors.

### Extended Data Fig. 8 Five of the bursts with a measurable drift rate.

**a–e**, For each burst, the top panel shows the pulse profile as a solid black line and the fitted multi-component Gaussian in grey (the burst name is given at top left.). Coloured regions indicate the subcomponent position. The main panels show the dynamic spectra, the subcomponent centroids

with  $1\sigma$  (s.d.) errors and the fitted drift rate  $\dot{\nu}$  (white line). The right panels display the spectra and the fitted Gaussian of each subcomponent, with the same colour as the shaded region of the pulse profile.

## Extended Data Fig. 9 Finding the best period.

**a–l**, The periodograms between 0.03 day and 20 day periods of four instrument combinations and three different period searching techniques. Each column corresponds, from left to right, to all detections combined (blue), Apertif detections (green), CHIME/FRB detections (yellow) and CHIME/FRB and Apertif detections combined (red). Each row corresponds to a different search technique, with Pearson’s  $\chi^2$  test<sup>7</sup> at the top, maximum continuous fraction in the centre<sup>21</sup>, and the normalized QMIEU method<sup>54</sup> at the bottom. The vertical grey lines mark the position of the aliased periods, solid lines for  $f_N = (Nf_{\text{sid}} + f_0)$  and dotted lines for  $f_N = (Nf_{\text{sid}} - f_0)$ . The number in the top left corner of each plot indicates the best period using the given burst data set and periodicity search method, with errors giving the full-width at half-maximum.

**Extended Data Table 1 Summary of LOFAR burst properties**  
**Extended Data Table 2 Summary of Apertif burst properties**

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Pastor-Marazuela, I., Connor, L., van Leeuwen, J. *et al.* Chromatic periodic activity down to 120 megahertz in a fast radio burst. *Nature* **596**, 505–508 (2021). <https://doi.org/10.1038/s41586-021-03724-8>

- Received: 15 December 2020

- Accepted: 14 June 2021
  - Published: 25 August 2021
  - Issue Date: 26 August 2021
  - DOI: <https://doi.org/10.1038/s41586-021-03724-8>
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03724-8>

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

- Article
- [Published: 25 August 2021](#)

# Universal pair polaritons in a strongly interacting Fermi gas

- [Hideki Konishi](#) [ORCID: orcid.org/0000-0001-5803-9133<sup>1 na1</sup>](#),
- [Kevin Roux](#) [ORCID: orcid.org/0000-0003-4832-7726<sup>1 na1</sup>](#),
- [Victor Helson](#) [ORCID: orcid.org/0000-0003-0499-4498<sup>1</sup>](#) &
- [Jean-Philippe Brantut](#) [ORCID: orcid.org/0000-0002-4610-6682<sup>1</sup>](#)

[Nature](#) volume **596**, pages 509–513 (2021)

- 2382 Accesses
- 45 Altmetric
- [Metrics details](#)

## Subjects

- [Atomic and molecular interactions with photons](#)
- [Ultracold gases](#)

## Abstract

Cavity quantum electrodynamics (QED) manipulates the coupling of light with matter, and allows several emitters to couple coherently with one light mode<sup>1</sup>. However, even in a many-body system, the light–matter coupling mechanism has so far been restricted to one-body processes. Leveraging cavity QED for the quantum simulation of complex, many-body systems has thus far relied on multi-photon processes, scaling down the light–matter

interaction to the low energy and slow time scales of the many-body problem<sup>2,3,4,5</sup>. Here we report cavity QED experiments using molecular transitions in a strongly interacting Fermi gas, directly coupling cavity photons to pairs of atoms. The interplay of strong light–matter and strong interparticle interactions leads to well-resolved pair polaritons—hybrid excitations coherently mixing photons, atom pairs and molecules. The dependence of the pair-polariton spectrum on interatomic interactions is universal, independent of the transition used, demonstrating a direct mapping between pair correlations in the ground state and the optical spectrum. This represents a magnification of many-body effects by two orders of magnitude in energy. In the dispersive regime, it enables fast, minimally destructive measurements of pair correlations, and opens the way to their measurement at the quantum limit and their coherent manipulation using dynamical, quantized optical fields.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

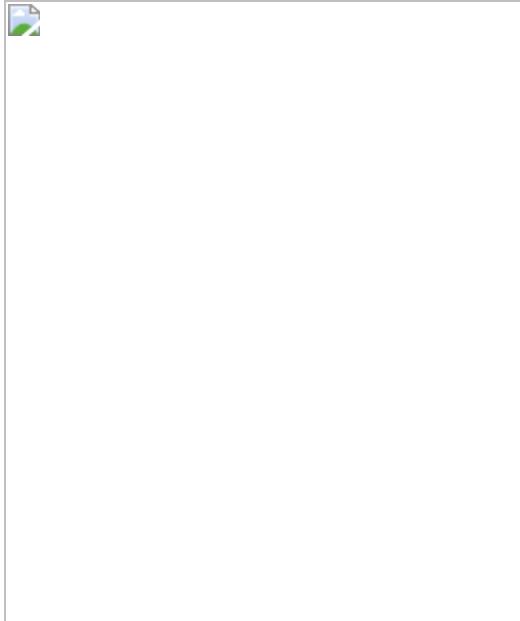
[Rent or Buy](#)

All prices are NET prices.

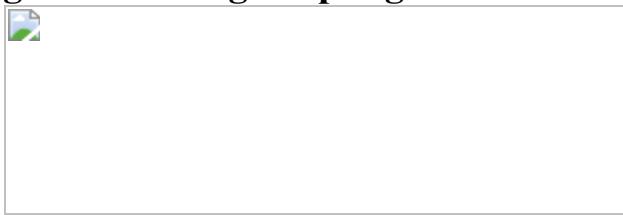
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

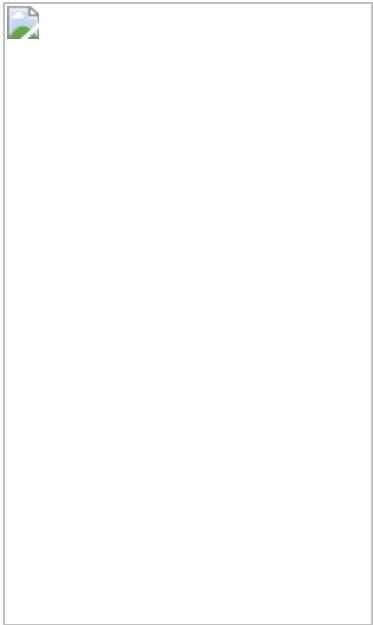
**Figure 1: Concept of the experiment.**



**Figure 2: Strong coupling on PA transitions.**



**Figure 3: Interaction-dependent photon–pair coupling.**



**Figure 4: Single-shot, repeated measurement of pair correlations.**



## Data availability

All data files are available from the corresponding author upon request. Accompanying data, including those for figures, are available from the Zenodo repository (<https://doi.org/10.5281/zenodo.4896757>).

## References

1. 1.

Haroche, S. & Raimond, J.-M. *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford Univ. Press, 2006).

2. 2.

Klinder, J., Keßler, H., Bakhtiari, M. R., Thorwart, M. & Hemmerich, A. Observation of a superradiant mott insulator in the dicke-hubbard model. *Phys. Rev. Lett.* **115**, 230403 (2015).

3. 3.

Landig, R. *et al.* Quantum phases from competing short- and long-range interactions in an optical lattice. *Nature* **532**, 476–479 (2016).

4. 4.

Vaidya, V. D. *et al.* Tunable-range, photon-mediated atomic interactions in multimode cavity qed. *Phys. Rev. X* **8**, 011002 (2018).

5. 5.

Norcia, M. A. *et al.* Cavity-mediated collective spin-exchange interactions in a strontium superradiant laser. *Science* **361**, 259–262 (2018).

6. 6.

Tanji-Suzuki, H. *et al.* in *Advances in Atomic, Molecular, and Optical Physics* Vol. 60 (eds Arimondo, E., P.R. Berman, P. R. & Lin, C. C.) Ch. 4 (Elsevier, 2011).

7. 7.

Reiserer, A. & Rempe, G. Cavity-based quantum networks with single atoms and optical photons. *Rev. Mod. Phys.* **87**, 1379–1418 (2015).

8. 8.

Krantz, P. *et al.* A quantum engineer’s guide to superconducting qubits. *Appl. Phys. Rev.* **6**, 021318 (2019).

9. 9.

Mivehvar, F., Piazza, F., Donner, T. & Ritsch, H. Cavity qed with quantum gases: new paradigms in many-body physics. Preprint at <https://arxiv.org/abs/1904.06560> (2021).

10. 10.

Münstermann, P., Fischer, T., Maunz, P., Pinkse, P. W. H. & Rempe, G. Observation of cavity-mediated long-range light forces between strongly coupled atoms. *Phys. Rev. Lett.* **84**, 4068–4071 (2000).

11. 11.

Mottl, R. *et al.* Roton-type mode softening in a quantum gas with cavity-mediated long-range interactions. *Science* **336**, 1570–1573 (2012).

12. 12.

Ritsch, H., Domokos, P., Brennecke, F. & Esslinger, T. Cold atoms in cavity-generated dynamical optical potentials. *Rev. Mod. Phys.* **85**, 553–601 (2013).

13. 13.

Dogra, N. *et al.* Dissipation-induced structural instability and chiral dynamics in a quantum gas. *Science* **366**, 1496–1499 (2019).

14. 14.

Mekhov, I. B., Maschler, C. & Ritsch, H. Probing quantum phases of ultracold atoms in optical lattices by transmission spectra in cavity quantum electrodynamics. *Nat. Phys.* **3**, 319–323 (2007).

15. 15.

Jones, K. M., Tiesinga, E., Lett, P. D. & Julienne, P. S. Ultracold photoassociation spectroscopy: long-range molecules and atomic scattering. *Rev. Mod. Phys.* **78**, 483–535 (2006).

16. 16.

Partridge, G. B., Strecker, K. E., Kamar, R. I., Jack, M. W. & Hulet, R. G. Molecular probe of pairing in the bcs-bcs crossover. *Phys. Rev. Lett.* **95**, 020404 (2005).

17. 17.

Kinoshita, T., Wenger, T. & Weiss, D. S. Local pair correlations in one-dimensional bose gases. *Phys. Rev. Lett.* **95**, 190406 (2005).

18. 18.

Semczuk, M., Gunton, W., Bowden, W. & Madison, K. W. Anomalous behavior of dark states in quantum gases of  ${}^6\text{Li}$ . *Phys. Rev. Lett.* **113**, 055302 (2014).

19. 19.

Liu, X.-P. *et al.* Observation of the density effect on the closed-channel fraction in a  ${}^6\text{Li}$  superfluid. Preprint at <https://arxiv.org/abs/1903.12321> (2019).

20. 20.

Paintner, T. *et al.* Pair fraction in a finite-temperature fermi gas on the bcs side of the bcs-bec crossover. *Phys. Rev. A* **99**, 053617 (2019).

21. 21.

Yan, M., DeSalvo, B. J., Huang, Y., Naidon, P. & Killian, T. C. Rabi oscillations between atomic and molecular condensates driven with coherent one-color photoassociation. *Phys. Rev. Lett.* **111**, 150402 (2013).

22. 22.

Taie, S., Watanabe, S., Ichinose, T. & Takahashi, Y. Feshbach-resonance-enhanced coherent atom-molecule conversion with

ultranarrow photoassociation resonance. *Phys. Rev. Lett.* **116**, 043202 (2016).

23. 23.

Roux, K., Konishi, H., Helson, V. & Brantut, J.-P. Strongly correlated fermions strongly coupled to light. *Nature Communications* **11**, 2974 (2020).

24. 24.

Giorgini, S., Pitaevskii, L. P. & Stringari, S. Theory of ultracold atomic fermi gases. *Rev. Mod. Phys.* **80**, 1215–1274 (2008).

25. 25.

Zwerger, W. (ed.) *The BCS-BEC Crossover and the Unitary Fermi Gas* Vol. 836 (Springer, 2012).

26. 26.

Stamper-Kurn, D. M. *Cavity Optomechanics with Cold Atoms*, 283–325 (Springer, 2014).

27. 27.

Abraham, E. R. I., Ritchie, N. W. M., McAlexander, W. I. & Hulet, R. G. Photoassociative spectroscopy of long range states of ultracold  ${}^6\text{Li}_2$  and  ${}^7\text{Li}_2$ . *J. Chem. Phys.* **103**, 7773–7778 (1995).

28. 28.

Tan, S. Energetics of a strongly correlated fermi gas. *Ann. Phys.* **323**, 2952–2970 (2008).

29. 29.

Hu, H., Liu, X.-J. & Drummond, P. D. Universal contact of strongly interacting fermions at finite temperatures. *N. J. Phys.* **13**, 035007 (2011).

30. 30.

Hoinka, S. *et al.* Precise determination of the structure factor and contact in a unitary fermi gas. *Phys. Rev. Lett.* **110**, 055305 (2013).

31. 31.

Kuhnle, E. D. *et al.* Universal behavior of pair correlations in a strongly interacting fermi gas. *Phys. Rev. Lett.* **105**, 070402 (2010).

32. 32.

Sagi, Y., Drake, T. E., Paudel, R. & Jin, D. S. Measurement of the homogeneous contact of a unitary fermi gas. *Phys. Rev. Lett.* **109**, 220402 (2012).

33. 33.

Chang, R. *et al.* Momentum-resolved observation of thermal and quantum depletion in a bose gas. *Phys. Rev. Lett.* **117**, 235303 (2016).

34. 34.

Laurent, S. *et al.* Connecting few-body inelastic decay to quantum correlations in a many-body system: a weakly coupled impurity in a resonant fermi gas. *Phys. Rev. Lett.* **118**, 103403 (2017).

35. 35.

Carcy, C. *et al.* Contact and sum rules in a near-uniform fermi gas at unitarity. *Phys. Rev. Lett.* **122**, 203401 (2019).

36. 36.

Mukherjee, B. *et al.* Spectral response and contact of the unitary fermi gas. *Phys. Rev. Lett.* **122**, 203402 (2019).

37. 37.

Roux, K., Helson, V., Konishi, H. & Brantut, J. P. Cavity-assisted preparation and detection of a unitary Fermi gas. *New J. Phys.* **23**, 043029 (2021).

38. 38.

Uchino, S., Ueda, M. & Brantut, J.-P. Universal noise in continuous transport measurements of interacting fermions. *Phys. Rev. A* **98**, 063619 (2018).

39. 39.

Yamazaki, R., Taie, S., Sugawa, S. & Takahashi, Y. Submicron spatial modulation of an interatomic interaction in a bose-einstein condensate. *Phys. Rev. Lett.* **105**, 050405 (2010).

40. 40.

Enss, T. & Thywissen, J. H. Universal spin transport and quantum bounds for unitary fermions. *Annu. Rev. Condens. Matter Phys.* **10**, 85–106 (2019).

41. 41.

Amico, A. *et al.* Time-resolved observation of competing attractive and repulsive short-range correlations in strongly interacting fermi gases. *Phys. Rev. Lett.* **121**, 253602 (2018).

42. 42.

Krinner, S., Esslinger, T. & Brantut, J.-P. Two-terminal transport measurements with cold atoms. *J. Phys. Condens. Matter* **29**, 343003 (2017).

43. 43.

Zeiher, J., Wolf, J., Isaacs, J. A., Kohler, J. & Stamper-Kurn, D. M. Tracking evaporative cooling of a mesoscopic atomic quantum gas in real time. Preprint at <https://arxiv.org/abs/2012.01280> (2020).

44. 44.

Eisele, M., Maier, R. A. W. & Zimmermann, C. Fast in situ observation of atomic feshbach resonances by photoassociative ionization. *Phys. Rev. Lett.* **124**, 123401 (2020).

45. 45.

Bohn, J. L., Rey, A. M. & Ye, J. Cold molecules: progress in quantum engineering of chemistry and quantum matter. *Science* **357**, 1002–1010 (2017).

46. 46.

Pérez-Ros, J., Kim, M. E. & Hung, C.-L. Ultracold molecule assembly with photonic crystals. *N. J. Phys.* **19**, 123035 (2017).

47. 47.

Kampschulte, T. & Denschlag, J. H. Cavity-controlled formation of ultracold molecules. *N. J. Phys.* **20**, 123015 (2018).

48. 48.

Wellnitz, D., Schütz, S., Whitlock, S., Schachenmayer, J. & Pupillo, G. Collective dissipative molecule formation in a cavity. *Phys. Rev. Lett.* **125**, 193201 (2020).

49. 49.

Côté, R., Dalgarno, A., Sun, Y. & Hulet, R. G. Photoabsorption by ultracold atoms and the scattering length. *Phys. Rev. Lett.* **74**, 3581–3584 (1995).

# Acknowledgements

We thank T. Donner for discussions and careful reading of the manuscript; R. Hulet, P. Julienne and J. Hutson for discussions; C. Vale, H. Hu and J. Drut for providing the contact data; and T. Zwettler for experimental assistance. We acknowledge funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 714309), the Swiss National Science Foundation (grant no. 184654), the Sandoz Family Foundation-Monique de Meuron program for Academic Promotion and EPFL.

# Author information

## Author notes

1. These authors contributed equally: H. Konishi, K. Roux

## Affiliations

1. Institute of Physics, EPFL, Lausanne, Switzerland

Hideki Konishi, Kevin Roux, Victor Helson & Jean-Philippe Brantut

## Contributions

H.K., K.R. and V.H. performed the experiments and processed the data; H.K., K.R. and J.-P.B. wrote the paper; J.-P.B. planned and supervised the project.

## Corresponding author

Correspondence to [Jean-Philippe Brantut](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

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

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

## Extended data figures and tables

### [Extended Data Figure 1 Fit to the spectrum.](#)

**a**, Spectrum of PA4 at 832 G averaged over three realizations. **b**, Spectrum reconstructed by equation (2) using the fit results. The solid and dashed lines indicate the fitted positions of the PA resonance and the dispersively shifted cavity resonance. The colour scale is identical to that in the main text.

### [Extended Data Figure 2 Magnetic field dependence of the binding energies.](#)

Positions of PA1–PA4 (green open circles, purple filled diamonds, orange open squares and light blue crosses, respectively) as a function of magnetic fields. The value at 730 G is subtracted for clarity. Linear fits presented by the solid lines yield 0.31, 0.67, 0.89 and  $-0.83 \text{ MHz G}^{-1}$  for the four PA resonances, respectively.

## Rights and permissions

### [Reprints and Permissions](#)

## About this article

### Cite this article

Konishi, H., Roux, K., Helson, V. *et al.* Universal pair polaritons in a strongly interacting Fermi gas. *Nature* **596**, 509–513 (2021).  
<https://doi.org/10.1038/s41586-021-03731-9>

- Received: 31 January 2021
- Accepted: 16 June 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03731-9>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03731-9>

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

- Article
- Open Access
- [Published: 25 August 2021](#)

# Sympathetic cooling of a trapped proton mediated by an LC circuit

- [M. Bohman](#) ORCID: [orcid.org/0000-0003-1322-1489<sup>1,2</sup>](https://orcid.org/0000-0003-1322-1489),
- [V. Grunhofer<sup>3</sup>](#),
- [C. Smorra](#) ORCID: [orcid.org/0000-0001-5584-7960<sup>2,3</sup>](https://orcid.org/0000-0001-5584-7960),
- [M. Wiesinger](#) ORCID: [orcid.org/0000-0003-0111-8812<sup>1,2</sup>](https://orcid.org/0000-0003-0111-8812),
- [C. Will](#) ORCID: [orcid.org/0000-0003-4622-7799<sup>1</sup>](https://orcid.org/0000-0003-4622-7799),
- [M. J. Borchert<sup>2,4,5</sup>](#),
- [J. A. Devlin<sup>2,6</sup>](#),
- [S. Erlewein<sup>2,6</sup>](#),
- [M. Fleck](#) ORCID: [orcid.org/0000-0003-4114-1902<sup>2,1</sup>](https://orcid.org/0000-0003-4114-1902),
- [S. Gavranovic<sup>3</sup>](#),
- [J. Harrington<sup>1,2</sup>](#),
- [B. Latacz](#) ORCID: [orcid.org/0000-0003-2320-1713<sup>2</sup>](https://orcid.org/0000-0003-2320-1713),
- [A. Mooser<sup>1</sup>](#),
- [D. Popper<sup>3</sup>](#),
- [E. Wursten](#) ORCID: [orcid.org/0000-0002-2413-2214<sup>2,6</sup>](https://orcid.org/0000-0002-2413-2214),
- [K. Blaum](#) ORCID: [orcid.org/0000-0003-4468-9316<sup>1</sup>](https://orcid.org/0000-0003-4468-9316),
- [Y. Matsuda](#) ORCID: [orcid.org/0000-0002-9847-3791<sup>1</sup>](https://orcid.org/0000-0002-9847-3791),
- [C. Ospelkaus](#) ORCID: [orcid.org/0000-0002-4170-2936<sup>4,5</sup>](https://orcid.org/0000-0002-4170-2936),
- [W. Quint<sup>8</sup>](#),
- [J. Walz<sup>3,9</sup>](#),
- [S. Ulmer](#) ORCID: [orcid.org/0000-0002-4185-4147<sup>2</sup>](https://orcid.org/0000-0002-4185-4147) &
- [BASE Collaboration](#)

*Nature* volume **596**, pages 514–518 (2021)

- 1336 Accesses
- 89 Altmetric
- [Metrics details](#)

## Subjects

- [Atomic and molecular physics](#)
- [Exotic atoms and molecules](#)
- [Quantum physics](#)

## Abstract

Efficient cooling of trapped charged particles is essential to many fundamental physics experiments<sup>1,2</sup>, to high-precision metrology<sup>3,4</sup> and to quantum technology<sup>5,6</sup>. Until now, sympathetic cooling has required close-range Coulomb interactions<sup>7,8</sup>, but there has been a sustained desire to bring laser-cooling techniques to particles in macroscopically separated traps<sup>5,9,10</sup>, extending quantum control techniques to previously inaccessible particles such as highly charged ions, molecular ions and antimatter. Here we demonstrate sympathetic cooling of a single proton using laser-cooled Be<sup>+</sup> ions in spatially separated Penning traps. The traps are connected by a superconducting LC circuit that enables energy exchange over a distance of 9 cm. We also demonstrate the cooling of a resonant mode of a macroscopic LC circuit with laser-cooled ions and sympathetic cooling of an individually trapped proton, reaching temperatures far below the environmental temperature. Notably, as this technique uses only image-current interactions, it can be easily applied to an experiment with antiprotons<sup>1</sup>, facilitating improved precision in matter–antimatter comparisons<sup>11</sup> and dark matter searches<sup>12,13</sup>.

[Download PDF](#)

## Main

Measurements of the charge-to-mass ratio and *g*-factor of the proton and antiproton, a prominent, stable particle–antiparticle system, are limited by cryogenic particle temperatures<sup>14,15,16</sup>. However, with no electronic structure, (anti)protons (protons and antiprotons) are not amenable to standard laser cooling techniques. Moreover, ions that are easily laser cooled are not readily trapped in the same potential well as negatively charged antiprotons or antimatter molecular ions (for example,  $\langle \bar{H} \rangle^2 \langle - \rangle$ )<sup>17</sup>. Sympathetic laser cooling with negatively charged ions<sup>18,19,20</sup> and with microscopically fabricated trapping potentials<sup>1,21,22</sup> have been proposed. Another technique, proposed over 30 years ago, extended laser cooling to exotic systems by coupling via induced image currents in trap electrodes to ions with a well-suited cooling transition<sup>9</sup>. Similarly, coupling laser addressable ions to systems

with no optical structure is sought after in precision spectroscopy<sup>23,24</sup>, mass measurements<sup>25</sup>, quantum information<sup>10</sup> and quantum engineering<sup>5</sup>.

We demonstrate sympathetic cooling of a single proton, extending the image-current coupling technique with a superconducting LC circuit that resonantly enhances energy exchange between the proton and laser-cooled ions. We use a cryogenic multi-Penning-trap system to store a single proton in the proton trap and a cloud of  $\text{Be}^+$  ions in a beryllium trap, separated axially by around 9 cm (Fig. 1a). A homogeneous magnetic field  $B$  parallel to the electrode axis and an electric quadrupole potential at voltage  $V_0$  confines the particles and gives rise to circular magnetron and modified cyclotron motion in the radial plane, and harmonic axial motion, at frequencies  $v_-$ ,  $v_+$  and  $v_z$ , respectively<sup>26</sup>. LC resonators with high quality factors (in our case  $Q \sim 15,000$ ) are commonly used to detect image currents in trap electrodes<sup>27</sup> and, as shown in Fig. 1a, we connect the resonator to both traps so that the two ion-trap systems are coupled via particle-induced image currents. The LC circuit with total capacitance  $C_R \approx 36 \text{ pF}$  and inductance  $L_R \approx 3.0 \text{ mH}$  has an equivalent parallel resistance, at resonance frequency  $v_0$ , of  $R_p = 2\pi v_0 L_R Q$ . The entire system is modelled by an equivalent circuit where the proton and  $\text{Be}^+$  ions are series LC circuits with capacitances and inductances  $C_p$ ,  $C_{\text{Be}}$ ,  $L_p$  and  $L_{\text{Be}}$  (ref. 28) connected in parallel to the superconducting LC circuit (Fig. 1b). The  $\text{Be}^+$  ions are also damped by the cooling laser, represented as a variable virtual resistance  $R_L$  (ref. 9). In contrast to the proposal in which energy is exchanged between the ion-trap systems via only a shared electrode<sup>9</sup>, we use the LC-circuit resonator to couple the axial modes of the trapped particles. On resonance, the large inductance of the resonator coil compensates the electrode capacitance and enhances the ion-induced image current by the  $Q$ -value. With mechanically machined traps used for precision Penning-trap experiments, the  $\sim 10\text{-mHz}$  coupling rates expected from the non-resonant proposal<sup>11</sup> require minute-scale cooling cycles and are limited by the loss of resonant coupling; for example, from voltage fluctuations of the trapping potential. For the parameters used in our resonant cooling demonstration, energy is exchanged between the proton and sympathetically laser-cooled resonator at a rate of 2.6 Hz (measured by the dip width on resonance) so that thermal equilibrium is reached within seconds and the axial frequencies of the two species are easily matched. Notably, by coupling the ion-trap systems via the resonator, the coupling does not rely on a shared electrode, so the energy exchange rate is not limited by the trap capacitance. Consequently, this cooling scheme can be realised over long distances and with several distributed ion traps.

**Fig. 1: Experimental setup.**

---

 **figure1**

---

**a**, A single proton is stored in the proton trap while one or more  $\text{Be}^+$  ions are stored in the beryllium trap. The two ion traps, with inner diameters  $d = 9 \text{ mm}$  and  $d = 5 \text{ mm}$ , respectively, are connected to a cryogenic LC circuit with resonance frequency near their axial frequencies. One end of the resonance circuit is connected to a cryogenic amplifier while the other is connected to rf ground. **b**, This system is described by a three-part equivalent circuit, two series LC circuits representing the trapped particles, and the LC resonator with effective parallel resistance  $R_p$  as a parallel RLC circuit. The amplifier is used to read out the image-current signal of the circuit, and drives the system with voltage noise at an effective temperature  $T_0$ . **c**, The resulting FFT spectrum consists of the broad resonance of the detector and two narrow ‘dips’ with  $\sim 0.8 \text{ Hz}$  FWHM for a single  $\text{Be}^+$  ion and  $\sim 2.6 \text{ Hz}$  FWHM for the proton at the axial oscillation frequencies of the two trapped-ion systems.

### Source data

The noise spectrum of this coupled system is shown in a fast Fourier transform (FFT) of the voltage signal of the resonator (Fig. 1c). The entire system is driven by a combination of the Johnson noise of the resonator and additional voltage noise from the cryogenic amplifier, resulting in an effective noise temperature  $T_0 = 17.0(2.4) \text{ K}$  (where the number in parentheses is the  $1\sigma$  uncertainty). Here, the axial frequencies of the proton in the proton trap and the  $\text{Be}^+$  ion in the beryllium trap were set close to resonance with the LC circuit by adjusting the voltage of the axial potential in each trap,  $v_z \propto V_0^{1/2}$ . In the measured noise spectrum, the detector appears as a broad  $\sim 40 \text{ Hz}$  (full-width at half-maximum (FWHM)) resonance while the proton and the  $\text{Be}^+$  ions short the parallel resistance of the resonator and appear as narrow dips with widths determined by the charge-to-mass ratios and the trap diameter<sup>28</sup>.

We demonstrate that the proton,  $\text{Be}^+$  ions and resonator form a system of three coupled oscillators by measuring the noise spectrum at thermal equilibrium. We detune both ion species around three resonator linewidths away from the LC-circuit resonance frequency  $v_0 \approx 479,000 \text{ Hz}$  to observe coupling signatures via the FFT lineshape. In these measurements, we store the proton in the proton trap at constant axial frequency

and gradually increase the axial frequency of a single  $\text{Be}^+$  ion in the beryllium trap. The resulting FFT spectra show observed particle frequencies at the dip positions (Fig. 2a, dark blue) and two of the normal modes of the coupled three-oscillator system at the maxima (Fig. 2a, red). Near the p- $\text{Be}^+$  resonance, the axial motion of both particles is no longer determined by the trapping potential alone, and we observe the coupling signature in two of the normal modes of the three-oscillator system. This feature is consistent with the analytical solution derived from the impedance of the circuit model in Fig. 2b (see [Methods](#)) and appears when the three oscillators exchange energy. Using simulations (described in the Methods), we show the corresponding time-domain behaviour in the presence of the environmental noise (Fig. 2c). In the absence of environmental noise, the energy of each oscillator as a function of time is deterministic and can be found from the initial phases and energy exchange rates. With environmental noise included, energy is still exchanged and, as shown here when laser cooling is absent, the oscillator energies are determined by this equivalent noise temperature.

**Fig. 2: Remote coupling measurements.**



- a**, Measured FFT spectra of a single  $\text{Be}^+$  ion and the proton are presented as vertical cuts of the heatmap. The axial frequency of the proton is fixed and the ring voltage of the trapping electrodes of the Beryllium trap (BT) is scanned one step upwards after each FFT is recorded. A vertical cross-cut looks similar to the spectrum shown in Fig. 1a, although as the dip features are shifted approximately 130 Hz upwards, they exhibit dispersive behaviour. **b**, The noise spectra expected from the impedance  $Z(\nu)$  of the equivalent circuit in Fig. 1b are presented in the same way as in **a** and are

calculated using the same parameters as in the experiment. **c**, An example of the energy exchange between the three oscillators at the resonance frequency of the LC circuit is simulated and includes the energy fluctuations from environmental noise of the system. The energy of each oscillator depends on the phase relation with the environmental noise, so the time series shown here is one of infinitely many that could produce the FFT spectrum used in Fig. 2a. **d**, The evolution of the standard deviation of the proton energy scatter is shown, measured via energy dependent axial frequency shifts. We see an unambiguous increase in energy when the excited  $\text{Be}^+$  ions are tuned to resonance with the proton (orange points), compared to when the  $\text{Be}^+$  ions are detuned (blue points). The excitation drive remains on for both points and can be compared to the drive off condition (green points). Error bars,  $1\sigma$  error of the standard deviation.

### Source data

We further demonstrate that the temperature of the proton can be modified by coupling to a cloud of excited  $\text{Be}^+$  ions, here consisting of around 15 ions. To this end, we apply a parametric rf drive at  $2\nu_0$  in the beryllium trap, which excites the  $\text{Be}^+$  ions if  $\nu_{z,\text{Be}} = \nu_0$  but, as confirmed by background measurements (see [Methods](#)), has no direct effect on the proton in the proton trap. By bringing the proton into resonance with the weakly excited  $\text{Be}^+$  ions the  $\text{Be}^+$  ions appear as a broad, shallow dip and the sympathetically excited proton appears as a narrow peak (see [Methods](#)). To quantify the energy transferred to the proton, we measure the axial frequency of the proton before and after coupling to the excited  $\text{Be}^+$  ions. Coupling the excited axial mode to the cyclotron mode with a sideband drive at  $\nu_+ - \nu_z$  transfers the energy of the axial mode to the cyclotron mode with resulting energy  $E_+ = (\nu_+/\nu_z)E_z$  (ref. [29](#)). Similar to the continuous Stern–Gerlach effect<sup>[30](#)</sup>, the quadratic component of the magnetic field in the proton trap,  $B_2 = -0.39(11) \text{ T m}^{-2}$ , interacts with the magnetic moment of the modified cyclotron mode at energy  $E_{+,p}$ , producing the axial frequency shift  $\Delta\nu_z \propto B_2\Delta E_{+,p}$  (see [Methods](#)), which we measure to determine the change in axial energy of the proton. We show the evolution of the standard deviation of the change in the proton energy  $E_{+,p}$  while the excited  $\text{Be}^+$  ions are tuned to resonance with the proton (Fig. 2d, orange) and when detuned from the resonator (Fig. 2d, blue). During this experiment we interleave on-resonance and off-resonance measurements and see a clear increase in proton energy that arises from the remote, resonator-mediated coupling to the excited ions. In contrast to the off-resonant points, the resonance points exhibit scatter that is nearly three orders of magnitude larger, with a statistical significance of more than  $20\sigma$ . The excitation drive remains on during both of the interleaved measurements to ensure that the increased scatter is due only to the ion–proton coupling. In addition, we constrain variations in the resonator temperature due

to off-resonant coupling to the excitation drive (see [Methods](#)). For comparison, the scatter when the drive is off is also shown in Fig. [2d](#) (green points).

Our demonstration of sympathetic cooling employs similar axial frequency shift measurements in the presence of a continuously laser-cooled Be<sup>+</sup> ion cloud. The Be<sup>+</sup> ions are cooled with the closed  $^2\text{S}_{1/2} \rightarrow ^2\text{P}_{3/2}$  transition and tuned to resonance with the superconducting circuit and the proton (Fig. [3a](#)). The cooling laser damps the axial motion, increasing the equivalent resistance  $R_L$  (Fig. [1b](#)) and reducing the signal of the broad Be<sup>+</sup> dip. The laser-cooled ions reduce the effective noise temperature in the entire circuit and lower the temperature in a narrow frequency range. Using the narrow proton dip as a temperature sensor for the cooled common mode of the system, we determine the temperature reduction experimentally with well understood energy-dependent shifts of the axial dip and develop further insight into the cooling using time-domain simulations.

**Fig. 3: Sympathetic cooling demonstration.**

---

 figure3

**a**, A cloud of laser-cooled  $\text{Be}^+$  ions appears on the resonator as a broad dip that reduces the temperature of that resonator mode. The proton spectral dip is much narrower and continues to short the resonator noise, and can be used as a temperature probe via equation (3). **b**, The frequency shift  $\Delta\nu_z$ , from which  $\Delta T$  is extracted, is shown for  $N_{\text{Be}} \approx 10$  and laser detuning  $\delta \approx 100$ , as a function of the tuning ratio  $\Delta\text{TR}$ , given by  $V_{\text{CE}}/V_0$ . Each data point is an average of several measurements. Error bars, standard deviation of the temperature reduction that arises from the scatter of the final proton temperature. The change in temperature is extracted from the slope, which,

when combined with the resonator temperature, allows us to extract the final temperature of the proton.

### Source data

A symmetric, cylindrical Penning trap provides a high degree of control over the trapping potential. We use a deliberately introduced trap anharmonicity in the proton trap that shifts the axial frequency by

$$\Delta \nu_z(\text{TR}, T) = \frac{1}{2} \frac{4\pi^2 m \nu_z}{C_4 (\text{TR})} \frac{C_2 (\text{TR}) k_B T}{(1)}$$

Here,  $C_n(\text{TR})$  are the coefficients of the expansion of the local trapping potential along the trap axis that depend on the ratio of voltages applied to the central ring electrode ( $V_0$ ), and the two nearest correction electrodes ( $V_{\text{CE}}$ )<sup>26,31,32</sup>, referred to as the tuning ratio,  $\text{TR} = V_{\text{CE}}/V_0$ . When the laser-cooled  $\text{Be}^+$  ions are tuned to resonance, the noise energy of the common mode of the proton, resonator and  $\text{Be}^+$  ions is reduced from the noise temperature of the environment resulting in an axial frequency shift,

$$\Delta \nu_z(T_0, T_p) = \nu_{z,1}(\text{TR}, T_p) - \nu_{z,2}(\text{TR}, T_p), \quad (2)$$

$$= \kappa \Delta T \times (T_0 - T_p), \quad (3)$$

where  $\nu_{z,1}(\text{TR}, T_0) = \nu_z + \delta\nu_z(\text{TR}, T_0)$  is the axial frequency measured at  $T_0$  when the  $\text{Be}^+$  ions are detuned and  $\nu_{z,2}(\text{TR}, T_p) = \nu_z + \delta\nu_z(\text{TR}, T_p)$  is the axial frequency measured when laser-cooled ions are in resonance and reduce the temperature to  $T_p$ . The trap anharmonicity is characterized by an offset from the ideal tuning ratio  $\Delta\text{TR} = \text{TR} - \text{TR}(C_4 = 0)$  and a constant determined from the trap geometry  $\kappa = 45.4 \text{ Hz K}^{-1}$  that we crosscheck with additional measurements that use electronic feedback to change the temperature of the resonator. We measure  $\Delta\nu_z$  as a function of  $\Delta\text{TR}$  and the measured slope  $s$  determines the change in temperature,  $\Delta T = T_0 - T_p = -s/\kappa$ . The results of an example measurement are shown in blue in Fig. 3b. With ten  $\text{Be}^+$  ions in resonance, we measure a slope  $s = -350(14) \text{ Hz}$  and in a background measurement with the  $\text{Be}^+$  ions detuned (Fig. 3b, orange), obtain a slope  $s = 4(13) \text{ Hz}$ . This corresponds to a temperature reduction of  $\Delta T = 7.7(0.3) \text{ K}$  and demonstrates sympathetic laser cooling of a single trapped proton. With a significance of more than 20 standard deviations, this is also a demonstration of remote, image-current mediated

sympathetic cooling, applicable to any charged particle without convenient cooling transitions.

The temperature of the proton is determined by the noise power dissipated by the laser-cooled ions. In the circuit representation, increasing the damping of the laser cooling  $\gamma_L$  increases  $R_L$  and has the effect of lowering the coupling rate of the  $\text{Be}^+$  ions to the resonator  $(\{\mathop{\gamma}\limits^{\sim}\}_{\rm Be})$ , which, in the absence of laser cooling is given by the dip width  $\gamma_{\text{Be}} \propto N_{\text{Be}}$ . For a given number of laser-cooled ions  $N_{\text{Be}}$ ,  $\gamma_L$  must be optimized and in the limiting case when  $\gamma_L \ll \gamma_{\text{Be}}$  the  $\text{Be}^+$  ions are driven by the resonator and the dip signal is unchanged. Likewise, when  $\gamma_L \gg \gamma_{\text{Be}}$ , the  $\text{Be}^+$  ions are decoupled from the resonator and the dip signal vanishes. In both limiting cases, the temperature of the resonator and the proton remain unchanged.

However, increasing  $N_{\text{Be}}$  increases  $(\{\mathop{\gamma}\limits^{\sim}\}_{\rm Be})$  and laser cooling reduces the temperature of the resonator and the proton, even at large  $\gamma_L$ . To lower the temperature and to investigate the scaling of  $T_p$ , we performed a series of further measurements with varying  $N_{\text{Be}}$  and laser detuning  $\delta$  (Fig. 4). We additionally analysed the temperature scaling by comparing to a temperature model in which the common mode temperature  $T_{\text{CM}}$  of the equivalent circuit arises from competing dissipation sources; the noise temperature of the environment,  $T_0$ , at a coupling rate given by the width of the LC resonance  $\gamma_D$ , and to the  $\text{Be}^+$  ions at temperature  $T_{\text{Be}}$ . As a result, the system comes to thermal equilibrium at

$$\$ \$ \{T\}_{\rm p} = \{T\}_{\rm CM} \{T\}_0 + \frac{\{T\}_{\rm Be}}{\{T\}_{\rm D}} \quad (4)$$

$$\$ \$ \{T\}_{\rm CM} = \frac{\{T\}_0}{\{T\}_{\rm D}} + \frac{\{T\}_{\rm Be}}{\{T\}_{\rm D}} \quad (5)$$

When  $(\{\mathop{\gamma}\limits^{\sim}\}_{\rm Be}) \gg \{T\}_{\rm D}$ , the proton temperature is approximated as

$$\$ \$ \{T\}_{\rm p} \approx \frac{\{T\}_0}{\{T\}_{\rm D}} + \frac{\{T\}_{\rm Be}}{\{T\}_{\rm D}} \quad (6)$$

reproducing the  $1/N_{\text{Be}}$  scaling, by  $\gamma_{\text{Be}} \propto N_{\text{Be}}$ , appearing in the non-resonant proposal<sup>9,11</sup> and related proposals in the context of trapped ion quantum information<sup>10,33,34</sup>. In these measurements, the laser detuning  $\delta$  can be viewed as a tuning parameter that changes the  $\gamma_L$  and subsequently  $(\{\mathop{\gamma}\limits^{\sim}\}_{\rm Be})$  (see [Methods](#)). As a result, the lowest proton temperatures are not found by minimizing the  $\text{Be}^+$  temperature, which would correspond to lower laser detunings, but by maximizing the coupling of the ions to the detector, corresponding to larger  $(\{\mathop{\gamma}\limits^{\sim}\}_{\rm Be})$ . For the experimental parameters used here,  $\delta = -90$  MHz is the largest laser detuning at which the proton dip is still visible. The largest ion cloud  $\gamma_{\text{Be}} = 164(5)$  Hz and largest detuning from the centre of the cooling transition  $\delta = -90$  MHz (Fig. 4) is representative of the lowest temperatures observed in our measurements. We achieve a temperature reduction of

$$\$ \$ \Delta T_{\rm p} = 14.4(0.7) \text{ K}, \quad (7)$$

and using the environment temperature  $T_0$ , we obtain

$$\$ \$ T_{\rm p} = T_0 - \Delta T_{\rm p} = 2.6(2.5) \text{ K}, \quad (8)$$

with uncertainty dominated by the one of  $T_0$ . This measurement demonstrates a temperature reduction of 85%.

**Fig. 4: Temperature scaling.**



The temperature of the cooled proton-resonator system is shown as a function of the laser detuning  $\delta$  and the coupling of the laser-cooled  $\text{Be}^+$  ions to the resonator parametrized by the dip width. Vertical error bars, fit uncertainty of the measured slope. Smaller horizontal error bars, standard error of  $\gamma_{\text{Be}}$ . Each point consists of a minimum of 50 measured frequency shifts, although more data was collected at some detunings, resulting in lower uncertainties. Fits to the data are shown to illustrate the  $1/N_{\text{Be}}$  scaling when the dip width becomes larger than the resonator width; approximately 40 Hz.

### Source data

Lower temperatures can be achieved by lowering the noise temperature of the amplifier,  $T_0$ , increasing the  $Q$ -value of the resonator, or by operating with smaller traps that increase  $\gamma_{\text{Be}}$  quadratically with lower radius. In addition, by performing these demonstration measurements fully on-resonance for maximal coupling rates, the balance of heating by the resonator to cooling by the  $\text{Be}^+$  ions is maximally inefficient, and future cooling work will be done off-resonantly to balance the coupling rate and the temperature limit with engineered cooling sequences<sup>35</sup>.

In the context of our experimental goals, this technique can be readily applied to sympathetically laser cool protons and antiprotons in the same large macroscopic traps that enable precision measurements of the charge-to-mass ratio and  $g$ -factor<sup>1,11</sup>. In addition, while we measure the axial temperature, sideband coupling<sup>29</sup> or axialization<sup>36</sup> can be used to cool the radial motion of the antiproton. In measurements of nuclear magnetic moments, this will enable nearly 100% spin-flip fidelity<sup>11,15,16,37</sup>, and can reduce the dominant systematic effect proportional to the particle temperature in the highest precision mass measurements<sup>38,39,40</sup>. In addition, this technique can be used to cool other exotic systems such as highly charged<sup>23,24</sup> or molecular ions<sup>17,41</sup> and the sympathetically cooled resonator can enhance the sensitivity of dark matter searches<sup>13,41,42</sup>. Ultimately, this demonstration realises a long-sought experimental technique that will enable precision experiments of any charged species at lower temperatures.

## Methods

### Equations of motion for the coupled ion-trap systems

The axial motions of the trapped proton and  $\text{Be}^+$  ions are described by a harmonic oscillator driven by the oscillating voltage on the trap electrodes connected to the

resonant LC circuit,  $V_{LC}$ , and in the case of the  $Be^+$  ion(s) by an additional photon scattering force from the cooling laser,  $F_L$ :

$$\begin{aligned} & \text{\$\$}\begin{array}{cc}\{m\}_{\{\rm{p}\}}\{\ddot{z}\}_{\{\rm{p}\}}+\{m\}_{\{\rm{p}\}}\\\{\omega\}_{z,\{\rm{p}\}}^2\{z\}_{\{\rm{p}\}}&=\frac{q}{D}\{\rm{p}\}\\V_{\{\rm{L}\}\{\rm{C}\}}\backslash\{m\}_{\{\rm{B}\}}\{\rm{e}\}\{\ddot{z}\}_{\{\rm{B}\}}\\\{\rm{e}\}+\{m\}_{\{\rm{B}\}}\{\rm{e}\}\{\omega\}_{z,\{\rm{B}\}}\{\rm{e}\}^2\\z_{\{\rm{B}\}}\{\rm{e}\}&=\frac{q}{D}\{\rm{B}\}\{\rm{e}\}\{V\}_{\{\rm{L}\}}\\\{\rm{C}}\}+\{F\}_{\{\rm{L}\}}.\end{array}\$\$ \\ (9) \end{aligned}$$

$V_{LC}$  is composed of voltage noise from the environment,  $V_{noise}$ , and the voltage arising from image currents induced by the proton and  $Be^+$  ions,  $I_p$  and  $I_{Be}$ , respectively. On resonance  $\omega_R = \omega_{z,p} = \omega_{z,Be}$ , the impedance of the LC circuit is given by its equivalent parallel resistance,  $R_p$ , and

$$\begin{aligned} & \text{\$\$}\begin{array}{c}\frac{1}{\omega_R}\{\rm{R}\}^2\{\ddot{V}\}_{\{\rm{L}\}}\\\{\rm{C}}\}+\frac{1}{Q\omega_R}\{\dot{V}\}_{\{\rm{L}\}\{\rm{C}\}}+\{V\}_{\{\rm{L}\}\{\rm{C}\}}=\{L\}_{\{\rm{R}\}}(\{\dot{I}\}_{\{\rm{n}\}}\{\rm{o}\}\\\{\rm{i}\}\{\rm{s}\}\{\rm{e}\}+\{\dot{I}\}_{\{\rm{p}\}}+\{\dot{I}\}_{\{\rm{B}\}}\\\{\rm{e}}).\end{array}\$\$ \\ (10) \end{aligned}$$

$I_{noise}$  is the noise current from the environment, and  $I_p$  and  $I_{Be}$  are the induced image currents of the proton and the  $Be^+$  ions, respectively:

$$\begin{aligned} & \$\{I\}_{\{\rm{p}\},\{\rm{B}\}}=\frac{q}{D}\{\rm{p}\},\{\rm{B}\}\\\{\rm{e}}\}\{\dot{z}\}_{\{\rm{p}\},\{\rm{B}\}}\{\rm{e}},\$\$ \\ (11) \end{aligned}$$

where  $D_{p,Be}$  are the trap-dependent effective electrode distances<sup>28</sup>.

The equations of motion and the equation for the voltage in the LC circuit form a set of coupled stochastic differential equations without closed analytical solutions available. As a result, we analyse the frequency response of the system by calculating the impedance of the equivalent circuit in Fig. 1b and estimate the energy of the proton by calculating the temperatures of each component based on their energy exchange rates. Finally, we numerically integrate the differential equations in simulations that allow the comparison of FFT spectra and the visualization of the time-domain behaviour in the system.

## Impedance analysis of the equivalent circuit

The FFT spectrum in Fig. 1c results from the noise on the image-current detector ( $\langle \{u\}_n^2 \rangle = 4 \{k\}_{\{\rm rm B\}} \{T\}_0 \{\rm rm R\} \{\rm rm e\} [Z(\omega)] \Delta f$ ) at effective noise temperature  $T_0$ , FFT bandwidth  $\Delta f$ , and the impedance  $Z(\omega)$  of the circuit in Fig. 1b. The lineshape of resistively cooled particles stored in a single trap based on the impedance of the equivalent circuit is well understood<sup>28,43</sup>. Here, we evaluate the impedance for two independently biased ion-trap systems as:

where  $k_L = R_L/R_p$  allows for additional damping in one of the traps. The lineshapes of the individual components arise from  $\delta_i(\omega) = 2(\omega - \omega_i)/\gamma_i$ , which are parameters proportional to the ratio of the frequency detuning ( $\omega - \omega_i$ ) to the oscillator linewidth,  $\gamma_i$ . The index  $i \in \{R, Be, p\}$  relates to the resonator, the  $Be^+$  ions and the proton, respectively. In the absence of additional damping  $k_L = 0$ , the impedance simplifies to

$$\begin{aligned}
 & \frac{\frac{1}{1+\delta}}{\left(\frac{R^2}{R^2+2\delta} - \frac{2\delta}{(R^2+2\delta)^2} + \frac{2}{R^2+2\delta}\right)} \\
 & = \frac{1}{R^2+2\delta} \cdot \frac{1}{1-\frac{2\delta^2}{R^2+2\delta} + \frac{2}{R^2+2\delta}} \\
 & = \frac{1}{R^2+2\delta} \cdot \frac{1}{\frac{R^2+2\delta-2\delta^2+2}{R^2+2\delta}} \\
 & = \frac{1}{R^2+2\delta} \cdot \frac{R^2+2\delta}{R^2+2\delta-2\delta^2+2} \\
 & = \frac{1}{R^2+2\delta-2\delta^2+2}
 \end{aligned}$$

which describes the lineshape of the data shown in Fig. 1c. Similarly, the heat maps in Fig. 2a and Fig. 2b compare the FFT spectra from experiment to ones calculated with  $Z(\omega)$ , and show consistent behaviour.

With laser cooling included,  $R_L > 0$  and the dip feature of the  $\text{Be}^+$  ions is modified as shown in Fig. 3a. The corresponding impedance is calculated for varying  $R_L$  in Extended Data Fig. 1. In both cases, regardless of the value of  $R_L$ , the proton shorts the noise of the LC circuit on resonance. The  $\text{Be}^+$  ions decouple from the LC circuit as  $R_L$  reduces the fraction of noise power dissipated in the series LC circuit of the  $\text{Be}^+$  ions—ultimately leading to a vanishing dip signal. This decoupling effect is well known from other coupled oscillator systems<sup>44</sup> and motivates the reduced coupling of the  $\text{Be}^+$  ions to the LC circuit,  $\langle \langle \mathop{\gamma} \limits^{\sim} \rangle \rangle_{\text{rm}\{\text{Be}\}} < \langle \mathop{\gamma} \limits^{\sim} \rangle_{\langle \langle \text{rm}\{\text{Be}\} \rangle \rangle}$ .

## Temperature model

The temperature model presented here is described in ref. 45 and assumes that each component of the three-oscillator system consisting of the trapped proton, the trapped  $\text{Be}^+$  ion(s) and the resonator comes to thermal equilibrium with the rest of the system at temperatures defined by the energy exchange rates in the system.

The  $\text{Be}^+$  ions are damped by the resonator as well as the cooling laser, and the power transmitted by the  $\text{Be}^+$  ions is then written

$$\langle \langle \frac{dE}{dt} \rangle \rangle = k_{\text{B}} T_{\text{Be}} + \langle \langle \frac{dE}{dt} \rangle \rangle_{\text{laser}}, \quad (14)$$

where  $\langle \langle \frac{dE}{dt} \rangle \rangle_{\text{laser}}$  is the power dissipated by scattered photons. An identical analysis applies to the resonator, which is coupled to the environment with a coupling rate  $\gamma_D$  given by the width of the resonance, or the  $Q$ -value, and to the  $\text{Be}^+$  ions with a coupling rate  $\langle \langle \mathop{\gamma} \limits^{\sim} \rangle \rangle_{\text{rm}\{\text{Be}\}}$ . These relations produce the system of equations shown in the main text,

$$T_p = T_C = T_M, \quad (15)$$

$$\begin{aligned} T_D &= \frac{T_0 \mathop{\gamma} \limits^{\sim} \text{rm}\{D\}}{\mathop{\gamma} \limits^{\sim} \text{rm}\{B\} \mathop{\gamma} \limits^{\sim} \text{rm}\{e\} + \mathop{\gamma} \limits^{\sim} \text{rm}\{D\} \mathop{\gamma} \limits^{\sim} \text{rm}\{B\} \mathop{\gamma} \limits^{\sim} \text{rm}\{e\}} \\ &= \frac{T_0 \mathop{\gamma} \limits^{\sim} \text{rm}\{D\}}{\mathop{\gamma} \limits^{\sim} \text{rm}\{B\} \mathop{\gamma} \limits^{\sim} \text{rm}\{e\}}. \end{aligned} \quad (16)$$

The power dissipated by the resonator while the  $\text{Be}^+$  ions are laser cooled can be written as

$$\begin{aligned} \text{\$}\{k\}_{\{\{\text{rm}\{B\}\}\}}\{\text{T}\}_{\{\{\text{rm}\{C\}\}\{\text{rm}\{M\}\}\}}\{\mathop{\gamma}\limits^{\sim}\}_{\{\{\text{rm}\{B\}\}\{\text{rm}\{e\}\}\}}=\langle\langle\mathop{J}\limits^{\sim}\rangle\rangle_z \\ \{\text{R}\}_{\{\{\text{rm}\{p\}\}\}}, \end{aligned} \quad (17)$$

and in combination with the power dissipated by the resonator in the absence of laser cooling,

$$\begin{aligned} \text{\$}\{k\}_{\{\{\text{rm}\{B\}\}\}}\{\text{T}\}_{\{0\}}\{\gamma\}_{\{\{\text{rm}\{B\}\}\{\text{rm}\{e\}\}\}}=\langle\langle J_z\rangle\rangle \\ \{\text{R}\}_{\{\{\text{rm}\{p\}\}\}}, \end{aligned} \quad (18)$$

allows the reduced coupling rate to be written as

$$\begin{aligned} \text{\$}\{\mathop{\gamma}\limits^{\sim}\}_{\{\{\text{rm}\{Be\}\}\}}=\frac{\{\text{T}\}_{\{0\}}}{\{\text{T}\}_{\{\{\text{rm}\{CM\}\}\}}}k\{\gamma\}_{\{\{\text{rm}\{Be\}\}\}}, \end{aligned} \quad (19)$$

where  $k$  is defined by the ratio

$$\begin{aligned} \text{\$}k=\frac{\langle\langle J_z^2\rangle\rangle}{\langle\langle J_z^2\rangle\rangle}. \end{aligned} \quad (20)$$

Although  $k$  can, in principle, be extracted from the FFT spectrum, the extraction of individual  $k$  values is imprecise and  $k$  and  $T_{\text{Be}}$  are treated as constant fit parameters in Fig. 4. A more accurate determination of  $\langle\langle dE_{\text{laser}}/dt\rangle\rangle$  can be performed by measuring the photon scattering rate<sup>9</sup> and is planned for future measurements.

## Simulations and time-domain behaviour

We access the time-domain behaviour of the proton–ion–resonator system through simulations, which are performed by numerically integrating equation (9) and equation (10). By replacing  $V_{\text{LC}}=L_R\langle\langle\dot{I}_L\rangle\rangle$ , where  $I_L$  is the current flowing through the inductance  $L_R$ , these equations can be rewritten as

$$\begin{aligned} \text{\$}\{m\}_{\{\{\text{rm}\{p\}\}\}}\{\ddot{z}\}_{\{\{\text{rm}\{p\}\}\}}+\{m\}_{\{\{\text{rm}\{p\}\}\}}\{\omega\}_{z,} \\ \{\text{rm}\{p\}\}^2\{z\}_{\{\{\text{rm}\{p\}\}\}}=\frac{q\{\{D\}_{\{\{\text{rm}\{p\}\}\}}\}}{\{\{\text{rm}\{L\}\}\}_{\{\{\text{rm}\{R\}\}\}}}, \end{aligned}$$

$$\{\dot{I}_L\}_{\{L\}}, \{m\}_{\{\{\text{rm}\{B\}\}\{\text{rm}\{e\}\}\}}$$

The integration is performed in time steps of  $\Delta t = 1$  ns for most simulations and the equivalent thermal noise  $\langle \langle I \rangle \rangle_{\{n\}} = \langle \langle I \rangle \rangle_{\{0\}} \langle \langle I \rangle \rangle_{\{s\}} \langle \langle I \rangle \rangle_{\{e\}}^2 / (k_B T_0 \Delta f R_p)$  is computed in each step  $n$  as

$$\text{\$}\{\text{I}_\text{-}\{\{\text{rm}\{\text{n}\}\}, \{\text{rm}\{\text{o}\}\}, \{\text{rm}\{\text{i}\}\}, \{\text{rm}\{\text{s}\}\}, \{\text{rm}\{\text{e}\}\}\}, \text{n}\} = \sqrt{2\{\text{k}_\text{-}\{\{\text{rm}\{\text{B}\}\}\}} \{\text{T}\}_\text{-}\{0\}/(\{\text{R}\}_\text{-}\{\{\text{rm}\{\text{p}\}\}\} \Delta t)} \cdot \text{G}_\text{-}\{n\} (\mu = 0, \sigma = 1), \text{\$}$$
  
 (22)

where, owing to the discrete time steps, the noise bandwidth is defined as  $\Delta f = 1/(2\Delta t)$ .  $G_n(\mu = 0, \sigma = 1)$  is a Gaussian distribution with mean  $\mu = 0$  and standard deviation  $\sigma = 1$  which is sampled every step, conserving the standard deviation of the noise while fulfilling the criterion that two subsequent values must be uncorrelated.

We implement laser cooling in the simulations by assigning a photon absorption probability to an ion in its electronic ground state at each time step. The laser, at wavelength  $\lambda$ , with wave vector  $k_L$  in the axial direction, is detuned from the centre of the transition frequency at  $f_0$  by a detuning  $\delta$  and we assume that the linewidth of the laser is negligible compared to the transition linewidth  $\Gamma$ . The discrete photon absorption probability depends on the velocity of the ion due to the Doppler effect and, in the low saturation limit, can be written as

$$\frac{\frac{1}{2} + \left( \frac{2\pi \Delta k L}{v_z} \right)^2}{\Gamma(\Delta t)} \Gamma(\Delta t) \quad (23)$$

where the saturation intensity is  $\langle I \rangle_{\rm s} = \frac{2\pi}{\hbar\Gamma\lambda^3}$  and  $I$  is a free parameter that is tuned to match the laser intensity in the experiment. Upon absorption of a photon the ion transitions to the excited electronic state and receives a momentum kick  $-\hbar k_L$ . The ion decays to the ground state via spontaneous emission with probability  $\Gamma\Delta t$  and receives a momentum kick in the axial direction of  $\hbar k \cos\theta$  where the angle  $\theta$  accounts for radial

momentum of the emitted photon. The ion can also decay via stimulated emission, in which case the momentum kick is  $+\hbar k_L$ .

Data preparation and analysis are performed in R<sup>46</sup>, while the intensive part of the calculation is performed using C++ via the Rcpp-package<sup>47</sup>. We use a fourth order symplectic integrator<sup>48</sup> to calculate the particle trajectories and the voltage across the RLC circuit to ensure that energy is, on average, conserved for numeric integration with more than  $10^{10}$  steps.

In simulations of the p-Be<sup>+</sup> resonator system, we apply the conditions of the experiments described in the main text to reproduce the frequency-domain behaviour in Extended Data Fig. 2, with Figs. 2a, b corresponding to the experimental results shown in the inset of Figs. 2d and 3a), respectively. The evolution of the oscillator energies with ten Be<sup>+</sup> ions,  $N_{\text{Be}} = 10$ , is shown in Extended Data Fig. 2c. Starting from  $t = 7.5$  s, a parametric drive is applied, resulting in a significant increase in the energy of the proton and the Be<sup>+</sup> ions from an initial temperature of 17 K. Similarly, Extended Data Fig. 2d shows the energy exchange between a single proton, 80 Be<sup>+</sup> ions, and the resonator all on resonance, where the cooling laser is applied from  $t = 10$  s, resulting in rapid cooling of the Be<sup>+</sup> ions and a temperature reduction of the proton.

## Axial frequency shifts

A particle in a Penning trap is subjected to shifts in the mode frequencies due to the inhomogeneity of the magnetic field and the anharmonic contributions to trapping potential<sup>26,32</sup>. The magnetic field in the trap centre can be written with the lowest order corrections as

$$\$ \$ \{ \bf{B} \} (z,r) = \{ B \}_0 \hat{z} + \{ B \}_2 ((z^2 - r^2/2) \hat{z} - r z \hat{r}), \$ \$ \\ (24)$$

where a quadratic gradient  $B_2$  shifts the axial frequency as a function of the radial energy by

$$\$ \$ \Delta \nu_z = \frac{1}{4\pi^2 m} \frac{\{ B \}_2}{\{ B \}_0} E_+ \\ (25)$$

We use this effect to demonstrate the energy exchange between the heated Be<sup>+</sup> ions and the proton in Fig. 2c. Here, the proton axial mode and modified cyclotron mode are sideband-coupled with a quadrupolar rf drive, so that after the sideband coupling,

the proton cyclotron energy freezes out at an energy  $E_+ = (\nu_+/v_z)E_z$ , where  $E_z$  is the axial energy while coupling the axial mode to the excited Be<sup>+</sup> ions.

timized to have a homogeneous magnetic field and is unsuited for energy measurements using equation (25) at low energy, with a temperature resolution of <0.1 mHz K<sup>-1</sup>. In the the sympathetic cooling measurements presented here, we instead used the trapping potential anharmonicity that we introduced in the proton trap to determine the temperature of the trapped particle. The trapping potential can be expanded in terms of  $C_n$  coefficients<sup>26,31,32</sup>, and the higher-order terms  $C_{2n}$ ,  $n \geq 2$  shift the trap frequencies  $\nu_i$  by

$$\Delta\nu_i \propto C_{2n} E_j^{n-1}, \quad (26)$$

where  $E_j$  is the energy of a trap mode. The coefficients  $C_n$  can be written in terms of a ‘tuning ratio’, TR, defined by the ratio of the voltage applied to the central ring electrode to the voltage applied to a correction electrode, as

$$C_n = E_j + D_n \text{rm}\{TR\}. \quad (27)$$

$D_n$  can be calculated from the trap geometry, and the axial frequency shift due to the leading energy-dependent trap anharmonicity  $C_4$  can be written as

$$\Delta\nu_z = \frac{1}{4\pi^2} \frac{m\nu_z}{\{C_2\}k_B T_z} \Delta \text{rm}\{T\} / \text{rm}\{R\}, \quad (28)$$

where  $\Delta\text{TR}$  is the offset in applied tuning ratio from the ideal tuning ratio at which  $C_4 = 0$ . Ultimately, the axial frequency shift as a function of TR and the axial energy  $E_z$  can be expressed as

$$\Delta\nu_z = \kappa_{D4} \{E_z\}, \quad (29)$$

where, for a proton stored in the proton trap,  $\kappa_{D4} = 45.4\Delta\text{TR}$  Hz K<sup>-1</sup>. This effect is used to determine the change of the proton axial temperature while the resonator is cooled with the laser-cooled Be<sup>+</sup> ions, and is the underlying method for the data shown in Figs. 3 and 4.

Temperature measurements using this method are limited by the determination of  $T_0$  to the  $\sim 1$  K level. We have previously performed higher precision temperature measurements using a dedicated, spatially distant trap with a ferromagnetic ring electrode that uses the shift of equation (25) to obtain a cyclotron energy resolution of up to 80 Hz K $^{-1}$  (refs. 15,49) and have developed a similar trap to reach 10 mK temperature resolution in future cooling measurements.

We also note that equation (28) and equation (29) cause the axial dip to spread out during an FFT averaging window and decrease the dip signal-to-noise ratio. This is reflected in the increased uncertainties at larger  $\Delta T$  that can be seen in Fig. 3b.

## Parasitic drive heating

The demonstration of remote energy exchange was performed by exciting a small cloud of Be $^+$  with an rf drive at twice the resonance frequency,  $2\nu_0$ , with results presented in Fig. 2d. To confirm that the proton is excited only by the resonantly coupled ions, we performed a series of background measurements. These control measurements show that the proton is excited only when resonant with the Be $^+$  ions and that the proton is unaffected by the excitation drive when the Be $^+$  ions are detuned (Extended Data Fig. 3).

Although the measurements presented in Fig. 2d are performed by interleaving the on-and off-resonant configurations, we additionally analysed the temperature of the proton in the presence of the drive. During these measurements, we transferred the axial energy to the modified cyclotron mode and measured the resulting energy dependent axial frequency shift, described in the main text. We measured an axial temperature of

$$\$ \$ \{T\}_{z,\{\text{rm}\{\text{on}\}\}} = 32(2)\backslash,\{\text{rm}\{K\}\} \$ \$ \\ (30)$$

in the presence of the drive, and an axial temperature of

$$\$ \$ \{T\}_{z,\{\text{rm}\{\text{off}\}\}} = 20(5)\backslash,\{\text{rm}\{K\}\} \$ \$ \\ (31)$$

in the absence of the drive, where the error comes from the fit uncertainty of the frequency scatter distribution. As a result, we constrain the possible increase in axial temperature due to the excitation drive to no more than a factor of two. From the spectra shown in Extended Data Fig. 3, we see that the signal-to-noise ratio of the proton dip is unaffected by the drive and conclude that an increase in axial temperature would come not from direct coupling of the excitation drive to the proton but via an

increase in the equivalent noise temperature of the resonator. We further note that the resonator temperature of 17.0 K given in the main text comes from the weighted mean of several temperature measurements performed with several methods. Importantly, residual heating due to the drive is far lower than the energy scatter shown in Fig. 2d; approximately 40,000 K  $k_B^{-1}$ .

## Data availability

The datasets generated and/or analysed during this study are available from the corresponding authors on request. [Source data](#) are provided with this paper.

## Code availability

The code used during this study is available from the corresponding authors on reasonable request.

## References

1. 1.

Smorra, C. et al. BASE—the baryon antibaryon symmetry experiment. *Eur. Phys. J. Spec. Top.* **224**, 3055–3108 (2015).

2. 2.

Sturm, S. et al. The ALPHATRAP experiment. *Eur. Phys. J. Spec. Top.* **227**, 1425–1491 (2019).

3. 3.

Huntemann, N., Sanner, C., Lipphardt, B., Tamm C.H.R. & Peik, E. Single-ion atomic clock with  $3 \times 10^{-18}$  systematic uncertainty. *Phys. Rev. Lett.* **116**, 063001 (2016).

4. 4.

Brewer, S. M. et al.  $^{27}\text{Al}^+$  quantum-logic clock with a systematic uncertainty below  $10^{-18}$ . *Phys. Rev. Lett.* **123**, 033201 (2019).

5. 5.

Kotler, S., Simmonds, R. W., Leibfried, D. & Wineland, D. J. Hybrid quantum systems with trapped charged particles. *Phys. Rev. A* **95**, 022327 (2017).

6. 6.

Bruzewicz, C. D., Chiaverini, J., McConnell, R. & Sage, J. M. Trapped-ion quantum computing: progress and challenges. *Appl. Phys. Rev.* **6**, 021314 (2019).

7. 7.

Brown, K. R. et al. Coupled quantized mechanical oscillators. *Nature* **471**, 196–199 (2011).

8. 8.

Harlander, M., Lechner, R., Brownnutt, M., Blatt, R. & Hänsel, W. Trapped-ion antennae for the transmission of quantum information. *Nature* **471**, 200–203 (2011).

9. 9.

Heinzen, D. J. & Wineland, D. J. Quantum-limited cooling and detection of radio-frequency oscillations by laser-cooled ions. *Phys. Rev. A*, **42**, 2977 (1990).

10. 10.

Daniilidis, N., Lee, T., Clark, R., Narayanan, S. & Haffner, H. Wiring up trapped ions to study aspects of quantum information. *J. Phys. B* **42**, 154012 (2009).

11. 11.

Bohman, M. et al. Sympathetic cooling of protons and antiprotons with a common end-cap Penning trap. *J. Mod. Opt.* **65**, 568–576 (2018).

12. 12.

Smorra, C. et al. Direct limits on the interaction of antiprotons with axionlike dark matter. *Nature* **575**, 310–314 (2019).

13. 13.

Devlin, J. A. et al. Constraints on the coupling between axionlike dark matter and photons using an antiproton superconducting tuned detection circuit in a cryogenic Penning trap. *Phys. Rev. Lett.* **126**, 041301 (2021).

14. 14.

Ulmer, S. et al. High-precision comparison of the antiproton-to-proton charge-to-mass ratio. *Nature* **524**, 196–199 (2015).

15. 15.

Schneider, G. et al. Double-trap measurement of the proton magnetic moment at 0.3 parts per billion precision. *Science* **358**, 1081–1084 (2017).

16. 16.

Smorra, C. et al. A parts-per-billion measurement of the antiproton magnetic moment. *Nature* **550**, 371–374 (2017).

17. 17.

Myers, E. G. CPT tests with the antihydrogen molecular ion. *Phys. Rev. A* **98**, 010101 (2018).

18. 18.

Yzombard, P., Hamamda, M., Gerber, S., Doser, M. & Comparat, D. Laser cooling of molecular anions. *Phys. Rev. Lett.* **114**, 213001 (2015).

19. 19.

Tang, R. et al. Candidate for laser cooling of a negative ion: high-resolution photoelectron imaging of  $\text{Th}^-$ . *Phys. Rev. Lett.* **123**, 203002 (2019).

20. 20.

Cerchiari, G., Yzombard, P. & Kellerbauer, A. Laser-assisted evaporative cooling of anions. *Phys. Rev. Lett.* **123**, 103201 (2019).

21. 21.

Wineland, D. J. et al. Experimental issues in coherent quantum-state manipulation of trapped atomic ions. *J. Res. Natl Inst. Stand. Technol.* **103**, 259–328 (1998).

22. 22.

Niemann, M., Paschke, A. -G., Dubielzig, T., Ulmer, S. & Ospelkaus, C. CPT test with (anti)proton magnetic moments based on quantum logic cooling and readout. In *Proc. Sixth Meeting on CPT and Lorentz Symmetry* (Ed. Kostelecký, V. A.) 41–44 (World Scientific, 2014);  
[https://www.worldscientific.com/doi/abs/10.1142/9789814566438\\_0011](https://www.worldscientific.com/doi/abs/10.1142/9789814566438_0011)

23. 23.

Egl, A. et al. Application of the continuous Stern-Gerlach effect for laser spectroscopy of the  $^{40}\text{Ar}^{13+}$  fine structure in a Penning trap. *Phys. Rev. Lett.* **123**, 123001 (2019).

24. 24.

Micke, P. et al. Coherent laser spectroscopy of highly charged ions using quantum logic. *Nature* **578**, 60–65 (2020).

25. 25.

Gutiérrez, M. J. et al. Dynamics of an unbalanced two-ion crystal in a Penning trap for application in optical mass spectrometry. *Phys. Rev. A* **100**, 063415 (2019).

26. 26.

Brown, L. S. & Gabrielse, G. Geonium theory: physics of a single electron or ion in a Penning trap. *Rev. Mod. Phys.* **58**, 233–311 (1986).

27. 27.

Nagahama, H. et al. Highly sensitive superconducting circuits at 700 kHz with tunable quality factors for image-current detection of single trapped antiprotons. *Rev. Sci. Instrum.* **87**, 113305 (2016).

28. 28.

Wineland, D. J. & Dehmelt, H. G. Principles of the stored ion calorimeter. *J. Appl. Phys.* **46**, 919–930 (1975).

29. 29.

Cornell, E. A., Weisskoff, R. M., Boyce, K. R. & Pritchard, D. E. Mode coupling in a Penning trap:  $\pi$  pulses and a classical avoided crossing. *Phys. Rev. A* **41**, 312–315 (1990).

30. 30.

Dehmelt, H. Continuous Stern-Gerlach effect: principle and idealized apparatus. *Proc. Natl Acad. Sci.* **83**, 2291–2294 (1986).

31. 31.

Gabrielse, G., Haarsma, L. & Rolston, S. L. Open-endcap Penning traps for high precision experiments. *Int. J. Mass Spectrom. Ion Processes* **88**, 319–332 (1989).

32. 32.

Ketter, J., Eronen, T., Höcker, M., Streubel, S. & Blaum, K. First-order perturbative calculation of the frequency-shifts caused by static cylindrically-symmetric electric and magnetic imperfections of a Penning trap. *Int. J. Mass spectrom.* **358**, 1–16 (2014).

33. 33.

Daniilidis, N. & Häffner, H. Quantum interfaces between atomic and solid-state systems. *Annu. Rev. Condens. Matter Phys.* **4**, 83–112 (2013).

34. 34.

Möller, S. A. *How to Couple Trapped Ions to Superconducting Resonators: Towards Hybrid Quantum Devices*. PhD thesis, Univ. California (2016).

35. 35.

Tu, B. et al. Tank-circuit assisted coupling method for sympathetic laser cooling. *Adv. Quantum Technol.* **4**, 2100029 (2021).

36. 36.

Powell, H. F., Segal, D. M. & Thompson, R. C. Axialization of laser cooled magnesium ions in a Penning trap. *Phys. Rev. Lett.* **89**, 093003 (2002).

37. 37.

Schneider, A. et al. A novel Penning-trap design for the high-precision measurement of the  ${}^3\text{He}^{2+}$  nuclear magnetic moment. *Ann. Phys.* **531**, 1800485 (2019).

38. 38.

Heiße, F. et al. High-precision measurement of the proton's atomic mass. *Phys. Rev. Lett.* **119**, 033001 (2017).

39. 39.

Smith, J. A., Hamzeloui, S., Fink, D. J. & Myers, E. G. Rotational energy as mass in H<sub>3</sub><sup>+</sup> and lower limits on the atomic masses of d and <sup>3</sup>He. *Phys. Rev. Lett.* **120**, 143002 (2018).

40. 40.

Rau, S. et al. Penning trap mass measurements of the deuteron and the HD+ molecular ion. *Nature* **585**, 43–47 (2020).

41. 41.

Cairncross, W. B. et al. Precision measurement of the electron's electric dipole moment using trapped molecular ions. *Phys. Rev. Lett.* **119**, 153001 (2017).

42. 42.

Roussy, T. S., Palken, D. A., Cairncross, W. B. et al. Experimental constraint on axion-like particle coupling over seven orders of magnitude in mass. Preprint at <https://arxiv.org/abs/2006.15787> (2020).

43. 43.

Feng, X., Charlton, M., Holzscheiter, M., Lewis, R.A. & Yamazaki, Y. Tank circuit model applied to particles in a Penning trap. *J. Appl. Phys.* **79**, 8–13 (1996).

44. 44.

Zeuthen, E. *Extending Cavity-Mechanical Cooling via a Hot LC Electrical Circuit Framework for Electromechanical Coupling Calculations*. MSc thesis, Niels Bohr Institute, Univ. Copenhagen (2012).

45. 45.

Bohman, M. *Sympathetic Cooling of a Proton with Resonant Image Current Coupling*. PhD thesis, Ruperto-Carola Univ. Heidelberg (2020).

46. 46.

R Core Team. *R: a Language and Environment for Statistical Computing*. <https://www.R-project.org/> (R Foundation for Statistical Computing, 2020).

47. 47.

Eddelbuettel, D. & François, R. Rcpp: seamless R and C++ integration. *J. Stat. Softw.* **40**, 1–18 (2011).

48. 48.

Yoshida, H. Construction of higher order symplectic integrators. *Phys. Lett. A* **150**, 262–268 (1990).

49. 49.

Borchert, M. J. et al. Measurement of ultralow heating rates of a single antiproton in a cryogenic Penning trap. *Phys. Rev. Lett.* **122**, 043201 (2019).

## Acknowledgements

This study comprises parts of the PhD thesis work of M.B. We acknowledge the contributions of G. L. Schneider and N. Schön to the design and construction of the Penning-trap system. We thank S. Sturm for helpful discussions regarding the cooling method presented here and acknowledge similar developments toward cooling highly charged ions in the ALPHATRAP collaboration. We acknowledge financial support from the RIKEN Chief Scientist Program, RIKEN Pioneering Project Funding, the RIKEN JRA Program, the Max-Planck Society, the Helmholtz-Gemeinschaft, the DFG through SFB 1227 ‘DQ-mat’, the European Union (Marie Skłodowska-Curie grant agreement number 721559), the European Research Council under the European Union’s Horizon 2020 research and innovation programme (Grant agreement numbers 832848-FunI and 852818-STEP) and the Max-Planck-RIKEN-PTB Center for Time, Constants and Fundamental Symmetries.

## Funding

Open access funding provided by Max Planck Society.

## Author information

### Affiliations

1. Max-Planck-Institut für Kernphysik, Heidelberg, Germany

M. Bohman, M. Wiesinger, C. Will, J. Harrington, A. Mooser & K. Blaum

2. RIKEN, Fundamental Symmetries Laboratory, Saitama, Japan

M. Bohman, C. Smorra, M. Wiesinger, M. J. Borchert, J. A. Devlin, S. Erlewein, M. Fleck, J. Harrington, B. Latacz, E. Wursten & S. Ulmer

3. Institut für Physik, Johannes Gutenberg-Universität, Mainz, Germany

V. Grunhofer, C. Smorra, S. Gavranovic, D. Popper & J. Walz

4. Institut für Quantenoptik, Leibniz Universität Hannover, Hannover, Germany

M. J. Borchert & C. Ospelkaus

5. Physikalisch-Technische Bundesanstalt, Braunschweig, Germany

M. J. Borchert & C. Ospelkaus

6. CERN, Geneva, Switzerland

J. A. Devlin, S. Erlewein & E. Wursten

7. Graduate School of Arts and Sciences, University of Tokyo, Tokyo, Japan

M. Fleck & Y. Matsuda

8. GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, Germany

W. Quint

9. Helmholtz-Institut Mainz, Mainz, Germany

J. Walz

## **Consortia**

### **BASE Collaboration**

- M. Bohman
- , V. Grunhofer
- , C. Smorra

- , M. Wiesinger
- , C. Will
- , M. J. Borchert
- , J. A. Devlin
- , S. Erlewein
- , M. Fleck
- , S. Gavranovic
- , J. Harrington
- , B. Latacz
- , A. Mooser
- , D. Popper
- , E. Wursten
- , K. Blaum
- , Y. Matsuda
- , C. Ospelkaus
- , W. Quint
- , J. Walz
- & S. Ulmer

## Contributions

M.B., A.M., S.U., J.W. and M.W. designed the experimental apparatus. M.B., A.M. and M.W. assembled the trap and laser systems. M.B., V.G., C.S. and M.W. contributed to the experiment run. M.B. and C.S. implemented the methods, and recorded and evaluated the experimental data. C.W. developed the simulation code and provided the simulation results. M.B., C.S., K.B. and S.U. prepared the manuscript, which was discussed and approved by all authors.

## Corresponding author

Correspondence to [M. Bohman](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

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

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

## Extended data figures and tables

### Extended Data Fig. 1 Calculated impedance of the equivalent circuit.

The normalised impedance  $\left(\{\mathrm{R}\}\{\mathrm{e}\}[Z(\omega)]/\{R\}_p\right)$  is plotted for different values of  $R_L$ , where no damping corresponds to  $R_L = 0$ , and no Be<sup>+</sup> ions to  $R_L \rightarrow \infty$ . Here,  $\gamma_R = 2\gamma_{\text{Be}} = 20\gamma_p$ , corresponding in experiment to about 30 Be<sup>+</sup> ions and a single proton.

[Source data](#)

### Extended Data Fig. 2 Simulation results.

**a)** A computed FFT spectrum is shown simulating the experimental conditions of Fig. 2d) in the main text. **b)** Representative time domain behaviour for these measurements is shown where the excitation drive is applied at time  $t = 10\text{s}$ . **c)** A computed FFT spectrum is shown simulating the experimental conditions of Fig. 3a) in the main text. **d)** Representative time domain behaviour for these measurements is shown where the cooling laser is applied at time  $t = 10\text{s}$ .

[Source data](#)

### Extended Data Fig. 3 Excitation drive background measurements.

**a)** An FFT spectrum while the excitation drive is off, the proton is on resonance with the resonator and the Be<sup>+</sup> ions are off resonance. **b)** An FFT spectrum while the excitation drive is on, the proton is off resonance with the resonator and the Be<sup>+</sup> ions are on resonance. **c)** An FFT spectrum while the excitation drive is on, the proton is on resonance with the resonator and the Be<sup>+</sup> ions are on resonance. **d)** An FFT spectrum while the excitation drive is on, the proton is on resonance with the resonator and the Be<sup>+</sup> ions are off resonance.

[Source data](#)

## Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[Reprints and Permissions](#)

## About this article

### Cite this article

Bohman, M., Grunhofer, V., Smorra, C. *et al.* Sympathetic cooling of a trapped proton mediated by an LC circuit. *Nature* **596**, 514–518 (2021).  
<https://doi.org/10.1038/s41586-021-03784-w>

- Received: 21 March 2021

- Accepted: 29 June 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03784-w>

## Single proton cooled by distant ions

- Manas Mukherjee

News & Views 25 Aug 2021

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03784-w>

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

- Article
- [Published: 25 August 2021](#)

# Single-crystal, large-area, fold-free monolayer graphene

- [Meihui Wang](#) [ORCID: orcid.org/0000-0001-5497-5709](#)<sup>1,2 na1</sup>,
- [Ming Huang](#) [ORCID: orcid.org/0000-0002-9188-4619](#)<sup>1 na1 nAff5</sup>,
- [Da Luo](#) [ORCID: orcid.org/0000-0002-9128-6782](#)<sup>1</sup>,
- [Yunqing Li](#)<sup>1,3</sup>,
- [Myeonggi Choe](#)<sup>1,3</sup>,
- [Won Kyung Seong](#)<sup>1</sup>,
- [Minhyeok Kim](#)<sup>1,2</sup>,
- [Sunghwan Jin](#) [ORCID: orcid.org/0000-0003-3750-8599](#)<sup>1,4</sup>,
- [Mengran Wang](#)<sup>1</sup>,
- [Shahana Chatterjee](#)<sup>1</sup>,
- [Youngwoo Kwon](#)<sup>1</sup>,
- [Zonghoon Lee](#) [ORCID: orcid.org/0000-0003-3246-4072](#)<sup>1,3</sup> &
- [Rodney S. Ruoff](#) [ORCID: orcid.org/0000-0002-6599-6764](#)<sup>1,2,3,4</sup>

[Nature](#) volume **596**, pages 519–524 (2021)

- 5766 Accesses
- 107 Altmetric
- [Metrics details](#)

## Subjects

- [Graphene](#)
- [Synthesis of graphene](#)

## Abstract

Chemical vapour deposition of carbon-containing precursors on metal substrates is currently the most promising route for the scalable synthesis of large-area, high-quality graphene films<sup>1</sup>. However, there are usually some imperfections present in the resulting films: grain boundaries, regions with additional layers (adlayers), and wrinkles or folds, all of which can degrade the performance of graphene in various applications<sup>2,3,4,5,6,7</sup>. There have been numerous studies on ways to eliminate grain boundaries<sup>8,9</sup> and adlayers<sup>10,11,12</sup>, but graphene folds have been less investigated. Here we explore the wrinkling/folding process for graphene films grown from an ethylene precursor on single-crystal Cu–Ni(111) foils. We identify a critical growth temperature (1,030 kelvin) above which folds will naturally form during the subsequent cooling process. Specifically, the compressive stress that builds up owing to thermal contraction during cooling is released by the abrupt onset of step bunching in the foil at about 1,030 kelvin, triggering the formation of graphene folds perpendicular to the step edge direction. By restricting the initial growth temperature to between 1,000 kelvin and 1,030 kelvin, we can produce large areas of single-crystal monolayer graphene films that are high-quality and fold-free. The resulting films show highly uniform transport properties: field-effect transistors prepared from these films exhibit average room-temperature carrier mobilities of around  $(7.0 \pm 1.0) \times 10^3$  centimetres squared per volt per second for both holes and electrons. The process is also scalable, permitting simultaneous growth of graphene of the same quality on multiple foils stacked in parallel. After electrochemical transfer of the graphene films from the foils, the foils themselves can be reused essentially indefinitely for further graphene growth.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

## Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

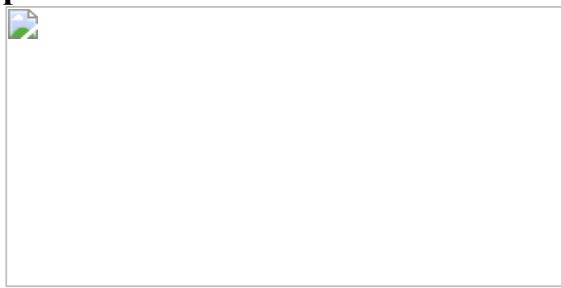
\$19.99/month

[\*\*Learn more\*\*](#)

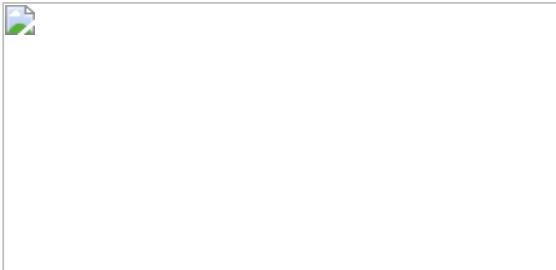
### Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

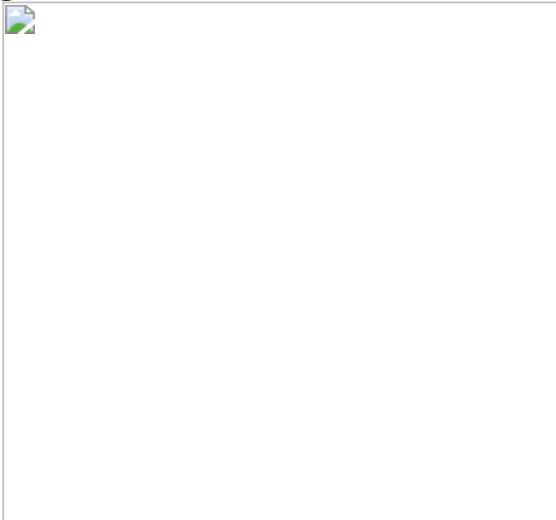
**Fig. 1: Investigation of the mechanism of graphene fold formation by cycling experiments.**



**Fig. 2: Fold evolution as a function of growth temperature.**



**Fig. 3: Characterizations of the fold-free graphene films.**



**Fig. 4: Transport properties of the fold-free graphene films.**

## Data availability

The data sets generated during the current study, and/or analysed during the current study, are available from the corresponding author upon reasonable request.

## References

1. 1.

Li, X. et al. Large-area synthesis of high-quality and uniform graphene films on copper foils. *Science* **324**, 1312–1314 (2009).

2. 2.

Zhu, W. et al. Structure and electronic transport in graphene wrinkles. *Nano Lett.* **12**, 3431–3436 (2012).

3. 3.

Wang, B. et al. Camphor-enabled transfer and mechanical testing of centimeter-scale ultrathin films. *Adv. Mater.* **30**, 1800888 (2018).

4. 4.

Huang, M. et al. Highly oriented monolayer graphene grown on a Cu/Ni(111) alloy foil. *ACS Nano* **12**, 6117–6127 (2018).

5. 5.

Chen, S. et al. Thermal conductivity measurements of suspended graphene with and without wrinkles by micro-Raman mapping. *Nanotechnology* **23**, 365701 (2012).

6. 6.

Li, X., Colombo, L. & Ruoff, R. S. Synthesis of graphene films on copper foils by chemical vapor deposition. *Adv. Mater.* **28**, 6247–6252 (2016).

7. 7.

Luo, D. et al. Adlayer-free large-area single crystal graphene grown on a Cu(111) foil. *Adv. Mater.* **31**, 1903615 (2019).

8. 8.

Wang, M., Luo, D., Wang, B. & Ruoff, R. S. Synthesis of large-area single-crystal graphene. *Trends Chem.* **3**, 15–33 (2021).

9. 9.

Zhang, J. et al. Controlled growth of single-crystal graphene films. *Adv. Mater.* **32**, 1903266 (2020).

10. 10.

Pang, J. et al. Oxidation as a means to remove surface contaminants on Cu foil prior to graphene growth by chemical vapor deposition. *J. Phys. Chem. C* **119**, 13363–13368 (2015).

11. 11.

Han, Z. et al. Homogeneous optical and electronic properties of graphene due to the suppression of multilayer patches during CVD on copper foils. *Adv. Funct. Mater.* **24**, 964–970 (2014).

12. 12.

Han, G. H. et al. Influence of copper morphology in forming nucleation seeds for graphene growth. *Nano Lett.* **11**, 4144–4148 (2011).

13. 13.

Zhang, Y. et al. Defect-like structures of graphene on copper foils for strain relief investigated by high-resolution scanning tunneling microscopy. *ACS Nano* **5**, 4014–4022 (2011).

14. 14.

Wang, W., Yang, S. & Wang, A. Observation of the unexpected morphology of graphene wrinkle on copper substrate. *Sci. Rep.* **7**, 8244 (2017).

15. 15.

Tripathi, M. et al. Structural defects modulate electronic and nanomechanical properties of 2D materials. *ACS Nano* **15**, 2520–2531 (2021).

16. 16.

Hattab, H. et al. Interplay of wrinkles, strain, and lattice parameter in graphene on iridium. *Nano Lett.* **12**, 678–682 (2012).

17. 17.

Lee, J.-H. et al. Wafer-scale growth of single-crystal monolayer graphene on reusable hydrogen-terminated germanium. *Science* **344**, 286–289 (2014).

18. 18.

Choi, J.-K. et al. Growth of wrinkle-free graphene on texture-controlled platinum films and thermal-assisted transfer of large-scale patterned graphene. *ACS Nano* **9**, 679–686 (2015).

19. 19.

Zhang, X. et al. Epitaxial growth of 6 in. single-crystalline graphene on a Cu/Ni (111) film at 750 °C via chemical vapor deposition. *Small* **15**, 1805395 (2019).

20. 20.

Deng, B. et al. Wrinkle-free single-crystal graphene wafer grown on strain-engineered substrates. *ACS Nano* **11**, 12337–12345 (2017).

21. 21.

Zhang, Y. et al. Batch production of uniform graphene films via controlling gas-phase dynamics in confined space. *Nanotechnology* **32**, 105603 (2021).

22. 22.

Polsen, E. S., McNerny, D. Q., Viswanath, B., Pattinson, S. W. & John Hart, A. High-speed roll-to-roll manufacturing of graphene using a concentric tube CVD reactor. *Sci. Rep.* **5**, 10257 (2015).

23. 23.

Wang, Y. et al. Electrochemical delamination of CVD-grown graphene film: toward the recyclable use of copper catalyst. *ACS Nano* **5**, 9927–9933 (2011).

24. 24.

Deng, B. et al. Anisotropic strain relaxation of graphene by corrugation on copper crystal surfaces. *Small* **14**, 1800725 (2018).

25. 25.

Kang, J. H. et al. Strain relaxation of graphene layers by Cu surface roughening. *Nano Lett.* **16**, 5993–5998 (2016).

26. 26.

Suh, I.-K., Ohta, H. & Waseda, Y. High-temperature thermal expansion of six metallic elements measured by dilatation method and X-ray diffraction. *J. Mater. Sci.* **23**, 757–760 (1988).

27. 27.

Ferrari, A. C. et al. Raman spectrum of graphene and graphene layers. *Phys. Rev. Lett.* **97**, 187401 (2006).

28. 28.

Lee, J. E., Ahn, G., Shim, J., Lee, Y. S. & Ryu, S. Optical separation of mechanical strain from charge doping in graphene. *Nat. Commun.* **3**, 1024 (2012).

29. 29.

Zabel, J. et al. Raman spectroscopy of graphene and bilayer under biaxial strain: bubbles and balloons. *Nano Lett.* **12**, 617–621 (2012).

30. 30.

Lee, G.-H. et al. High-strength chemical-vapor-deposited graphene and grain boundaries. *Science* **340**, 1073–1076 (2013).

31. 31.

Yi, D. et al. What drives metal-surface step bunching in graphene chemical vapor deposition? *Phys. Rev. Lett.* **120**, 246101 (2018).

32. 32.

Leong, W. S. et al. Paraffin-enabled graphene transfer. *Nat. Commun.* **10**, 867 (2019).

33. 33.

Huang, M. et al. Large-area single-crystal AB-bilayer and ABA-trilayer graphene grown on a Cu/Ni(111) foil. *Nat. Nanotechnol.* **15**, 289–295 (2020).

34. 34.

Kim, S. et al. Realization of a high mobility dual-gated graphene field-effect transistor with Al<sub>2</sub>O<sub>3</sub> dielectric. *Appl. Phys. Lett.* **94**, 062107 (2009).

35. 35.

Jin, S. et al. Colossal grain growth yields single-crystal metal foils by contact-free annealing. *Science* **362**, 1021 (2018).

## Acknowledgements

This work was supported by the Institute for Basic Science (IBS-R019-D1). We appreciate discussions with R. Huang (UT Austin) and K. Duck Park and H. Lee of UNIST.

## Author information

### Author notes

1. Ming Huang

Present address: School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore, Singapore

2. These authors contributed equally: Meihui Wang, Ming Huang

## Affiliations

1. Center for Multidimensional Carbon Materials (CMCM), Institute for Basic Science (IBS), Ulsan, Republic of Korea

Meihui Wang, Ming Huang, Da Luo, Yunqing Li, Myeonggi Choe, Won Kyung Seong, Minhyeok Kim, Sunghwan Jin, Mengran Wang, Shahana Chatterjee, Youngwoo Kwon, Zonghoon Lee & Rodney S. Ruoff

2. Department of Chemistry, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Meihui Wang, Minhyeok Kim & Rodney S. Ruoff

3. Department of Materials Science and Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Yunqing Li, Myeonggi Choe, Zonghoon Lee & Rodney S. Ruoff

4. School of Energy and Chemical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Sunghwan Jin & Rodney S. Ruoff

## Contributions

R.S.R., D.L. and Meihui Wang conceived the experiments. R.S.R. supervised the project. Meihui Wang did CVD growths. Meihui Wang and D.L. characterized films. S.J., Y.K. and M.K. prepared Cu(111) foils. M.H. and Mengran Wang prepared Cu–Ni(111) alloy foils. Y.L. made and tested the GFET devices. M.C., S.C. and Z.L. acquired and analysed TEM/SAED data. W.S. measured the LEED patterns and analysed the high-temperature XRD data (acquired by staff member M. J. Woo of the Korea Advanced Institute of Science and Technology (KAIST)). W.S. designed, built and tested the 6-inch CVD system which is now used by Meihui Wang and others. Meihui Wang wrote a draft manuscript and R.S.R., Meihui Wang and D.L. revised it. All co-authors commented on the manuscript prior to its submission.

## Corresponding authors

Correspondence to [Da Luo](#) or [Rodney S. Ruoff](#).

## Ethics declarations

## Competing interests

The Institute for Basic Science has filed a patent application (KR 10-2021-0095514) that lists Meihui Wang, D.L. and R.S.R. as inventors. Other than this, the authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Cedric Huyghebaert, Jeehwan Kim and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

# Supplementary information

## Supplementary Information

This file contains supplementary text, supplementary notes, supplementary figures s1 – s29, supplementary equations, supplementary tables s1 – s5 and supplementary references.

## Peer Review File

## Rights and permissions

### Reprints and Permissions

## About this article

### Cite this article

Wang, M., Huang, M., Luo, D. *et al.* Single-crystal, large-area, fold-free monolayer graphene. *Nature* **596**, 519–524 (2021). <https://doi.org/10.1038/s41586-021-03753-3>

- Received: 18 January 2021
- Accepted: 22 June 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03753-3>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03753-3>

- Article
- [Published: 25 August 2021](#)

# Rechargeable Na/Cl<sub>2</sub> and Li/Cl<sub>2</sub> batteries

- [Guanzhou Zhu](#)<sup>1 na1</sup>,
- [Xin Tian](#)<sup>1 na1</sup>,
- [Hung-Chun Tai](#) ORCID: [orcid.org/0000-0001-9963-3518](https://orcid.org/0000-0001-9963-3518)<sup>2</sup>,
- [Yuan-Yao Li](#)<sup>2</sup>,
- [Jiachen Li](#)<sup>1</sup>,
- [Hao Sun](#)<sup>1</sup>,
- [Peng Liang](#)<sup>1</sup>,
- [Michael Angell](#)<sup>1</sup>,
- [Cheng-Liang Huang](#)<sup>2</sup>,
- [Ching-Shun Ku](#)<sup>3</sup>,
- [Wei-Hsuan Hung](#)<sup>4</sup>,
- [Shi-Kai Jiang](#)<sup>5</sup>,
- [Yongtao Meng](#)<sup>1</sup>,
- [Hui Chen](#)<sup>6</sup>,
- [Meng-Chang Lin](#)<sup>6</sup>,
- [Bing-Joe Hwang](#) ORCID: [orcid.org/0000-0002-3873-2149](https://orcid.org/0000-0002-3873-2149)<sup>5</sup> &
- [Hongjie Dai](#) ORCID: [orcid.org/0000-0002-4906-4502](https://orcid.org/0000-0002-4906-4502)<sup>1</sup>

[Nature](#) volume 596, pages 525–530 (2021)

- 4661 Accesses
- 127 Altmetric
- [Metrics details](#)

## Subjects

- [Batteries](#)

## Abstract

Lithium-ion batteries (LIBs) are widely used in applications ranging from electric vehicles to wearable devices. Before the invention of secondary LIBs, the primary lithium-thionyl chloride (Li-SOCl<sub>2</sub>) battery was developed in the 1970s using SOCl<sub>2</sub> as the catholyte, lithium metal as the anode and amorphous carbon as the cathode<sup>1,2,3,4,5,6,7</sup>. This battery discharges by lithium oxidation and catholyte reduction to sulfur, sulfur dioxide and lithium chloride, is well known for its high energy density and is widely used in real-world applications; however, it has not been made rechargeable since its invention<sup>8,9,10,11,12,13</sup>. Here we show that with a highly microporous carbon positive electrode, a starting electrolyte composed of aluminium chloride in SOCl<sub>2</sub> with fluoride-based additives, and either sodium or lithium as the negative electrode, we can produce a rechargeable Na/Cl<sub>2</sub> or Li/Cl<sub>2</sub> battery operating via redox between mainly Cl<sub>2</sub>/Cl<sup>-</sup> in the micropores of carbon and Na/Na<sup>+</sup> or Li/Li<sup>+</sup> redox on the sodium or lithium metal. The reversible Cl<sub>2</sub>/NaCl or Cl<sub>2</sub>/LiCl redox in the microporous carbon affords rechargeability at the positive electrode side and the thin alkali-fluoride-doped alkali-chloride solid electrolyte interface stabilizes the negative electrode, both are critical to secondary alkali-metal/Cl<sub>2</sub> batteries.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

## Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

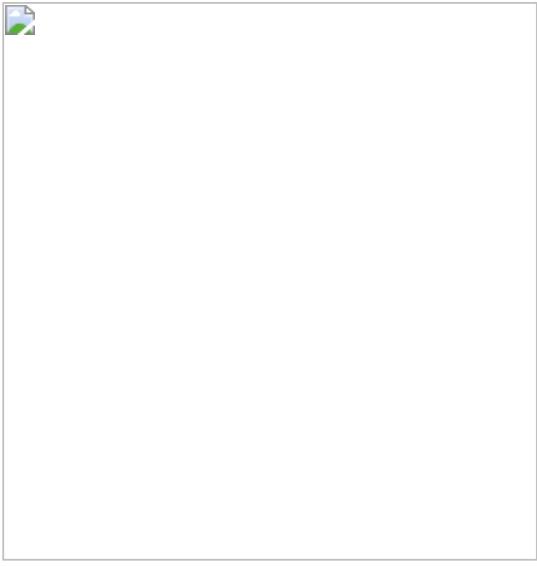
\$19.99/month

[\*\*Learn more\*\*](#)

### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

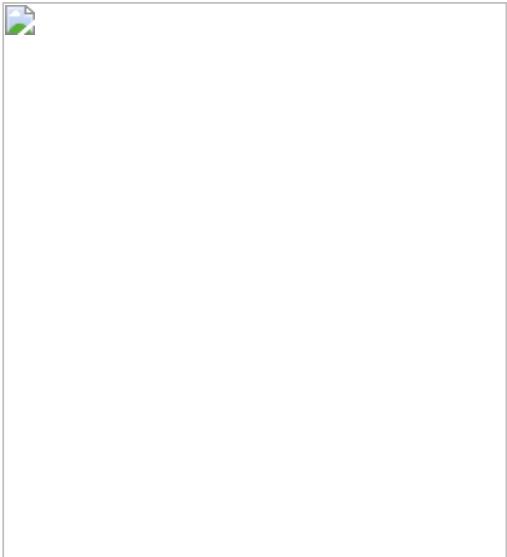
**Fig. 1: A high-capacity Na/Cl<sub>2</sub> battery through the first discharge.**



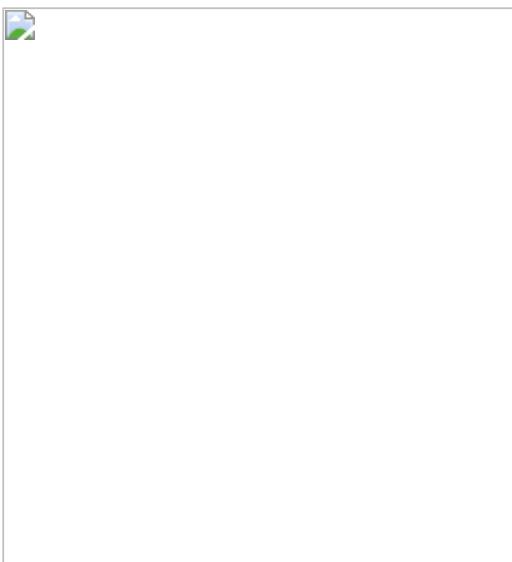
**Fig. 2: Rechargeable Na/Cl<sub>2</sub> battery at different battery states through cycling.**



**Fig. 3: Cycling performance of Na/Cl<sub>2</sub> battery at capacities up to the capacity of the first lower discharge plateau (1,860 mAh g<sup>-1</sup>).**



**Fig. 4: Importance of stable SEI on sodium anode and aCNS cathode for Na/Cl<sub>2</sub> and Li/Cl<sub>2</sub> batteries.**



## Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. Source data are provided with this paper.

## References

1. 1.

Venkatasetty, H. V. & Saathoff, D. J. Properties of LiAlCl<sub>4</sub>–SOCl<sub>2</sub> solutions for Li/SOCl<sub>2</sub> battery. *J. Electrochem. Soc.* **128**, 773–777 (1981).

2. 2.

Tsaur, K. C. & Pollard, R. Mathematical modeling of the lithium, thionyl chloride static cell: II. Acid electrolyte. *J. Electrochem. Soc.* **131**, 984–990 (1984).

3. 3.

Istone, W. K. & Brodd, R. J. The mechanisms of thionyl chloride reduction at solid electrodes. *J. Electrochem. Soc.* **131**, 2467–2470 (1984).

4. 4.

Gangadharan, R., Namboodiri, P. N. N., Prasad, K. V. & Viswanathan, R. The lithium-thionyl chloride battery—a review. *J. Power Sources* **4**, 1–9 (1979).

5. 5.

Madou, M. J. & Szpak, S. Investigation of SOCl<sub>2</sub> reduction by cyclic voltammetry and ac impedance measurements. *J. Electrochem. Soc.* **131**, 2471–2475 (1984).

6. 6.

Bedfer, Y., Corset, J., Dhamelincourt, M. C., Wallart, F. & Barbier, P. Raman spectroscopic studies of the structure of electrolytes used in the Li/SOCl<sub>2</sub> battery. *J. Power Sources* **9**, 267–272 (1983).

7. 7.

Carter, B. J. et al. Mechanistic studies related to the safety of Li/SOCl<sub>2</sub> cells. *J. Electrochem. Soc.* 525–528 (1985).

8. 8.

Marinčić, N. Materials balance in primary batteries. II. Lithium inorganic batteries at high discharge rates. *J. Appl. Electrochem.* **6**, 51–58 (1976).

9. 9.

Wang, D. et al. The effects of pore size on electrical performance in lithium-thionyl chloride batteries. *Front. Mater.* **6**, 245 (2019).

10. 10.

Klinedinst, K. A. & Domeniconi, M. J. High rate discharge characteristics of Li/SOCl<sub>2</sub> cells. *J. Electrochem. Soc.* **127**, 539–544 (1980).

11. 11.

Abraham, K. M. & Mank, R. M. Some chemistry in the Li/SOCl<sub>2</sub> cell. *J. Electrochem. Soc.* **127**, 2091–2096 (1980).

12. 12.

Spotnitz, R. M., Yeduvaka, G. S., Nagasubramanian, G. & Jungst, R. Modeling self-discharge of Li/SOCl<sub>2</sub> cells. *J. Power Sources* **163**, 578–583 (2006).

13. 13.

Morrison, M. M. & Marincic, N. Studies in lithium oxyhalide cells for downhole instrumentation use of lithium tetrachlorogallate electrolyte in Li/SOCl<sub>2</sub> cells. *J. Power Sources* **45**, 343–352 (1993).

14. 14.

Sun, H. et al. A safe and non-flammable sodium metal battery based on an ionic liquid electrolyte. *Nature Commun.* **10**, 3302 (2019).

15. 15.

Sun, H. et al. High-safety and high-energy-density lithium metal batteries in a novel ionic-liquid electrolyte. *Adv. Mater.* **32**, 2001741 (2020).

16. 16.

Zhu, G. et al. Rechargeable aluminum batteries: effects of cations in ionic liquid electrolytes. *RSC Adv.* **9**, 11322–11330 (2019).

17. 17.

Lin, M.-C. et al. An ultrafast rechargeable aluminium-ion battery. *Nature* **520**, 324–328 (2015).

18. 18.

Angell, M. et al. High Coulombic efficiency aluminum-ion battery using an AlCl<sub>3</sub>-urea ionic liquid analog electrolyte. *Proc. Natl Acad. Sci. USA* **114**, 834–839 (2017).

19. 19.

Pan, C.-J. et al. An operando X-ray diffraction study of chloroaluminate anion-graphite intercalation in aluminum batteries. *Proc. Natl Acad. Sci. USA* **115**, 5670–5675 (2018).

20. 20.

Di Lecce, D., Carbone, L., Gancitano, V. & Hassoun, J. Rechargeable lithium battery using non-flammable electrolyte based on tetraethylene glycol dimethyl ether and olivine cathodes. *J. Power Sources* **334**, 146–153 (2016).

21. 21.

Agostini, M., Xiong, S., Matic, A. & Hassoun, J. Polysulfide-containing glyme-based electrolytes for lithium sulfur battery. *Chem. Mater.* **27**, 4604–4611 (2015).

22. 22.

Cai, K., Song, M.-K., Cairns, E. J. & Zhang, Y. Nanostructured Li<sub>2</sub>S–C composites as cathode material for high-energy lithium/sulfur batteries. *Nano Lett.* **12**, 6474–6479 (2012).

23. 23.

Angell, M., Zhu, G., Lin, M.-C., Rong, Y. & Dai, H. Ionic liquid analogs of AlCl<sub>3</sub> with urea derivatives as electrolytes for aluminum batteries. *Adv. Funct. Mater.* **30**, 1901928 (2020).

24. 24.

Barpanda, P., Oyama, G., Nishimura, S.-i., Chung, S.-C. & Yamada, A. A 3.8-V earth-abundant sodium battery electrode. *Nat. Commun.* **5**, 4358 (2014).

25. 25.

Zhu, C., Kopold, P., van Aken, P. A., Maier, J. & Yu, Y. High power–high energy sodium battery based on threefold interpenetrating network. *Adv. Mater.* **28**, 2409–2416 (2016).

26. 26.

Liu, J. et al. Extension of the Stöber method to the preparation of monodisperse resorcinol–formaldehyde resin polymer and carbon spheres. *Angew. Chem. Int. Ed.* **50**, 5947–5951 (2011).

27. 27.

Tsai, C.-Y., Tai, H.-C., Su, C.-A., Chiang, L.-M. & Li, Y.-Y. Activated microporous carbon nanospheres for use in supercapacitors. *ACS Appl. Nano Mater.* **3**, 10380–10388 (2020).

28. 28.

Gross, S. & Society, E. *Proc. Symposium on Battery Design and Optimization* (Battery Division, Electrochemical Society, 1979).

29. 29.

Abraham, K. M., Mank, R. M. & Holleck, G. L. *Investigations of the Safety of Li/SOCl<sub>2</sub> Batteries* (1979).

30. 30.

Dey, A. N. Lithium anode film and organic and inorganic electrolyte batteries. *Thin Solid Films* **43**, 131–171 (1977).

31. 31.

Mogensen, M. B. & Hennesø, E. Properties and structure of the LiCl-films on lithium anodes in liquid cathodes. *Acta Chim. Slov.* **63**, 519–534 (2016).

32. 32.

Alvarado, J. et al. Bisalt ether electrolytes: a pathway towards lithium metal batteries with Ni-rich cathodes. *Energy Environ. Sci.* **12**, 780–794 (2019).

33. 33.

Eshetu, G. G. et al. Ultrahigh performance all solid-state lithium sulfur batteries: salt anion's chemistry-induced anomalous synergistic effect. *J. Am. Chem. Soc.* **140**, 9921–9933 (2018).

34. 34.

Evans, T. I., Nguyen, T. V. & White, R. E. A mathematical model of a lithium/thionyl chloride primary cell. *J. Electrochem. Soc.* **136**, 328–339 (1989).

35. 35.

Gilman, S. The reduction of sulfuryl chloride at teflon-bonded carbon cathodes. *J. Electrochem. Soc.* **127**, 1427–1433 (1980).

36. 36.

Xu, X. et al. A room-temperature sodium–sulfur battery with high capacity and stable cycling performance. *Nat. Commun.* **9**, 3870 (2018).

37. 37.

Lee, M. et al. High-performance sodium–organic battery by realizing four-sodium storage in disodium rhodizonate. *Nat. Energy* **2**, 861–868 (2017).

38. 38.

Hu, L. et al. Dually decorated  $\text{Na}_3\text{V}_2(\text{PO}_4)_2\text{F}_3$  by carbon and 3D graphene as cathode material for sodium-ion batteries with high energy and power densities. *ChemElectroChem* **7**, 3975–3983 (2020).

39. 39.

Hwang, J.-Y., Kim, J., Yu, T.-Y. & Sun, Y.-K. A new P2-type layered oxide cathode with extremely high energy density for sodium-ion batteries. *Adv. Energy Mater.* **9**, 1803346 (2019).

40. 40.

Liu, X., Ma, W., Lei, X., Zhang, S. & Ding, Y. Rechargeable Na– $\text{SO}_2$  battery with ethylenediamine additive in ether-based electrolyte. *Adv. Funct. Mater.* **30**, 2002120 (2020).

41. 41.

Jeong, G. et al. A room-temperature sodium rechargeable battery using an SO<sub>2</sub>-based nonflammable inorganic liquid catholyte. *Sci. Rep.* **5**, 12827–12827 (2015).

## Acknowledgements

This work was supported by a Stanford Bits and Watts Fellowship. Part of this work was performed at the Stanford Nano Shared Facilities (SNSF) supported by the National Science Foundation under award ECCS-2026822.

## Author information

### Author notes

1. These authors contributed equally: Guanzhou Zhu, Xin Tian

### Affiliations

1. Department of Chemistry and Bio-X, Stanford University, Stanford, CA, USA  
Guanzhou Zhu, Xin Tian, Jiachen Li, Hao Sun, Peng Liang, Michael Angell, Yongtao Meng & Hongjie Dai
2. Department of Chemical Engineering, National Chung Cheng University, Chia-Yi, Taiwan  
Hung-Chun Tai, Yuan-Yao Li & Cheng-Liang Huang
3. National Synchrotron Radiation Research Center, Hsinchu, Taiwan  
Ching-Shun Ku
4. Institute of Materials Science and Engineering, National Central University, Taoyuan City, Taiwan  
Wei-Hsuan Hung
5. Department of Chemical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

Shi-Kai Jiang & Bing-Joe Hwang

6. College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, People's Republic of China

Hui Chen & Meng-Chang Lin

## Contributions

G.Z. and H.D. conceived the main idea of the project. G.Z. and X.T. performed the experiments and contributed equally to this work. G.Z. performed the mass spectrometry measurement of species in the battery. H.-C.T., C.-L.H. and Y.-Y.L. prepared the aCNS raw material. H.-C.T. performed the characterizations of aCNS (SEM, TEM, XRD and so on). J.L., C.-S.K., W.-H.H., S.-K.J. and B.-J.H. performed characterizations of electrodes in battery. G.Z. and H.S. performed the LED demo of the Na/Cl<sub>2</sub> battery. G.Z., P.L. and M.A. performed X-ray photoelectron spectroscopy measurements. H.C. and M.-C.L. prepared the chemicals used as electrolyte additives and thinner glass fibre separators (60 µm). G.Z. and H.D. prepared the manuscript. All authors participated in experimental data/results analysis and discussion.

## Corresponding author

Correspondence to [Hongjie Dai](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks the K. M. Abraham, Jiulin Wang, Jianshe Zhao and the anonymous reviewer(s) for their contribution to the peer review of this work.

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

## Extended data figures and tables

### Extended Data Fig. 1 SEM images of aCNS at different battery stages.

**a**, SEM images of aCNS through the first discharge (from 950 mAh g<sup>-1</sup>, to 2,100 mAh g<sup>-1</sup>, and then full discharge) of the Na/Cl<sub>2</sub> battery and atomic percentage of C, Na and Cl at these stages measured by SEM/energy dispersive X-ray spectroscopy (EDS) mapping (right bar graph). As discharge continued, more and more NaCl was formed on the aCNS and the discharge stopped when NaCl passivated the aCNS. Some of the NaCl formed was very large in size (tens of micrometres). **b**, SEM images of aCNS when the Na/Cl<sub>2</sub> battery was re-charged to different capacities (375 mAh g<sup>-1</sup>, 600 mAh g<sup>-1</sup>, 900 mAh g<sup>-1</sup>) and the atomic percentage of C, Na and Cl at these stages measured by SEM/EDS mapping (right bar graph). As charging increased, more and more NaCl was removed from the aCNS, exposing the nanospheres underlying the NaCl coating. The active sites of the battery (the sites at which oxidation reactions happened) were in the gaps in the NaCl microcrystal coating that remained intact during battery operations. **c**, SEM images of aCNS when the Na/Cl<sub>2</sub> battery was charged to 900 mAh g<sup>-1</sup> then discharged to different capacities (375 mAh g<sup>-1</sup>, 600 mAh g<sup>-1</sup>, 900 mAh g<sup>-1</sup>) and atomic percentage of C, Na and Cl at these stages measured by SEM/EDS mapping (right bar graph). As discharge increased, more and more NaCl formed on the aCNS. When the battery was fully discharged, all the nanospheres were covered and passivated by the NaCl. To take these SEM images, batteries stopping at the designated states were opened inside an argon-filled glovebox and the electrodes were first dried under vacuum, then taken out of the glovebox and transferred into an SEM instrument for the measurements. See [Methods](#) for details.

### Extended Data Fig. 2 EIS of Na/Cl<sub>2</sub> battery with acidic 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% NaFSI + 2 wt% NaTFSI as the electrolyte through its first discharge and re-charging and first discharge curve of Na/Cl<sub>2</sub> battery using neutral 4 M AlCl<sub>3</sub> + 4 M NaCl in SOCl<sub>2</sub> as the electrolyte.

**a**, Impedance measurements at six points along the curve of first discharge of the battery when acidic 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% NaFSI + 2 wt% NaTFSI was used as the electrolyte. **b**, Charging curve of the Na/Cl<sub>2</sub> battery when the charging capacity was 500 mAh g<sup>-1</sup>. Each spike along the curve was a point at which battery charging was stopped for EIS measurements and then allowed to continue to charge.

**c**, Impedance measurements of the Na/Cl<sub>2</sub> battery at different charging capacities tracing the charging curve in **b**. As charging started, the impedance of the battery rapidly decreased due to removal of NaCl in the coating layer on the positive electrode. **d**, First discharge curve of Na/Cl<sub>2</sub> battery when neutral 4 M AlCl<sub>3</sub> + 4 M NaCl in SOCl<sub>2</sub> was used as the electrolyte. Only one discharge plateau was observed in neutral electrolyte case.

### Extended Data Fig. 3 Cycling performance of Na/Cl<sub>2</sub> battery at different capacities.

**a**, Cycling performance of a Na/Cl<sub>2</sub> battery at 500 mAh g<sup>-1</sup> (150 mA g<sup>-1</sup>). The battery was kept at open circuit in a discharged state for two weeks. We found that simply aging the battery in the discharged state for days could improve the battery's cycle life, probably due to the slower formation of a more uniform SEI layer on the electrode. The loading of aCNS was about 4.5 mg cm<sup>-2</sup>. **b**, Na/Cl<sub>2</sub> battery cycling at 1,200 mAh g<sup>-1</sup>. The electrolyte was 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 1 wt% NaFSI + 1 wt% NaTFSI. **c**, Na/Cl<sub>2</sub> battery cycling at 1,200 mAh g<sup>-1</sup>. The electrolyte was 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% NaFSI + 2 wt% NaTFSI. Both of the batteries in **b**, **c** were first cycling at 500 mAh g<sup>-1</sup> (150 mA g<sup>-1</sup>) for 15 cycles and the cycling capacity was gradually increased to 1,200 mAh g<sup>-1</sup> with 150 mA g<sup>-1</sup> and 100 mA g<sup>-1</sup> currents. The loading of both batteries was about 2.6 mg cm<sup>-2</sup>. **d**, Cycling performance of Na/Cl<sub>2</sub> battery as the charging current increased from 0.3 C (150 mA g<sup>-1</sup>) up to 3.9 C (1,950 mA/g<sup>-1</sup>) with 0.3 C (150 mA g<sup>-1</sup>) increased for every five cycles. The discharge current was kept at 0.3 C (150 mA g<sup>-1</sup>). The loading of aCNS was about 3 mg cm<sup>2</sup>. **e**, Cycling performance of Na/Cl<sub>2</sub> battery at 1,200 mAh g<sup>-1</sup> with charging current increased to 0.5 C (600 mA g<sup>-1</sup>) and discharging current kept at 0.08 C (100 mA g<sup>-1</sup>). Cycles 1–3: 0.0625 C (75 mA g<sup>-1</sup>), cycles 4 and 5: 0.08 C (100 mA g<sup>-1</sup>) for battery stabilization. The loading of the battery was about 3 mg cm<sup>-2</sup>. **f**, Typical charge–discharge curves of Na/Cl<sub>2</sub> battery at 1,200 mAh g<sup>-1</sup>. Black curve: 0.5 C (600 mA g<sup>-1</sup>) charging, 0.08 C (100 mA g<sup>-1</sup>) discharging. Red curve: 0.08 C (100 mA g<sup>-1</sup>) charging and discharging. Only a slight increase in overpotential (about 182 mV at 0.08 C versus about 298 mV at 0.5 C) was observed. The loading of the battery was about 3 mg cm<sup>-2</sup>.

### Extended Data Fig. 4 SEM images of aCNS after charging to 1,860 mAh g<sup>-1</sup>.

Left image: the nanospheres in aCNS were readily observed as NaCl depositing on the surface of aCNS were oxidized. Middle and right images: NaCl microcrystals that were either loosely deposited on top of the nanospheres clusters (not inside the nanospheres, middle image) or deposited in the gaps between the aCNS clusters (right image) were not oxidizable, and could not contribute to the battery's rechargeable capacity. Middle and right images have different magnifications.

**Extended Data Fig. 5 Na/Cl<sub>2</sub> battery performances when 2 wt% FEC and 2 wt% NaPF<sub>6</sub> were used as the electrolyte additives and XPS of Na metal immersing in electrolytes with different additives (2 wt% NaFSI + 2 wt% NaTFSI, 2 wt% NaPF<sub>6</sub>, and 2 wt% FEC) and after battery cycling.**

**a**, Na/Cl<sub>2</sub> battery cycling performance at 500 mAh g<sup>-1</sup>, 150 mA g<sup>-1</sup> when 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% FEC was used as the electrolyte. The battery behaved poorly and died after cycle 9. **b**, Na/Cl<sub>2</sub> battery cycling performance at 1,200 mAh g<sup>-1</sup>, 100 mA g<sup>-1</sup> when 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% NaPF<sub>6</sub> was used as the electrolyte. The battery showed worse cycling performance than when 2 wt% NaFSI + 2 wt% NaTFSI was used as the electrolyte additive. **c**, Atomic percentage of different elements, calculated from XPS survey spectrum, on the Na metal after immersing in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with different additives (2 wt% NaFSI + 2 wt% NaTFSI, 2 wt% NaPF<sub>6</sub> and 2 wt% FEC). **d**, Cl 2p spectrum of Na metal after immersing in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with different additives (2 wt% NaPF<sub>6</sub> and 2 wt% FEC). **e**, F 1s spectrum of Na metal after immersing in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with different additives (2 wt% NaPF<sub>6</sub> and 2 wt% FEC). **f**, S 2p spectrum of Na metal after immersing in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with different additives (2 wt% NaPF<sub>6</sub> and 2 wt% FEC). **g**, Atomic percentage of different elements, calculated from XPS survey spectrum, on the Na electrode after cycling in batteries using 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with different additives (2 wt% NaFSI + 2 wt% NaTFSI, 2 wt% NaPF<sub>6</sub> and 2 wt% FEC) as the electrolyte. **h**, F 1s spectrum of Na electrode after cycling in batteries using 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with different additives (2 wt% NaPF<sub>6</sub> and 2 wt% FEC) as the electrolyte. The batteries using 2 wt% NaFSI + 2 wt% NaTFSI and 2 wt% NaPF<sub>6</sub> as the electrolyte additives in **g**, **h** were stopped at cycle 21. The battery using 2 wt% FEC as the electrolyte additive was stopped at cycle 9 when the battery died.

**Extended Data Fig. 6 Characterizations of Na anode immersed and cycled in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with and without 2 wt% NaFSI/NaTFSI,**

**and charge–discharge curves of the normal battery versus decayed battery.**

**a**, Atomic percentage of different elements on the Na metal when immersed in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with and without 2 wt% NaFSI/NaTFSI as additives. **b**, F 1s spectrum of Na immersed in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with/without additives. **c**, S 2p spectrum of Na immersed in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with/without additives. **d**, Cl 2p spectrum of Na immersed in 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with/without additives. **e**, Atomic percentage of different elements on the Na metal after cycling for 21 cycles in Na/Cl<sub>2</sub> battery when 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with and without 2 wt% NaFSI/NaTFSI as additives were used as the electrolyte. **f**, F 1s spectrum of Na cycled in Na/Cl<sub>2</sub> battery when 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with/without additives were used as the electrolyte. **g**, SEM images of Na anode from actual Na/Cl<sub>2</sub> battery in charged state (top images) and when lost cycling capability (bottom images). Note that in the case of battery without fluoride additive, the Na anode surface was coated by more densely packed NaCl particles, eventually leading to the loss of re-chargeability. **h**, Charge–discharge curves of the battery at normal state and after the battery started to decay.

**Extended Data Fig. 7 SEM images of Na electrodes after cycling in batteries using 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> with different additives (2 wt% NaFSI + 2 wt% NaTFSI, 2 wt% NaPF<sub>6</sub>, and 2 wt% FEC) as the electrolytes.**

Top row: SEM images of Na electrode after cycling in battery using 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% NaFSI + 2 wt% NaTFSI as the electrolyte. The SEI layer contained loosely packed, square-shaped NaCl crystals and abundant voids still present in the SEI (indicated by circles). Middle row: SEM images of Na electrode after cycling in battery using 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% NaPF<sub>6</sub> as the electrolyte. The SEI layer contained closely packed, square-shaped NaCl crystals that were grown on top of a uniform layer of NaCl crystals. Such morphology made ions penetrations much less efficient. Bottom row: SEM images of Na electrode after cycling in battery using 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% FEC as the electrolyte. The SEI layer was made of very large NaCl crystals (tens of micrometres in size) packed together. Such morphology made ions penetrations only possible via the small cracks between these crystals and the least efficient. The batteries using 2 wt% NaFSI + 2 wt% NaTFSI and 2 wt% NaPF<sub>6</sub> as the electrolyte additives were both stopped at cycle 21. The battery using 2 wt% FEC as the electrolyte additive was stopped at cycle 9 when the battery died.

**Extended Data Fig. 8 Na/Cl<sub>2</sub> battery cycling performance using less electrolyte (4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% NaFSI + 2 wt% NaTFSI) and thinner separators down to 60 µm.**

**a**, Na/Cl<sub>2</sub> battery cycling performance at 500 mAh g<sup>-1</sup> using 100 µl electrolyte with one layer of QR-100 separator. The loading of the battery was about 5 mg cm<sup>-2</sup>. **b**, Na/Cl<sub>2</sub> battery cycling performance at 500 mAh g<sup>-1</sup> using 75 µl electrolyte with one layer of QR-100 separator. The loading of the battery was about 5 mg cm<sup>-2</sup>. **c**, Na/Cl<sub>2</sub> battery cycling performance at 500 mAh g<sup>-1</sup> using 50 µl electrolyte with one layer of 60-µm glass fibre separator. The loading of the battery was about 5 mg cm<sup>-2</sup>. **d**, Charge–discharge curve of Na/Cl<sub>2</sub> battery at 500 mAh g<sup>-1</sup> using 50 µl electrolyte. **e**, Na/Cl<sub>2</sub> battery cycling performance at 1,200 mAh g<sup>-1</sup> using 100 µl electrolyte with one layer of QR-100 separator. The loading of the battery was about 3.6 mg cm<sup>-2</sup>. **f**, Charge–discharge curve of Na/Cl<sub>2</sub> battery at 1,200 mAh g<sup>-1</sup> using 100 µl electrolyte.

**Extended Data Fig. 9 Li/Cl<sub>2</sub> battery cycling at 500 mAh g<sup>-1</sup> with 4 M AlCl<sub>3</sub> in SOCl<sub>2</sub> + 2 wt% LiFSI + 2 wt% LiTFSI as the electrolyte.**

**a**, Cycling performance of Li/Cl<sub>2</sub> battery at 500 mAh g<sup>-1</sup> with 150 mA g<sup>-1</sup> and 100 mA g<sup>-1</sup> currents (the first five cycles were cycling at 150 mA g<sup>-1</sup> and starting from cycle 6 the current was 100 mA g<sup>-1</sup>). The loading of the battery was about 4.5 mg cm<sup>-2</sup>. **b**, Typical charge–discharge curve of Li/Cl<sub>2</sub> battery at 500 mAh g<sup>-1</sup> cycling capacity. The loading of the battery was about 4.5 mg cm<sup>-2</sup>.

**Extended Data Table 1 Average surface area, pore volume (micropores and mesopores) of different carbon materials, and the first discharge capacity of Na/Cl<sub>2</sub> battery using AB, KJ and aCNS as the positive electrode**

## Supplementary information

### **Supplementary Information**

This file contains Supplementary Discussion, Supplementary Figures and Tables, and additional references. The Supplementary Discussion includes discussions related to the main manuscript such as details on mass spectrometry experiments, battery working mechanisms, and effects of different electrolyte additives. The 14 display

items – Supplementary Figures 1-12 and Supplementary Tables 1-2 – show data relating to battery performances, battery characterizations, and mass spectrometry experiments.

## Rights and permissions

### [Reprints and Permissions](#)

## About this article

### Cite this article

Zhu, G., Tian, X., Tai, HC. *et al.* Rechargeable Na/Cl<sub>2</sub> and Li/Cl<sub>2</sub> batteries. *Nature* **596**, 525–530 (2021). <https://doi.org/10.1038/s41586-021-03757-z>

- Received: 24 September 2020
- Accepted: 22 June 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03757-z>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03757-z>

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

- Article
- [Published: 25 August 2021](#)

# Direct observation of ultrafast hydrogen bond strengthening in liquid water

- [Jie Yang](#) [ORCID: orcid.org/0000-0002-5115-2413](#)<sup>1,2</sup> nAff11,
- [Riccardo Dettori](#)<sup>3</sup>,
- [J. Pedro F. Nunes](#) [ORCID: orcid.org/0000-0003-0670-6023](#)<sup>4</sup>,
- [Nanna H. List](#) [ORCID: orcid.org/0000-0002-0246-3995](#)<sup>1,2,5</sup>,
- [Elisa Biasin](#)<sup>1,2</sup>,
- [Martin Centurion](#) [ORCID: orcid.org/0000-0002-5662-2293](#)<sup>4</sup>,
- [Zhijiang Chen](#) [ORCID: orcid.org/0000-0002-2546-7075](#)<sup>1</sup>,
- [Amy A. Cordones](#) [ORCID: orcid.org/0000-0001-9897-5380](#)<sup>2</sup>,
- [Daniel P. Deponte](#)<sup>1</sup>,
- [Tony F. Heinz](#)<sup>2,6</sup>,
- [Michael E. Kozina](#)<sup>1,2</sup>,
- [Kathryn Ledbetter](#) [ORCID: orcid.org/0000-0002-3961-8524](#)<sup>2,7</sup>,
- [Ming-Fu Lin](#) [ORCID: orcid.org/0000-0001-8086-2484](#)<sup>1</sup>,
- [Aaron M. Lindenberg](#) [ORCID: orcid.org/0000-0003-3233-7161](#)<sup>2,8,9</sup>,
- [Mianzhen Mo](#) [ORCID: orcid.org/0000-0002-2962-0815](#)<sup>1</sup>,
- [Anders Nilsson](#)<sup>10</sup>,
- [Xiaozhe Shen](#) [ORCID: orcid.org/0000-0002-6844-608X](#)<sup>1</sup>,
- [Thomas J. A. Wolf](#) [ORCID: orcid.org/0000-0002-0641-1279](#)<sup>1,2</sup>,
- [Davide Donadio](#) [ORCID: orcid.org/0000-0002-2150-4182](#)<sup>3</sup>,
- [Kelly J. Gaffney](#) [ORCID: orcid.org/0000-0002-0525-6465](#)<sup>2</sup>,
- [Todd J. Martinez](#) [ORCID: orcid.org/0000-0002-4798-8947](#)<sup>1,2,5</sup> &
- [Xijie Wang](#) [ORCID: orcid.org/0000-0003-3324-4709](#)<sup>1</sup>

[Nature](#) volume **596**, pages 531–535 (2021)

- 3180 Accesses

- 276 Altmetric
- [Metrics details](#)

## Subjects

- [Chemical physics](#)
- [Photochemistry](#)

## Abstract

Water is one of the most important, yet least understood, liquids in nature. Many anomalous properties of liquid water originate from its well-connected hydrogen bond network<sup>1</sup>, including unusually efficient vibrational energy redistribution and relaxation<sup>2</sup>. An accurate description of the ultrafast vibrational motion of water molecules is essential for understanding the nature of hydrogen bonds and many solution-phase chemical reactions. Most existing knowledge of vibrational relaxation in water is built upon ultrafast spectroscopy experiments<sup>2,3,4,5,6,7</sup>. However, these experiments cannot directly resolve the motion of the atomic positions and require difficult translation of spectral dynamics into hydrogen bond dynamics. Here, we measure the ultrafast structural response to the excitation of the OH stretching vibration in liquid water with femtosecond temporal and atomic spatial resolution using liquid ultrafast electron scattering. We observed a transient hydrogen bond contraction of roughly 0.04 Å on a timescale of 80 femtoseconds, followed by a thermalization on a timescale of approximately 1 picosecond. Molecular dynamics simulations reveal the need to treat the distribution of the shared proton in the hydrogen bond quantum mechanically to capture the structural dynamics on femtosecond timescales. Our experiment and simulations unveil the intermolecular character of the water vibration preceding the relaxation of the OH stretch.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

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

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

## Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

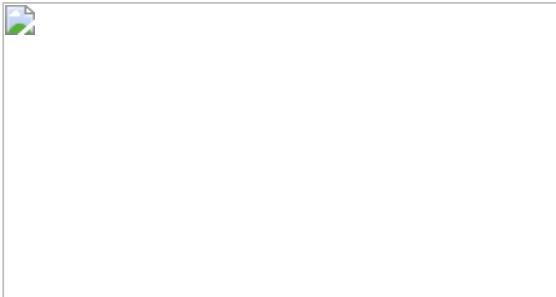
\$19.99/month

[\*\*Learn more\*\*](#)

### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

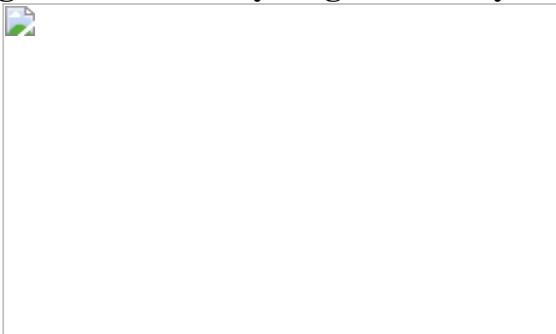
**Fig. 1: Experiment overview.**



**Fig. 2: Transient hydrogen bond strengthening.**



**Fig. 3: First shell hydrogen atom dynamics.**



**Fig. 4: Thermalization.**

## Data availability

Experimental data were generated at the MeV-UED facility at the SLAC National Accelerator Laboratory. Data behind each figure are available in Zenodo with the

identifier <https://doi.org/10.5281/zenodo.4678299>. Raw datasets are available from the corresponding authors on reasonable request. [Source data](#) are provided with this paper.

## Code availability

The non-commercial codes used for the simulation and analysis here are available from the corresponding authors on reasonable request.

## References

1. 1.

Stillinger, F. H. Water revisited. *Science* **209**, 451–457 (1980).

2. 2.

Perakis, F. et al. Vibrational spectroscopy and dynamics of water. *Chem. Rev.* **116**, 7590–7607 (2016).

3. 3.

Lindner, J. et al. Vibrational relaxation of pure liquid water. *Chem. Phys. Lett.* **421**, 329–333 (2006).

4. 4.

Ramasesha, K., De Marco, L., Mandal, A. & Tokmakoff, A. Water vibrations have strongly mixed intra- and intermolecular character. *Nat. Chem.* **5**, 935–940 (2013).

5. 5.

Bakker, H. J. et al. Transient absorption of vibrationally excited water. *J. Chem. Phys.* **116**, 2592–2598 (2002).

6. 6.

Fecko, C. J., Eaves, J. D., Loparo, J. J., Tokmakoff, A. & Geissler, P. L. Ultrafast hydrogen-bond dynamics in the infrared spectroscopy of water. *Science* **301**, 1698–1702 (2003).

7. 7.

Ashihara, S., Huse, N., Espagne, A., Nibbering, E. T. J. & Elsaesser, T. Ultrafast structural dynamics of water induced by dissipation of vibrational energy. *J. Phys. Chem. A* **111**, 743–746 (2007).

8. 8.

Auer, B. M. & Skinner, J. L. IR and Raman spectra of liquid water: theory and interpretation. *J. Chem. Phys.* **128**, 224511 (2008).

9. 9.

Amann-Winkel, K. et al. X-ray and neutron scattering of water. *Chem. Rev.* **116**, 7570–7589 (2016).

10. 10.

Sellberg, J. A. et al. Ultrafast X-ray probing of water structure below the homogeneous ice nucleation temperature. *Nature* **510**, 381–384 (2014).

11. 11.

Beyerlein, K. R. et al. Ultrafast nonthermal heating of water initiated by an X-ray free-electron laser. *Proc. Natl Acad. Sci. USA* **115**, 5652–5657 (2018).

12. 12.

Wen, H., Huse, N., Schoenlein, R. W. & Lindenberg, A. M. Ultrafast conversions between hydrogen bonded structures in liquid water observed by femtosecond x-ray spectroscopy. *J. Chem. Phys.* **131**, 234505 (2009).

13. 13.

Perakis, F. et al. Coherent X-rays reveal the influence of cage effects on ultrafast water dynamics. *Nat. Commun.* **9**, 1917 (2018).

14. 14.

Kim, K. H. et al. Direct observation of bond formation in solution with femtosecond X-ray scattering. *Nature* **518**, 385–389 (2015).

15. 15.

van Driel, T. B. et al. Atomistic characterization of the active-site solvation dynamics of a model photocatalyst. *Nat. Commun.* **7**, 13678 (2016).

16. 16.

Nunes, J. P. F. et al. Liquid-phase mega-electron-volt ultrafast electron diffraction. *Struct. Dyn.* **7**, 024301 (2020).

17. 17.

Koralek, J. D. et al. Generation and characterization of ultrathin free-flowing liquid sheets. *Nat. Commun.* **9**, 1353 (2018).

18. 18.

Ihee, H. Visualizing solution-phase reaction dynamics with time-resolved X-ray liquidography. *Acc. Chem. Res.* **42**, 356–366 (2009).

19. 19.

Skinner, L. B., Benmore, C. J., Neufeind, J. C. & Parise, J. B. The structure of water around the compressibility minimum. *J. Chem. Phys.* **141**, 214507 (2014).

20. 20.

Haldrup, K. et al. Observing solvation dynamics with simultaneous femtosecond X-ray emission spectroscopy and X-ray scattering. *J. Phys. Chem. B* **120**, 1158–1168 (2016).

21. 21.

Morawietz, T., Singraber, A., Dellago, C. & Behler, J. How van der Waals interactions determine the unique properties of water. *Proc. Natl Acad. Sci. USA* **113**, 8368–8373 (2016).

22. 22.

Dettori, R. et al. Simulating energy relaxation in pump–probe vibrational spectroscopy of hydrogen-bonded liquids. *J. Chem. Theory Comput.* **13**, 1284–1292 (2017).

23. 23.

Dettori, R., Ceriotti, M., Hunger, J., Colombo, L. & Donadio, D. Energy relaxation and thermal diffusion in infrared pump–probe spectroscopy of hydrogen-bonded liquids. *J. Phys. Chem. Lett.* **10**, 3447–3452 (2019).

24. 24.

Gilli, P., Bertolasi, V., Ferretti, V. & Gilli, G. Evidence for resonance-assisted hydrogen bonding. 4. Covalent nature of the strong homonuclear hydrogen bond. Study of the O-H-O system by crystal structure correlation methods. *J. Am. Chem. Soc.* **116**, 909–915 (1994).

25. 25.

Lippincott, E. R. & Schroeder, R. One-dimensional model of the hydrogen bond. *J. Chem. Phys.* **23**, 1099–1106(1955).

26. 26.

Staib, A. & Hynes, J. T. Vibrational predissociation in hydrogen-bonded OH...O complexes via OH stretch-OO stretch energy transfer. *Chem. Phys. Lett.* **204**, 197–205 (1993).

27. 27.

McKenzie, R. H., Bekker, C., Athokpam, B. & Ramesh, S. G. Effect of quantum nuclear motion on hydrogen bonding. *J. Chem. Phys.* **140**, 174508 (2014).

28. 28.

Grabowski, S. J. What is the covalency of hydrogen bonding? *Chem. Rev.* **111**, 2597–2625 (2011).

29. 29.

Shibata, S. & Bartell, L. S. Electron-diffraction study of water and heavy water. *J. Chem. Phys.* **42**, 1147–1151 (1965).

30. 30.

Ceriotti, M. et al. Nuclear quantum effects in water and aqueous systems: experiment, theory, and current challenges. *Chem. Rev.* **116**, 7529–7550 (2016).

31. 31.

Weathersby, S. P. et al. Mega-electron-volt ultrafast electron diffraction at SLAC National Accelerator Laboratory. *Rev. Sci. Instrum.* **86**, 073702 (2015).

32. 32.

Bertie, J. E. & Lan, Z. Infrared intensities of liquids XX: the intensity of the OH stretching band of liquid water revisited, and the best current values of the optical constants of H<sub>2</sub>O(l) at 25 °C between 15,000 and 1 cm<sup>-1</sup>. *Appl. Spectrosc.* **50**, 1047–1057 (1996).

33. 33.

Sorenson, J. M., Hura, G., Glaeser, R. M. & Head-Gordon, T. What can x-ray scattering tell us about the radial distribution functions of water? *J. Chem. Phys.* **113**, 9149–9161 (2000).

34. 34.

Brockway, L. O. Electron diffraction by gas molecules. *Rev. Mod. Phys.* **8**, 0231–0266 (1936).

35. 35.

Hubbell, J. H. et al. Atomic form factors, incoherent scattering functions, and photon scattering cross sections. *J. Phys. Chem. Ref. Data* **4**, 471–538 (1975).

36. 36.

Yang, J. et al. Structure retrieval in liquid-phase electron scattering. *Phys. Chem. Chem. Phys.* **23**, 1308–1316 (2021).

37. 37.

Horn, H. W. et al. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **120**, 9665–9678 (2004).

38. 38.

Levine, B. G., Stone, J. E. & Kohlmeyer, A. Fast analysis of molecular dynamics trajectories with graphics processing units—radial distribution function histogramming. *J. Comput. Phys.* **230**, 3556–3569 (2011).

39. 39.

Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

40. 40.

Dohn, A. O. et al. On the calculation of X-ray scattering signals from pairwise radial distribution functions. *J. Phys. B* **48**, 244010 (2015).

41. 41.

Bartell, L. S. & Gavin, R. M. Effects of electron correlation in x-ray and electron diffraction. *J. Am. Chem. Soc.* **86**, 3493–3498 (1964).

42. 42.

Wang, J. H., Tripathi, A. N. & Smith, V. H. Chemical-binding and electron correlation-effects in x-ray and high-energy electron-scattering. *J. Chem. Phys.* **101**, 4842–4854 (1994).

43. 43.

Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **59**, 7413–7421 (1999).

44. 44.

Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).

45. 45.

Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).

46. 46.

Zwanzig, R. in *Non-equilibrium Statistical Mechanics*, 1st edn, ch. 1, 18–21 (Oxford Univ. Press, 2001).

47. 47.

Ceriotti, M. & Parrinello, M. The  $\delta$ -thermostat: selective normal-modes excitation by colored-noise Langevin dynamics. *Procedia Comput. Sci.* **1**, 1607–1614 (2010).

48. 48.

Marx, D. & Parrinello, M. Ab initio path integral molecular dynamics: basic ideas. *J. Chem. Phys.* **104**, 4077–4082 (1996).

49. 49.

Cao, J. & Voth, G. A. The formulation of quantum statistical mechanics based on the Feynman path centroid density. III. Phase space formalism and analysis of centroid molecular dynamics. *J. Chem. Phys.* **101**, 6157–6167 (1994).

50. 50.

Markland, T. E. & Ceriotti, M. Nuclear quantum effects enter the mainstream. *Nature Reviews Chemistry* **2**, 0109, <https://doi.org/10.1038/s41570-017-0109> (2018).

51. 51.

Hockney, R. W. & Eastwood, J. W. *Computer simulation using particles*, 1st edn, ch. 1 (Taylor & Francis, 1988).

52. 52.

Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).

53. 53.

Dunning, T. H. Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90**, 1007–1023 (1989).

54. 54.

Ufimtsev, I. S. & Martínez, T. J. Quantum chemistry on graphical processing units. 1. Strategies for two-electron integral evaluation. *J. Chem. Theory Comput.* **4**, 222–231 (2008).

55. 55.

Ufimtsev, I. S. & Martinez, T. J. Quantum chemistry on graphical processing units. 2. Direct self-consistent-field implementation. *J. Chem. Theory Comput.* **5**, 1004–1015 (2009).

56. 56.

Ufimtsev, I. S. & Martinez, T. J. Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J. Chem. Theory Comput.* **5**, 2619–2628 (2009).

57. 57.

Yang, J. et al. Simultaneous observation of nuclear and electronic dynamics by ultrafast electron diffraction. *Science* **368**, 885–889 (2020).

58. 58.

Shibata, S., Sekiyama, H., Tachikawa, K. & Moribe, M. Chemical bonding effect in electron scattering by gaseous molecules. *J. Mol. Struct.* **641**, 1–6 (2002).

59. 59.

Jensen, P. Hamiltonians for the internal dynamics of triatomic molecules. *J. Chem. Soc. Faraday Trans.* **84**, 1315–1339 (1988).

60. 60.

Wilson, E. B. J., Decius, J. C. & Cross, P. C. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra* (Dover Publications, 2012).

61. 61.

Ceriotti, M., Bussi, G. & Parrinello, M. Colored-noise thermostats à la carte. *J. Chem. Theory Comput.* **6**, 1170–1180 (2010).

62. 62.

Ceriotti, M., Bussi, G. & Parrinello, M. Nuclear quantum effects in solids using a colored-noise thermostat. *Phys. Rev. Lett.* **103**, 030603 (2009).

63. 63.

Marsalek, O. & Markland, T. E. Ab initio molecular dynamics with nuclear quantum effects at classical cost: ring polymer contraction for density functional theory. *J. Chem. Phys.* **144**, 054112 (2016).

64. 64.

Habershon, S. & Manolopoulos, D. E. Zero point energy leakage in condensed phase dynamics: an assessment of quantum simulation methods for liquid water. *J. Chem. Phys.* **131**, 244518 (2009).

65. 65.

Salvat, F., Jablonski, A. & Powell, C. J. ELSEPA - Dirac partial-wave calculation of elastic scattering of electrons and positrons by atoms, positive ions and molecules. *Comput. Phys. Commun.* **165**, 157–190 (2005).

66. 66.

Kim, J. G. et al. Mapping the emergence of molecular vibrations mediating bond formation. *Nature* **582**, 520–524 (2020).

67. 67.

Császár, A. G. et al. First-principles prediction and partial characterization of the vibrational states of water up to dissociation. *J. Quant. Spectrosc. Radiat. Transfer* **111**, 1043–1064 (2010).

68. 68.

Soper, A. K. Joint structure refinement of X-ray and neutron diffraction data on disordered materials: application to liquid water. *J. Phys. Condens. Matter* **19**, 335206 (2007).

## Acknowledgements

We thank T. E. Markland for helpful discussions, and G. M. Stewart for help in producing Fig. 1a. The experiment was performed at the SLAC MeV-UED facility, which is supported in part by the US DOE BES SUF division Accelerator and Detector R&D program, the LCLS Facility, and SLAC under contract nos. DE-AC02-05-CH11231 and DE-AC02-76SF00515. J.P.F.N. and M.C. are supported by the US DOE Office of Science, Basic Energy Sciences under award no. DE-SC0014170. K.L. is supported by a Melvin and Joan Lane Stanford Graduate

Fellowship and a Stanford Physics Department fellowship. J.Y., T.F.H., A.A.C., T. J. A.W., E.B., N.H.L., T.J.M. and K.J.G. were supported by the US DOE Office BES, Chemical Sciences, Geosciences, and Biosciences division. A.M.L. acknowledges support from the DOE BES Materials Science and Engineering division under contract DE-AC02-76SF00515. Z.C. and M.M. are supported by the DOE Fusion Energy Sciences under fieldwork proposal no. 100182.

## Author information

### Author notes

#### 1. Jie Yang

Present address: Center of Basic Molecular Science, Department of Chemistry, Tsinghua University, Beijing, China

### Affiliations

#### 1. SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Jie Yang, Nanna H. List, Elisa Biasin, Zhijiang Chen, Daniel P. Deonte, Michael E. Kozina, Ming-Fu Lin, Mianzhen Mo, Xiaozhe Shen, Thomas J. A. Wolf, Todd J. Martinez & Xijie Wang

#### 2. Stanford PULSE Institute, SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Jie Yang, Nanna H. List, Elisa Biasin, Amy A. Cordones, Tony F. Heinz, Michael E. Kozina, Kathryn Ledbetter, Aaron M. Lindenberg, Thomas J. A. Wolf, Kelly J. Gaffney & Todd J. Martinez

#### 3. Department of Chemistry, University of California-Davis, Davis, CA, USA

Riccardo Dettori & Davide Donadio

#### 4. Department of Physics, University of Nebraska-Lincoln, Lincoln, NE, USA

J. Pedro F. Nunes & Martin Centurion

#### 5. Department of Chemistry, Stanford University, Stanford, CA, USA

Nanna H. List & Todd J. Martinez

6. Department of Applied Physics, Stanford University, Stanford, CA, USA

Tony F. Heinz

7. Department of Physics, Stanford University, Stanford, CA, USA

Kathryn Ledbetter

8. Department of Materials Science and Engineering, Stanford University, Stanford, CA, USA

Aaron M. Lindenberg

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

Aaron M. Lindenberg

10. Department of Physics, AlbaNova University Center, Stockholm University, Stockholm, Sweden

Anders Nilsson

## Contributions

J.Y., K.J.G., A.M.L. and X.W. proposed the study. J.P.F.N., K.L., E.B., M.C., D.P.D., M.F.-L., M.M., X.S., T.J.A.W., J.Y., A.A.C. and X.W. developed the experimental setup. M.E.K. developed the pump laser setup. J.Y., J.P.F.N., E.B., Z.C., A.A.C., T.F.H., K.L., M.F.-L., M.M., X.S., T.J.A.W and X.W performed the experiment. J.Y. analysed the experimental data and performed the  $\chi^2$  fitting. J.Y., A.N., T.J.M. and K.J.G. interpreted the experimental data. R.D. and D.D. performed the pump-probe molecular dynamics simulation. J.P.F.N. performed the equilibrium water simulation. N.H.L. and T.J.M. performed the 1D and 2D quantum simulations and the ab initio electron scattering simulation. J.Y., R.D., N.H.L., D.D., T.J.M., K.J.G. and X.W. wrote the manuscript with input from all authors.

## Corresponding authors

Correspondence to [Jie Yang](#) or [Davide Donadio](#) or [Kelly J. Gaffney](#) or [Todd J. Martinez](#) or [Xijie Wang](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

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

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

## Extended data figures and tables

### Extended Data Fig. 1 Extra information for data interpretation.

**a**, Ab initio simulation of the inelastic and elastic scattering signal change for  $\nu_{\text{OH}} = 1$  in comparison to  $\nu_{\text{OH}} = 0$ . The simulation is performed on a single water molecule with OH bond lengths adjusted to the equilibrium length for each vibrational state as predicted in ref. [5](#) (for more details, see [Methods](#)). **b**, Spectrum of the second and the third harmonics of the pump laser. **c**, Experimental  $g_2$  and temperature evolution up to 100 ps. **d**, Damped  $Q\Delta S$  from experimental data. This is related to Fig. [1c](#) by the damping term  $\backslash(\{e\}^{-0.03\{Q\}^2})$ ; equation (3) in [Methods](#)

[Source data.](#)

### Extended Data Fig. 2 Wigner sampling.

**a**, The three lowest eigenstates (coloured lines) and eigenvalues (horizontal grey lines) of the Lippincott–Schroeder model potential (black line). Inset, the probability distribution of the  $\nu_{\text{OH}} = 0$  and  $\nu_{\text{OH}} = 1$  states,  $\mu$  and  $\sigma$  represent mean and standard deviation. **b**, **c**, Wigner distribution for  $\nu_{\text{OH}} = 0$  (**b**) and  $\nu_{\text{OH}} = 1$  (**c**). The region of phase space with negative values of the  $\nu_{\text{OH}} = 1$  distribution (orange shades) was excluded from the sampling. Note the different colour gradient used for negative function values. Lippincott–Schroeder model ( $R_{\text{OO}} = 2.85 \text{ \AA}$ ) is used for sampling of the initial displacements and velocities along the OH bonds of the excited molecules

[Source data.](#)

### Extended Data Fig. 3 Probability density from classical and Wigner sampling.

**a**, Wigner sampling. Magenta represents  $v_{\text{OH}} = 0$ , yellow represents  $v_{\text{OH}} = 1$ . **b**, Classical sampling. Magenta and yellow represent unexcited and excited molecules, respectively, calculated by averaging over the final 10 fs window during the excitation phase. Dashed black line represents the equilibrium water before excitation. The vertical dotted lines represent the equilibrium distance for each curve, and  $\mu$  and  $\sigma$  represent the mean and standard deviation of each curve, respectively

[Source data](#).

### Extended Data Fig. 4 Examples of pair distances shift.

**a**,  $g_{\text{OO}}(r)$  around the first OO peak for four different  $\Delta R_1$ . **b**,  $\Delta \text{PDF}_{\text{OO}}$  for three different  $\Delta R_1$ . **c**,  $\Delta \text{PDF}_{\text{OH}}$  for three different  $\Delta r_2$ . **d**,  $\Delta \text{PDF}_{\text{OH}}$  for three different  $\Delta r_3$

[Source data](#).

### Extended Data Fig. 5 CPDF analysis.

**a**, A comparison of experimental and simulated CPDF. The overall scaling factor is achieved by matching the height of the first OO between experimental and simulated curves. The simulation is a 275 K water box under equilibrium condition. **b**, The simulated elastic and inelastic components of the CPDF, the inelastic component is concentrated to  $r < 2.5 \text{ \AA}$ . Exp., experimental; Sim., simulated. **c**, CPDF for five delay windows (see the key) in full  $r$  range. **d**, CPDF for five delay windows (see the key) around the second OO shell. The peak height around  $4.6 \text{ \AA}$  is used to extract  $g_2$  for Fig. 4a

[Source data](#).

### Extended Data Fig. 6 Comparison of equilibrium $\Delta \text{PDF}$ simulation.

$\Delta \text{PDF}$  from experiment at 2.2 ps (blue with error bars), simulation using Tip4p-Ew force field (orange) and simulation using machine-learning force field (yellow)

[Source data](#).

### Extended Data Fig. 7 $\Delta \text{PDF}$ simulated using different methods.

**a–c**,  $\Delta$ PDF consistency. **a**, The  $\Delta$ PDF simulated using the conventional method (that is, by first simulating the electron scattering pattern using equation (7), then transforming to real space using equation (3)). **b**, The  $\Delta$ PDF simulated by directly applying equation (4), and smoothed by convolution with a Gaussian kernel with a FWHM of 0.53 Å. The weight of OO, OH and HH pairs are chosen to be 1, 0.4 and 0.16, respectively, obtained by atomic scattering cross section and the relative number of each types of atom pairs. The 0.53 Å FWHM of the Gaussian Kernel is obtained using  $2\pi/Q_{\max}$ , where  $Q_{\max} = 11.8\text{\AA}^{-1}$  is the maximum  $Q$  range in this experiment. **c**, The  $\Delta$ PDF simulated by directly applying equation (4) without Gaussian smoothing. The vertical scales of all subpanels are identical. **d–f**, Comparison of the  $\Delta$ PDF in quantum simulations (**d**), classical simulations with  $h\nu$  excitation (**e**) and classical simulations with 3/2  $h\nu$  excitation (**f**)

[Source data.](#)

### Extended Data Fig. 8 Simulated instantaneous kinetic temperature evolution.

**a, b**, Classical excitation during the 100 fs excitation phase (**a**), and during the 3 ps relaxation phase (**b**). **c, d**, Quantum excitation, with  $v_{\text{OH}} = 1$  (**c**) and  $v_{\text{OH}} = 0$  (**d**).  $T_{\text{stretch}}$  and  $T_{\text{rot}}$  are defined in equation (11) and equation (12). In **c**, the subscript ‘Stretch1’ and ‘Rot1’ indicate the OH bond corresponding to  $v_{\text{OH}} = 1$  Wigner sampling, and ‘Stretch2’ and ‘Rot2’ indicate the OH bond corresponding to  $v_{\text{OH}} = 0$  Wigner sampling. The superscript ‘excited’ indicates Wigner sampling. Excited and unexcited molecules are calculated separately. The initial temperature before excitation is 300 K

[Source data.](#)

### Extended Data Fig. 9 Comparison of NNP-based 2D OH stretching vibrational modes in gas phase and frozen phonon liquid phase.

**a–h**, The lowest vibrational eigenstates  $((\{n\}_1, \{n\}_2))$  for a representative configuration (bond angle of 104.4°) among the 200 2D potential energy surfaces considered (**a–c, e–g**; dashed black lines indicate symmetric and antisymmetric displacements); and distribution of vibrational frequencies (defined as  $\backslash(\varDelta \{\nu\}_{\{n\}_1, \{n\}_2} = \{\nu\}_{\{n\}_1, \{n\}_2} - \{\nu\}_{0,0})$ ) for the two lowest OH stretching vibrationally excited states for the 200 configurations (**d, h**). The distribution in the gas phase originates from the variation in the bond angle. The vertical lines indicate the experimental gas-phase stretch frequencies<sup>67</sup> and \

(`\varDelta \{\nu\}_1`) from the 1D Lippincott–Schroeder model, respectively. **i**, Comparison of 1D OH stretch potentials for gas phase and liquid water as obtained from the NNP (blue and red, respectively) and the Lippincott–Schroeder model (black). The transparent thin lines correspond to the underlying  $2 \times 200$  NNP replicates while the corresponding thick lines indicate the average potentials

[Source data](#).

### **Extended Data Fig. 10 Zero-point energy leakage time.**

**a–c**, Comparison of the OO (**a**), OH (**b**) and HH (**c**) RDFs computed during an equilibrium run for a classical distribution of positions and momenta (NVT), during the coupling with the quantum GLE thermostat, from ab initio PIMD simulations<sup>63</sup> and measured from neutron diffraction experiments<sup>68</sup>. The inset in **b** is a zoom-in on the OH bond peak where, due to the absence of experimental data to compare with, we reported the comparison with DFT-based PIMD simulations. **d**, Kinetic energies computed during the coupling with the quantum thermostat. **e**, Kinetic energies computed during the NVE simulations. The inset in **d** shows a temporal fitting of the stretching temperature decay. **f**, Time-resolved RDF computed during the NVE relaxation. The black curve refers to the NVT-computed RDF, obtained at  $T = 300$  K. The inset shows the shift of the  $R_1$  distance during the system relaxation

[Source data](#).

## **Supplementary information**

### **Peer Review File**

## **Source data**

### **Source Data Fig. 3**

### **Source Data Fig. 4**

### **Source Data Extended Data Fig. 1**

### **Source Data Extended Data Fig. 2**

### **Source Data Extended Data Fig. 3**

[\*\*Source Data Extended Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 5\*\*](#)

[\*\*Source Data Extended Data Fig. 6\*\*](#)

[\*\*Source Data Extended Data Fig. 7\*\*](#)

[\*\*Source Data Extended Data Fig. 8\*\*](#)

[\*\*Source Data Extended Data Fig. 9\*\*](#)

[\*\*Source Data Extended Data Fig. 10\*\*](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Yang, J., Dettori, R., Nunes, J.P.F. *et al.* Direct observation of ultrafast hydrogen bond strengthening in liquid water. *Nature* **596**, 531–535 (2021).  
<https://doi.org/10.1038/s41586-021-03793-9>

- Received: 28 September 2020
- Accepted: 01 July 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03793-9>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03793-9>

- Article
- [Published: 25 August 2021](#)

# High aboveground carbon stock of African tropical montane forests

- [Aida Cuni-Sanchez](#) ORCID: [orcid.org/0000-0001-8619-1095](https://orcid.org/0000-0001-8619-1095)<sup>1,2</sup>,
- [Martin J. P. Sullivan](#) ORCID: [orcid.org/0000-0002-5955-0483](https://orcid.org/0000-0002-5955-0483)<sup>3,4</sup>,
- [Philip J. Platts](#) ORCID: [orcid.org/0000-0002-0153-0121](https://orcid.org/0000-0002-0153-0121)<sup>1,5,6</sup>,
- [Simon L. Lewis](#) ORCID: [orcid.org/0000-0002-8066-6851](https://orcid.org/0000-0002-8066-6851)<sup>4,7</sup>,
- [Rob Marchant](#)<sup>1</sup>,
- [Gérard Imani](#)<sup>8</sup>,
- [Wannes Hubau](#) ORCID: [orcid.org/0000-0003-3795-4986](https://orcid.org/0000-0003-3795-4986)<sup>9,10</sup>,
- [Iveren Abiem](#)<sup>11,12</sup>,
- [Hari Adhikari](#) ORCID: [orcid.org/0000-0002-9089-3249](https://orcid.org/0000-0002-9089-3249)<sup>13</sup>,
- [Tomas Albrecht](#) ORCID: [orcid.org/0000-0002-9213-0034](https://orcid.org/0000-0002-9213-0034)<sup>14,15</sup>,
- [Jan Altman](#) ORCID: [orcid.org/0000-0003-4879-5773](https://orcid.org/0000-0003-4879-5773)<sup>16</sup>,
- [Christian Amani](#)<sup>8</sup>,
- [Abreham B. Aneseyee](#)<sup>17,18</sup>,
- [Valerio Avitabile](#) ORCID: [orcid.org/0000-0003-3646-052X](https://orcid.org/0000-0003-3646-052X)<sup>19</sup>,
- [Lindsay Banin](#) ORCID: [orcid.org/0000-0002-1168-3914](https://orcid.org/0000-0002-1168-3914)<sup>20</sup>,
- [Rodrigue Batumike](#)<sup>21</sup>,
- [Marijn Bauters](#) ORCID: [orcid.org/0000-0003-0978-6639](https://orcid.org/0000-0003-0978-6639)<sup>22</sup>,
- [Hans Beeckman](#) ORCID: [orcid.org/0000-0001-8954-6277](https://orcid.org/0000-0001-8954-6277)<sup>9</sup>,
- [Serge K. Begne](#)<sup>4,23</sup>,
- [Amy C. Bennett](#)<sup>4</sup>,
- [Robert Bitariho](#) ORCID: [orcid.org/0000-0002-3461-0013](https://orcid.org/0000-0002-3461-0013)<sup>24</sup>,
- [Pascal Boeckx](#) ORCID: [orcid.org/0000-0003-3998-0010](https://orcid.org/0000-0003-3998-0010)<sup>22</sup>,
- [Jan Bogaert](#)<sup>25</sup>,
- [Achim Bräuning](#) ORCID: [orcid.org/0000-0003-3106-4229](https://orcid.org/0000-0003-3106-4229)<sup>26</sup>,
- [Franklin Bulonvu](#)<sup>27</sup>,
- [Neil D. Burgess](#)<sup>28</sup>,
- [Kim Calders](#) ORCID: [orcid.org/0000-0002-4562-2538](https://orcid.org/0000-0002-4562-2538)<sup>29</sup>,
- [Colin Chapman](#)<sup>30,31,32,33</sup>,
- [Hazel Chapman](#) ORCID: [orcid.org/0000-0001-8509-703X](https://orcid.org/0000-0001-8509-703X)<sup>12,34</sup>,

- [James Comiskey](#) ORCID: [orcid.org/0000-0001-6710-1269<sup>35</sup>](http://orcid.org/0000-0001-6710-1269),
- [Thales de Haulleville<sup>36</sup>](#),
- [Mathieu Decuyper](#) ORCID: [orcid.org/0000-0002-1713-8562<sup>37,38</sup>](http://orcid.org/0000-0002-1713-8562),
- [Ben DeVries<sup>39</sup>](#),
- [Jiri Dolezal](#) ORCID: [orcid.org/0000-0002-5829-4051<sup>16,40</sup>](http://orcid.org/0000-0002-5829-4051),
- [Vincent Droissart](#) ORCID: [orcid.org/0000-0001-9798-5616<sup>23,41</sup>](http://orcid.org/0000-0001-9798-5616),
- [Corneille Ewango<sup>42</sup>](#),
- [Senbeta Feyera<sup>43</sup>](#),
- [Aster Gebrekirstos<sup>44</sup>](#),
- [Roy Gereau](#) ORCID: [orcid.org/0000-0002-5863-4466<sup>45</sup>](http://orcid.org/0000-0002-5863-4466),
- [Martin Gilpin<sup>4</sup>](#),
- [Dismas Hakizimana<sup>46</sup>](#),
- [Jefferson Hall](#) ORCID: [orcid.org/0000-0003-4761-9268<sup>47</sup>](http://orcid.org/0000-0003-4761-9268),
- [Alan Hamilton<sup>48</sup>](#),
- [Olivier Hardy<sup>49</sup>](#),
- [Terese Hart<sup>50</sup>](#),
- [Janne Heiskanen](#) ORCID: [orcid.org/0000-0002-3899-8860<sup>13,51</sup>](http://orcid.org/0000-0002-3899-8860),
- [Andreas Hemp<sup>52</sup>](#),
- [Martin Herold](#) ORCID: [orcid.org/0000-0003-0246-6886<sup>38,53</sup>](http://orcid.org/0000-0003-0246-6886),
- [Ulrike Hiltner<sup>26,54</sup>](#),
- [David Horak](#) ORCID: [orcid.org/0000-0002-8073-1617<sup>55</sup>](http://orcid.org/0000-0002-8073-1617),
- [Marie-Noel Kamdem<sup>23</sup>](#),
- [Charles Kayijamahe<sup>56</sup>](#),
- [David Kenfack](#) ORCID: [orcid.org/0000-0001-8208-3388<sup>47</sup>](http://orcid.org/0000-0001-8208-3388),
- [Mwangi J. Kinyanjui<sup>57</sup>](#),
- [Julia Klein<sup>58</sup>](#),
- [Janvier Lisingo<sup>42</sup>](#),
- [Jon Lovett<sup>4</sup>](#),
- [Mark Lung<sup>59</sup>](#),
- [Jean-Remy Makana<sup>60</sup>](#),
- [Yadvinder Malhi](#) ORCID: [orcid.org/0000-0002-3503-4783<sup>61</sup>](http://orcid.org/0000-0002-3503-4783),
- [Andrew Marshall<sup>1,62,63</sup>](#),
- [Emanuel H. Martin](#) ORCID: [orcid.org/0000-0003-1801-5131<sup>64</sup>](http://orcid.org/0000-0003-1801-5131),
- [Edward T. A. Mitchard](#) ORCID: [orcid.org/0000-0002-5690-4055<sup>65</sup>](http://orcid.org/0000-0002-5690-4055),
- [Alexandra Morel](#) ORCID: [orcid.org/0000-0002-0905-8079<sup>66</sup>](http://orcid.org/0000-0002-0905-8079),
- [John T. Mukendi<sup>9</sup>](#),
- [Tom Muller<sup>67</sup>](#),
- [Felix Nchu](#) ORCID: [orcid.org/0000-0002-9910-5085<sup>68</sup>](http://orcid.org/0000-0002-9910-5085),

- [Brigitte Nyirambangutse](#) ORCID: orcid.org/0000-0002-4395-7641<sup>69,70</sup>,
- [Joseph Okello](#) ORCID: orcid.org/0000-0003-4462-3923<sup>22,71,72</sup>,
- [Kelvin S.-H. Peh](#) ORCID: orcid.org/0000-0002-2921-1341<sup>73,74</sup>,
- [Petri Pellikka](#)<sup>13,75</sup>,
- [Oliver L. Phillips](#) ORCID: orcid.org/0000-0002-8993-6168<sup>4</sup>,
- [Andrew Plumptre](#) ORCID: orcid.org/0000-0002-9333-4047<sup>76</sup>,
- [Lan Qie](#)<sup>77</sup>,
- [Francesco Rovero](#)<sup>78,79</sup>,
- [Moses N. Sainge](#)<sup>80</sup>,
- [Christine B. Schmitt](#) ORCID: orcid.org/0000-0002-8527-9682<sup>81,82</sup>,
- [Ondrej Sedlacek](#)<sup>55</sup>,
- [Alain S. K. Ngute](#) ORCID: orcid.org/0000-0001-7090-5452<sup>62,83</sup>,
- [Douglas Sheil](#) ORCID: orcid.org/0000-0002-1166-6591<sup>84</sup>,
- [Demisse Sheleme](#)<sup>85</sup>,
- [Tibebu Y. Simegn](#)<sup>86</sup>,
- [Murielle Simo-Droissart](#)<sup>23</sup>,
- [Bonaventure Sonké](#) ORCID: orcid.org/0000-0002-4310-3603<sup>23</sup>,
- [Teshome Soromessa](#) ORCID: orcid.org/0000-0002-3680-0240<sup>17</sup>,
- [Terry Sunderland](#)<sup>87,88</sup>,
- [Miroslav Svoboda](#)<sup>89</sup>,
- [Hermann Taedoumg](#)<sup>90,91</sup>,
- [James Taplin](#) ORCID: orcid.org/0000-0002-7639-9244<sup>92</sup>,
- [David Taylor](#) ORCID: orcid.org/0000-0002-6098-5636<sup>93</sup>,
- [Sean C. Thomas](#) ORCID: orcid.org/0000-0002-0686-2483<sup>94</sup>,
- [Jonathan Timberlake](#)<sup>95</sup>,
- [Darlington Tuagben](#)<sup>96</sup>,
- [Peter Umunay](#)<sup>97</sup>,
- [Eustate Uzabaho](#) ORCID: orcid.org/0000-0002-7563-6455<sup>56</sup>,
- [Hans Verbeeck](#) ORCID: orcid.org/0000-0003-1490-0168<sup>29</sup>,
- [Jason Vleminckx](#) ORCID: orcid.org/0000-0002-7600-9170<sup>98</sup>,
- [Göran Wallin](#) ORCID: orcid.org/0000-0002-5359-1102<sup>70</sup>,
- [Charlotte Wheeler](#)<sup>65</sup>,
- [Simon Willcock](#) ORCID: orcid.org/0000-0001-9534-9114<sup>99,100</sup>,
- [John T. Woods](#)<sup>101</sup> &
- [Etienne Zibera](#)<sup>69</sup>

*Nature* volume **596**, pages 536–542 (2021)

- 1792 Accesses

- 295 Altmetric
- [Metrics details](#)

## Subjects

- [Carbon cycle](#)
- [Ecosystem services](#)
- [Forest ecology](#)

## Abstract

Tropical forests store 40–50 per cent of terrestrial vegetation carbon<sup>1</sup>. However, spatial variations in aboveground live tree biomass carbon (AGC) stocks remain poorly understood, in particular in tropical montane forests<sup>2</sup>. Owing to climatic and soil changes with increasing elevation<sup>3</sup>, AGC stocks are lower in tropical montane forests compared with lowland forests<sup>2</sup>. Here we assemble and analyse a dataset of structurally intact old-growth forests (AfriMont) spanning 44 montane sites in 12 African countries. We find that montane sites in the AfriMont plot network have a mean AGC stock of 149.4 megagrams of carbon per hectare (95% confidence interval 137.1–164.2), which is comparable to lowland forests in the African Tropical Rainforest Observation Network<sup>4</sup> and about 70 per cent and 32 per cent higher than averages from plot networks in montane<sup>2,5,6</sup> and lowland<sup>7</sup> forests in the Neotropics, respectively. Notably, our results are two-thirds higher than the Intergovernmental Panel on Climate Change default values for these forests in Africa<sup>8</sup>. We find that the low stem density and high abundance of large trees of African lowland forests<sup>4</sup> is mirrored in the montane forests sampled. This carbon store is endangered: we estimate that 0.8 million hectares of old-growth African montane forest have been lost since 2000. We provide country-specific montane forest AGC stock estimates modelled from our plot network to help to guide forest conservation and reforestation interventions. Our findings highlight the need for conserving these biodiverse<sup>9,10</sup> and carbon-rich ecosystems.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

## Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

\$19.99/month

[\*\*Learn more\*\*](#)

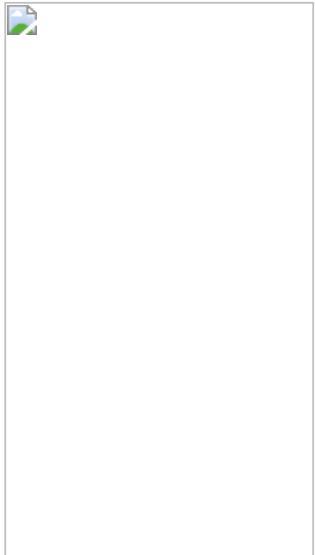
### Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Pantropical variation in AGC stocks sampled by plot networks in montane ( $\geq 800$  m a.s.l.) and lowland ( $< 800$  m a.s.l.) tropical forests.**



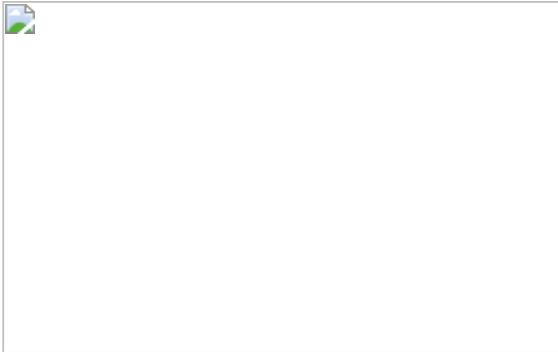
**Fig. 2: Proportion of plot-level AGC stock and stems accounted for by each size class in montane and lowland forests in Africa.**



**Fig. 3: Variables as a function of elevation.**



**Fig. 4: Old-growth evergreen humid forests in lowland and montane tropical Africa.**



## Data availability

Source data to generate figures and tables are available from  
[https://doi.org/10.5521/forestplots.net/2021\\_5](https://doi.org/10.5521/forestplots.net/2021_5).

## Code availability

The R code to generate figures and tables is available from  
[https://doi.org/10.5521/forestplots.net/2021\\_5](https://doi.org/10.5521/forestplots.net/2021_5).

## References

1. 1.

Erb, K. et al. Unexpectedly large impact of forest management and grazing on global vegetation biomass. *Nature* **553**, 73–76 (2018).

2. 2.

Spracklen, D. V. & Righelato, R. Tropical montane forests are a larger than expected global carbon store. *Biogeosciences* **11**, 2741–2754 (2014).

3. 3.

Fahey, T. J., Sherman, R. E. & Tanne, E. V. J. Tropical montane cloud forest: environmental drivers of vegetation structure and ecosystem function. *J. Trop. Ecol.* **32**, 355–367 (2016).

4. 4.

Lewis, S. L. et al. Above-ground biomass and structure of 260 African tropical forests. *Phil. Trans. R. Soc. Lond. B* **368**, 20120295 (2013).

5. 5.

Vilanova, E. et al. Environmental drivers of forest structure and stem turnover across Venezuelan tropical forests. *PLoS ONE* **13**, e0198489 (2018).

6. 6.

Álvarez-Dávila, E. et al. Forest biomass density across large climate gradients in northern South America is related to water availability but not with temperature. *PLoS ONE* **12**, e0171072 (2017).

7. 7.

Sullivan, M. J. P. et al. Long-term thermal sensitivity of Earth's tropical forests. *Science* **368**, 869–874 (2020).

8. 8.

Domke, G. et al. in 2019 *Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* Vol. 4 (eds Calvo Buendia, E. et al.) Ch. 4, 48 (IPCC, 2019).

9. 9.

*African Mountains Atlas* (UNEP, 2014).

10. 10.  
Rahbek, C. et al. Humboldt's enigma: what causes global patterns of mountain biodiversity? *Science* **365**, 1108–1113 (2019).
11. 11.  
Pan, Y. et al. A large and persistent carbon sink in the world's forests. *Science* **333**, 988–993 (2011).
12. 12.  
Booth, B. B. B. et al. High sensitivity of future global warming to land carbon cycle processes. *Environ. Res. Lett.* **7**, 024002 (2012).
13. 13.  
Hubau, W. et al. Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature* **579**, 80–87 (2020).
14. 14.  
Feldpausch, T. R. et al. Tree height integrated into pantropical forest biomass estimates. *Biogeosciences* **9**, 3381–3403 (2012).
15. 15.  
Bastin, J.-F. et al. Pan-tropical prediction of forest structure from the largest trees. *Glob. Ecol. Biogeogr.* **27**, 1366–1383 (2018).
16. 16.  
*CCI BIOMASS Product User Guide Year 1 Version 1.0*, [https://climate.esa.int/sites/default/files/biomass\\_D4.3\\_Product\\_User\\_Guide\\_V1.0.pdf](https://climate.esa.int/sites/default/files/biomass_D4.3_Product_User_Guide_V1.0.pdf) (Aberystwyth University and GAMMA Remote Sensing, 2019).
17. 17.  
Lefsky, M. A., Keller, M., Pang, Y., de Camargo, P. & Hunter, M. O. Revised method for forest canopy height estimation from the Geoscience Laser Altimeter System waveforms. *J. Appl. Remote Sens.* **1**, 013537 (2007).
18. 18.

Willcock, S. et al. Quantifying and understanding carbon storage and sequestration within the Eastern Arc Mountains of Tanzania, a tropical biodiversity hotspot. *Carbon Balance Manag.* **9**, 2 (2014); correction **12**, 20 (2017).

19. 19.

Bussmann, R. W. Vegetation zonation and nomenclature of African Mountains —an overview. *Lyonia* **11**, 41–66 (2006).

20. 20.

Hamilton, A. Vegetation, climate and soil, altitudinal relationships on the East Usambara Mountains of Tanzania. *J. East Afr. Nat. Hist.* **87**, 85–89 (1998).

21. 21.

Phillips, O., Baker, T., Brienen, R. & Feldpausch, T. *RAINFOR Field Manual for Plot Establishment and Remeasurement*, [http://www.rainfor.org/upload/ManualsEnglish/RAINFOR\\_field\\_manual\\_version\\_2016.pdf](http://www.rainfor.org/upload/ManualsEnglish/RAINFOR_field_manual_version_2016.pdf) (Univ. Leeds, 2016)

22. 22.

Jarvis, A. & Mulligan, M. The climate of cloud forests. *Hydrol. Process.* **25**, 327–343 (2011).

23. 23.

Platts, P. J., Omeny, P. A. & Marchant, R. AFRICLIM: high-resolution climate projections for ecological applications in Africa. *Afr. J. Ecol.* **53**, 103–108 (2015).

24. 24.

McInerny, G. J. & Purves, D. W. Fine-scale environmental variation in species distribution modelling: regression dilution, latent variables and neighbourly advice. *Methods Ecol. Evol.* **2**, 248–257 (2011).

25. 25.

Poulsen, J. R. et al. Ecological consequences of forest elephant declines for Afrotropical forests. *Conserv. Biol.* **32**, 559–567 (2018).

26. 26.
- Berzaghi, F. et al. Carbon stocks in central African forests enhanced by elephant disturbance. *Nat. Geosci.* **12**, 725–729 (2019); correction **12**, 1050 (2019).
27. 27.
- Enquist, B. J. et al. The megabiota are disproportionately important for biosphere functioning. *Nat. Commun.* **11**, 699 (2020).
28. 28.
- Lin, T.-C., Hogan, J. A. & Chang, C. T. Tropical cyclone ecology: a scale-link perspective. *Trends Ecol. Evol.* **35**, 594–604 (2020).
29. 29.
- Favalli, M. et al. Lava flow hazard and risk at Mt. Cameroon volcano. *Bull. Volcanol.* **74**, 423–439 (2012).
30. 30.
- Stanley, T. & Kirschbaum, D. B. A heuristic approach to global landslide susceptibility mapping. *Nat. Hazards* **87**, 145–164 (2017).
31. 31.
- Lovett, J. C. Elevational and latitudinal changes in tree associations and diversity in the Eastern Arc mountains of Tanzania. *J. Trop. Ecol.* **12**, 629–650 (1996).
32. 32.
- Hemp, A. et al. Africa’s highest mountain harbours Africa’s tallest trees. *Biodivers. Conserv.* **26**, 103–113 (2017).
33. 33.
- Culmsee, H., Leuschner, C., Moser, G. & Pitopang, R. Forest aboveground biomass along an elevational transect in Sulawesi, Indonesia, and the role of Fagaceae in tropical montane rain forests. *J. Biogeogr.* **37**, 960–974 (2010).
34. 34.

Enright, N. J. & Ogden, J. in *Ecology of the Southern Conifers* (eds Enright, N. J. & Hill, R. S.) 271–287 (Melbourne Univ. Press, 1995).

35. 35.

Neale, D. B. & Wheeler, N. C. in *The Conifers: Genomes, Variation and Evolution* (eds Neale, D. B. & Wheeler, N. C.) 1–21 (Springer, 2019).

36. 36.

Mill, R. R. Towards a biogeography of the Podocarpaceae. *Acta Hortic.* **615**, 137–147 (2003).

37. 37.

Helmer, E. H. et al. Neotropical cloud forests and paramo to contract and dry from declines in cloud immersion and frost. *PLoS ONE* **14**, e0213155 (2019).

38. 38.

Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

39. 39.

Turubanova, S., Potapov, P., Tyukavina, A. & Hansen, M. Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. *Environ. Res. Lett.* **13**, 074028 (2018).

40. 40.

Pellikka, P. K. E., Lötjönen, M., Siljander, M. & Lens, L. Airborne remote sensing of spatiotemporal change (1955–2004) in indigenous and exotic forest cover in the Taita Hills, Kenya. *Int. J. Appl. Earth Obs. Geoinf.* **11**, 221–232 (2009).

41. 41.

Zeng, Z. et al. Deforestation-induced warming over tropical mountain regions regulated by elevation. *Nat. Geosci.* **14**, 23–29 (2021).

42. 42.

Spira, C., Kirkby, A., Kujirakwinja, D. & Plumptre, A. J. The socio-economics of artisanal mining and bushmeat hunting around protected areas: Kahuzi–Biega National Park and Itombwe nature reserve, eastern Democratic Republic of Congo. *Oryx* **53**, 136–144 (2019).

43. 43.

Bebber, D. P. & Butt, N. Tropical protected areas reduced deforestation carbon emissions by one third from 2000–2012. *Sci. Rep.* **7**, 14005 (2017); **8**, 14845 (2018).

44. 44.

Tegha, K. C. & Sendze, Y. G. Soil organic carbon stocks in Mt Cameroon National Park under different land uses. *J. Ecol. Nat. Environ.* **8**, 20–30 (2016).

45. 45.

Munishi, P. K. T. & Shear, T. H. Carbon storage in afromontane rain forests of the eastern arc mountains of Tanzania: their net contribution to atmospheric carbon. *J. Trop. For. Sci.* **16**, 78–98 (2004).

46. 46.

Wheeler, C. E. et al. Carbon sequestration and biodiversity following 18 years of active tropical forest restoration. *For. Ecol. Manage.* **373**, 44–55 (2016).

47. 47.

Avitabile, V., Baccini, A., Friedl, M. A. & Schmullius, C. Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. *Remote Sens. Environ.* **117**, 366–380 (2012).

48. 48.

Aneseyee, B. A., Soromessa, T. & Belliethathan, S. Carbon stock of Gambella National Park: implication for climate change mitigation. *Int. J. Adv. Life Sci.* **35**, 41–56 (2015).

49. 49.

Lisboa, S. N. et al. Biomass allometric equation and expansion factor for a mountain moist evergreen forest in Mozambique. *Carbon Balance Manag.* **13**,

23 (2018).

50. 50.

Lopez-Gonzalez, G., Lewis, S. L., Burkitt, M. & Phillips, O. L. ForestPlots.net: a web application and research tool to manage and analyse tropical forest plot data. *J. Veg. Sci.* **22**, 610–613 (2011).

51. 51.

Lopez-Gonzalez, G., Lewis, S. L., Burkitt, M., Baker, T. R. & Phillips, O. L. *ForestPlots.net Database*, <http://www.forestplots.net> (ForestPlots, 2009).

52. 52.

Cavanaugh, K. et al. Carbon storage in tropical forests correlates with taxonomic diversity and functional dominance on a global scale. *Glob. Ecol. Biogeogr.* **23**, 563–573 (2014).

53. 53.

Lewis, S. L. et al. 2009 Increasing carbon storage in intact African tropical forests. *Nature* **457**, 1003–1006 (2009).

54. 54.

Chave, J. et al. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Change Biol.* **20**, 3177–3190 (2014).

55. 55.

Martin, A. R., Doraisami, M. & Thomas, S. C. Global patterns in wood carbon concentration across the world's trees and forests. *Nat. Geosci.* **11**, 915–920 (2018).

56. 56.

Chave, J. et al. Towards a worldwide wood economics spectrum. *Ecol. Lett.* **12**, 351–366 (2009).

57. 57.

Zanne, A. E. et al. Towards a worldwide wood economics spectrum. *Dryad* <https://doi.org/10.5061/dryad.234> (2009).

58. 58.

Clark, D. A. et al. Net primary production in tropical forests: an evaluation and synthesis of existing field data. *Ecol. Appl.* **11**, 371–384 (2001).

59. 59.

Paul, T. S. H., Kimberley, M. O. & Beets, P. N. Thinking outside the square: evidence that plot shape and layout in forest inventories can bias estimates of stand metrics. *Methods Ecol. Evol.* **10**, 381–388 (2019).

60. 60.

Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).

61. 61.

Wilson, A. M. & Jetz, W. Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. *PLoS Biol.* **14**, e1002415 (2016).

62. 62.

Albrecht, R., Goodman, S., Buechler, D., Blakeslee R. & Christian, H. *LIS 0.1 Degree Very High Resolution Gridded Lightning Climatology Data Collection*, <https://ghrc.nsstc.nasa.gov/pub/lis/climatology/LIS/> (NASA Global Hydrology Resource Center, 2016).

63. 63.

Hengl, T. et al. SoilGrids250m: global gridded soil information based on machine learning. *PLoS ONE* **12**, e0169748 (2017).

64. 64.

Harris, N. L. et al. Global maps of twenty-first century forest carbon fluxes. *Nat. Clim. Chang.* **11**, 234–240 (2021).

65. 65.

Santoro, M. & Cartus, O. *ESA Biomass Climate Change Initiative (Biomass\_cci): Global Datasets of Forest Above-ground Biomass for the year*

2017, v1 (Centre for Environmental Data Analysis, 2019).

66. 66.

Saatchi, S. et al. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proc. Natl Acad. Sci. USA* **108**, 9899–9904 (2011).

67. 67.

Avitabile, V. et al. An integrated pan-tropical biomass map using multiple reference datasets. *Glob. Change Biol.* **22**, 1406–1420 (2016).

68. 68.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team nlme: linear and nonlinear mixed effects models. R package version 3.1-151 (2020).

69. 69.

Bates, D., Maechler, M., Bolker B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

70. 70.

Lüdecke, D., Ben-Shachar, M., Patil, I. & Makowski, D. Parameters: extracting, computing and exploring the parameters of statistical models using R. *J. Open Source Softw.* **5**, 2445 (2020).

71. 71.

Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. *Biom. J.* **50**, 346–363 (2008).

72. 72.

Barton, K. MuMIn: multi-model inference. R package version 1.43.17 (2020)

## Acknowledgements

We thank the people of the many villages and local communities who welcomed our field teams and became our field assistants, without whose support the AfriMont dataset would not have been possible. Cameroon: villages Elak-Oku, Bokwoango, Bakingili, Muandelengoh, Enyandong, Ekangmbeng, Ngalmoa, Dikome Balue,

Muyange, Matamani; assistants E. Ndive, D. Wultof, F. Keming, E. Bafon, J. Meyeh, T. K. Konsum, J. Esembe, F. Luma, F. Teke, E. E. Dagobert, E. D. Ndode, N. F. Njikang; Democratic Republic of the Congo: Bunyakiri, J. Kalume, W. Gului, D. Cirhagaga, B. Mugisho. Kenya: assistants A. M. Aide, H. Lerapo, J. Harugura, R. A. Wamuro, J. Lekatap, L. Lemooli, D. Kimuzi, B. M. Lombo, J. Broas, J. Hietanen, V. Heikinheimo, E. Schäfer. Rwanda: assistants I. Rusizana, P. Niyontegereje, J. B. Gakima, F. Ngayabahiga. Tanzania: TEAM staff and affiliates. Uganda: K. Laughlin, X. Mugumya, L. Etwodu, M. Mugisa. For logistical and administrative support, we are indebted to international, national and local institutions: SOPISDEW, Mt Cameroon National Park, Tropical Plant Exploration Group (TroPEG), Institut Congolais de Conservation de la Nature, Kahuzi-Biega National Park, Itombwe Nature Reserve, NEMA Marsabit Office, Taita Research Station, Kenya Forest Service, Rwanda Development Board, Nyungwe National Park, Conservation International, the Smithsonian Institution, Wildlife Conservation Society, Sokoine University of Agriculture, Tanzania Wildlife Research Institute, Tanzania National Parks Authority, Kilimanjaro National Park, Tanzania Commission for Science and Technology, Royal Zoological Society of Scotland, Uganda Wildlife Authority, Makerere University Biological Field Station, Uganda National Forestry Authority and Uganda National Council for Science and Technology. Field campaigns for AfriMont were funded by Marie Skłodowska-Curie Actions Intra-European Fellowships (number 328075) and Global Fellowships (number 74356), National Geographic Explorer (NGS-53344R-18), Czech Science Foundation (number 21-17125S), Rufford Small Grant Foundation (16712-B, 19476-D), Ministry of Foreign Affairs of Finland (BIODEV project), the Academy of Finland (number 318645), Swedish International Development Cooperation Agency, the Leverhulme Trust, the Strategic Research Area Biodiversity and Ecosystem Services in a Changing Climate, the German Research Foundation (DFG), Gatsby Plants, Natural Science and Engineering Research Council of Canada and International Development Research Centre of Canada. This paper is also a product of the AfriTRON network, for which we are indebted to hundreds of institutions, field assistants and local communities for establishing and maintaining the plots, including the Forestry Development Authority of the Government of Liberia, the University of Liberia, University of Ibadan (Nigeria), the University of Abeokuta (Nigeria), the University of Yaounde I (Cameroon), the National Herbarium of Yaounde (Cameroon), the University of Buea (Cameroon), Bioversity International (Cameroon), Salonga National Park (Democratic Republic of Congo), The Centre de Formation et de Recherche en Conservation Forestière (CEFRECOF, Epulu, Democratic Republic of Congo), the Institut National pour l'Étude et la Recherche Agronomiques (INERA, Democratic Republic of Congo), the École Régionale Postuniversitaire d'Aménagement et de Gestion intégrés des Forêts et Territoires tropicaux (ERAIFT Kinshasa, Democratic Republic of Congo), WWF-Democratic Republic of Congo,

WCS-Democratic Republic of Congo, the Université de Kisangani (Democratic Republic of Congo), Université Officielle de Bukavu (Democratic Republic of Congo), Université de Mbujimayi (Democratic Republic of Congo), le Ministère de l'Environnement et Développement Durable (Democratic Republic of Congo), the FORETS project in Yangambi (CIFOR, CGIAR and the European Union; Democratic Republic of Congo), the Lukuru Wildlife Research Foundation (Democratic Republic of Congo), Mbarara University of Science and Technology (MUST, Uganda), WCS-Uganda, the Uganda Forest Department, the Commission of Central African Forests (COMIFAC), the Udzungwa Ecological Monitoring Centre (Tanzania) and the Sokoine University of Agriculture (Tanzania). The AfriTRON network has been supported by the European Research Council (291585, ‘T-FORCES’ – Tropical Forests in the Changing Earth System, Advanced Grant to O.L.P. and S.L.L.), the Gordon and Betty Moore Foundation, the David and Lucile Packard Foundation and the European Union’s Seventh Framework Programme (283080, ‘GEOCARBON’). We are grateful to A. Daniels, F. Mbayu, T. R. Feldpausch, E. Kearsley, J. Lloyd, R. Lowe, J. Mukinzi, L. Ojo, A. T. Peterson, J. Talbot and L. Zemagho for giving us access to their plot data. We also thank C. Chatelain (Geneva Botanic Gardens) for access to the African Plants Database and to H. Tang for helping to explore the use of GEDI data. Data from AfriTRON and most of AfriMont are stored and curated by ForestPlots.net, a long-term cyberinfrastructure initiative hosted at the University of Leeds that unites permanent plot records and their contributing scientists from the world’s tropical forests. The development of ForestPlots.net and curation of African data have been funded by many sources, including the ERC (principally from AdG 291585 ‘T-FORCES’), the UK Natural Environment Research Council (including NE/B503384/1, NE/F005806/1, NE/P008755/1, NE/N012542/1 and NE/I028122/1), the Gordon and Betty Moore Foundation (‘RAINFOR’, ‘MonANPeru’), the EU Horizon programme (especially ‘GEOCARBON’, ‘Amazalert’) and the Royal Society (University Research Fellowship to S.L.L.).

## Author information

### Affiliations

1. Department of Environment and Geography, University of York, York, UK  
Aida Cuni-Sanchez, Philip J. Platts, Rob Marchant & Andrew Marshall
2. Department of International Environmental and Development Studies (NORAGRIC), Norwegian University of Life Sciences, Ås, Norway

Aida Cuni-Sanchez

3. Department of Natural Sciences, Manchester Metropolitan University, Manchester, UK

Martin J. P. Sullivan

4. School of Geography, University of Leeds, Leeds, UK

Martin J. P. Sullivan, Simon L. Lewis, Serge K. Begne, Amy C. Bennett, Martin Gilpin, Jon Lovett & Oliver L. Phillips

5. Leverhulme Centre for Anthropocene Biodiversity, University of York, York, UK

Philip J. Platts

6. Climate Change Specialist Group, Species Survival Commission, International Union for Conservation of Nature, Gland, Switzerland

Philip J. Platts

7. Department of Geography, University College London, London, UK

Simon L. Lewis

8. Biology Department, Université Officielle de Bukavu, Bukavu, Democratic Republic of the Congo

Gérard Imani & Christian Amani

9. Service of Wood Biology, Royal Museum for Central Africa, Tervuren, Belgium

Wannes Hubau, Hans Beeckman & John T. Mukendi

10. Department of Environment, Laboratory of Wood Technology (Woodlab), Ghent University, Ghent, Belgium

Wannes Hubau

11. University of Jos, Jos, Nigeria

Iveren Abiem

12. Nigerian Montane Forest Project, Yelwa Village, Nigeria

Iveren Abiem & Hazel Chapman

13. Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland

Hari Adhikari, Janne Heiskanen & Petri Pellikka

14. Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic

Tomas Albrecht

15. Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic

Tomas Albrecht

16. Institute of Botany of the Czech Academy of Science, Třeboň, Czech Republic

Jan Altman & Jiri Dolezal

17. College of Natural and Computational Science, Addis Ababa University, Addis Ababa, Ethiopia

Abreham B. Aneseyee & Teshome Soromessa

18. Department of Natural Resource Management, College of Agriculture and Natural Resource, Wolkite University, Wolkite, Ethiopia

Abreham B. Aneseyee

19. European Commission, Joint Research Centre, Ispra, Italy

Valerio Avitabile

20. UK Centre for Ecology and Hydrology, Edinburgh, UK

Lindsay Banin

21. Université du Cinquantenaire Lwiro, Département de sciences de l'environnement, Kabare, Democratic Republic of the Congo

Rodrigue Batumike

22. Isotope Bioscience Laboratory (ISOFYS), Ghent University, Ghent, Belgium

Marijn Bauters, Pascal Boeckx & Joseph Okello

23. Plant Systematic and Ecology Laboratory, Higher Teachers' Training College, University of Yaoundé I, Yaoundé, Cameroon

Serge K. Begne, Vincent Droissart, Marie-Noel Kamdem, Murielle Simo-Droissart & Bonaventure Sonké

24. Institute of Tropical Forest Conservation, Mbarara University of Science and Technology, Mbarara, Uganda

Robert Bitariho

25. Biodiversity and Landscape Unit, Gembloux Agro-Bio Tech, Université de Liège, Liège, Belgium

Jan Bogaert

26. Institute for Geography, Friedrich Alexander University, Erlangen–Nuremberg, Germany

Achim Bräuning & Ulrike Hiltner

27. Département de Eaux et Forêts, Institut Supérieur d'Agroforesterie et de Gestion de l'Environnement de Kahuzi-Biega (ISAGE-KB), Kalehe, Democratic Republic of the Congo

Franklin Bulonvu

28. UN Environment World Conservation Monitoring Center (UNEP-WCMC), Cambridge, UK

Neil D. Burgess

29. Computational and Applied Vegetation Ecology (CAVElab), Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium

Kim Calders & Hans Verbeeck

30. Department of Anthropology, George Washington University, Washington DC, USA

Colin Chapman

31. School of Life Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa

Colin Chapman

32. Shaanxi Key Laboratory for Animal Conservation, Northwest University, Xi'an, China

Colin Chapman

33. International Centre of Biodiversity and Primate Conservation, Dali University, Dali, China

Colin Chapman

34. University of Canterbury, Canterbury, New Zealand

Hazel Chapman

35. Inventory and Monitoring Program, National Park Service, Fredericksburg, VA, USA

James Comiskey

36. University of Ghent, Ghent, Belgium

Thales de Haulleville

37. World Agroforestry (ICRAF), Nairobi, Kenya

Mathieu Decuyper

38. Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, Wageningen, The Netherlands

Mathieu Decuyper & Martin Herold

39. Geography, Environment and Geomatics, University of Guelph, Guelph, Ontario, Canada

Ben DeVries

40. Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

Jiri Dolezal

41. AMAP Lab, Université de Montpellier, IRD, CNRS, INRAE, CIRAD, Montpellier, France

Vincent Droissart

42. Faculté de Gestion de Ressources Naturelles Renouvelables, Université de Kisangani, Kisangani, Democratic Republic of the Congo

Corneille Ewango & Janvier Lisingo

43. College of Development Studies, Addis Ababa University, Addis Ababa, Ethiopia

Senbeta Feyera

44. Dendrochronology Laboratory, World Agroforestry Centre (ICRAF), Nairobi, Kenya

Aster Gebrekirstos

45. Missouri Botanical Garden, St Louis, MO, USA

Roy Gereau

46. Department of Biology, University of Burundi, Bujumbura, Burundi

Dismas Hakizimana

47. Smithsonian Institution Forest Global Earth Observatory (ForestGEO), Smithsonian Tropical Research Institute, Washington DC, USA

Jefferson Hall & David Kenfack

48. Kunming Institute of Botany, Kunming, China

Alan Hamilton

49. Université Libre de Bruxelles, Brussels, Belgium

Olivier Hardy

50. Division of Vertebrate Zoology, Yale Peabody Museum of Natural History, New Haven, CT, USA

Terese Hart

51. Institute for Atmospheric and Earth System Research, Faculty of Science, University of Helsinki, Helsinki, Finland

Janne Heiskanen

52. Department of Plant Systematics, University of Bayreuth, Bayreuth, Germany

Andreas Hemp

53. Helmholtz Center Potsdam GFZ German Research Centre for Geosciences, Section 1.4 Remote Sensing and Geoinformatics, Potsdam, Germany

Martin Herold

54. Helmholtz-Centre for Environmental Research (UFZ), Leipzig, Germany

Ulrike Hiltner

55. Department of Ecology, Faculty of Science, Charles University, Prague, Czech Republic

David Horak & Ondrej Sedlacek

56. International Gorilla Conservation Programme, Musanze, Rwanda

Charles Kayijamahe & Eustate Uzabaho

57. Department of Natural Resources, Karatina University, Karatina, Kenya

Mwangi J. Kinyanjui

58. Department of Ecosystem Science and Sustainability, Colorado State University, Fort Collins, CO, USA

Julia Klein

59. Eco2librium LLC, Boise, ID, USA

Mark Lung

60. Department of Ecology, Université de Kisangani, Kisangani, Democratic Republic of the Congo

Jean-Remy Makana

61. Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK

Yadvinder Malhi

62. Tropical Forests and People Research Centre, University of the Sunshine Coast, Sippy Downs, Queensland, Australia

Andrew Marshall & Alain S. K. Ngute

63. Flamingo Land Ltd, Malton, UK

Andrew Marshall

64. College of African Wildlife Management, Mweka, Tanzania

Emanuel H. Martin

65. School of GeoSciences, University of Edinburgh, Edinburgh, UK

Edward T. A. Mitchard & Charlotte Wheeler

66. Department of Geography and Environmental Sciences, University of Dundee, Dundee, UK

Alexandra Morel

67. Independent Botanist, Harare, Zimbabwe

Tom Muller

68. Department of Horticultural Sciences, Faculty of Applied Sciences, Cape Peninsula University of Technology, Bellville, South Africa

Felix Nchu

69. Biology Department, University of Rwanda, Kigali, Rwanda

Brigitte Nyirambangutse & Etienne Zibera

70. Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden

Brigitte Nyirambangutse & Göran Wallin

71. Mountains of the Moon University, Fort Portal, Uganda

Joseph Okello

72. National Agricultural Research Organisation, Mbarara Zonal Agricultural Research and Development Institute, Mbarara, Uganda

Joseph Okello

73. School of Biological Sciences, University of Southampton, Southampton, UK

Kelvin S.-H. Peh

74. Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge, UK

Kelvin S.-H. Peh

75. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

Petri Pellikka

76. Key Biodiversity Areas Secretariat, BirdLife International, Cambridge, UK

Andrew Plumptre

77. School of Life Sciences, University of Lincoln, Lincoln, UK

Lan Qie

78. Department of Biology, University of Florence, Sesto Fiorentino, Italy

Francesco Rovero

79. Tropical Biodiversity Section, Museo delle Scienze, Trento, Italy

Francesco Rovero

80. Tropical Plant Exploration Group (TroPEG), Mundemba, Cameroon

Moses N. Sainge

81. Center for Development Research (ZEF), University of Bonn, Bonn, Germany

Christine B. Schmitt

82. Conservation and Landscape Ecology, University of Freiburg, Freiburg, Germany

Christine B. Schmitt

83. Applied Biology and Ecology Research Unit, University of Dschang, Dschang, Cameroon

Alain S. K. Ngute

84. Forest Ecology and Forest Management Group, Wageningen University, Wageningen, The Netherlands

Douglas Sheil

85. Water and Land Resources Center, Addis Ababa University, Addis Ababa, Ethiopia

Demisse Sheleme

86. African Wildlife Foundation (AWF), Biodiversity Conservation and Landscape Management Program, Simien Mountains National Park, Debark, Ethiopia

Tibebu Y. Simegn

87. Faculty of Forestry, University of British Columbia, Vancouver, British Columbia, Canada

Terry Sunderland

88. Center for International Forestry Research (CIFOR), Bogor, Indonesia

Terry Sunderland

89. Department of Forest Ecology, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences, Prague, Czech Republic

Miroslav Svoboda

90. Department of Plant Biology, Faculty of Sciences, University of Yaoundé I, Yaoundé, Cameroon

Hermann Taedoumg

91. Bioversity International, Yaoundé, Cameroon

Hermann Taedoumg

92. UK Research and Innovation, London, UK

James Taplin

93. Department of Geography, National University of Singapore, Singapore, Singapore

David Taylor

94. Institute of Forestry and Conservation, University of Toronto, Toronto, Ontario, Canada

Sean C. Thomas

95. Biodiversity Foundation for Africa, East Dean, UK

Jonathan Timberlake

96. Forestry Development Authority of the Government of Liberia (FDA), Monrovia, Liberia

Darlington Tuagben

97. School of Forestry and Environmental Studies, Yale University, New Haven, CT, USA

Peter Umunay

98. Department of Biological Sciences, Florida International University, Miami, FL, USA

Jason Vleminckx

99. School of Natural Sciences, University of Bangor, Bangor, UK

Simon Willcock

100. Rothamsted Research, Harpenden, UK

Simon Willcock

101. University of Liberia, Monrovia, Liberia

John T. Woods

## Contributions

A.C.-S. conceived the study and assembled the AfriMont dataset. A.C.-S. and M.J.P.S. analysed the plot data (with contributions from S.L.L.) and wrote the manuscript. P.J.P. analysed forest extents and contributed to writing. S.L.L. conceived and managed the AfriTRON forest plot recensus programme. E.T.A.M. and V.A. helped compare plot data with remote sensing carbon maps. All co-authors read and approved the manuscript.

## Corresponding author

Correspondence to [Aida Cuni-Sanchez](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Nicolas Barbier and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

### Extended Data Fig. 1 Sensitivity of mean AGC stock estimates to data subsampling.

AfriMont plot data were resampled at different sample sizes either at plot level (sampling with replacement) or at site level (sampling without replacement).  $N = 1,000$  resamples for each sample size.

### Extended Data Fig. 2 Effect of plot area, aggregation procedure and plot design on estimates of AGC stocks across the AfriMont plot network.

**a**, Relationship between AGC stocks and plot area of plots before aggregation. The red line shows the fit of a locally weighted regression model (span 0.75) relating these variables, with dashed lines showing the standard errors. **b**, Variation in AGC stocks using either all plots before aggregation (unaggregated), plots before aggregation but excluding those  $<0.2$  ha (unaggregated,  $>0.2$  ha) or the aggregated plots used in the main analyses (aggregated). **c**, Effects of plot design on AGC stocks (each site represents one dot). Sampling strategies include random or stratified random, plots positioned along transects, plots established within elevation bands, subjective measures such as choosing an area of forest considered representative of the wider area, and other strategies (one plot sampled per site or unclear strategy). Carbon stocks (log transformed) did not differ significantly between sites with different sampling strategies (analysis of variance  $F_{4,39} = 0.432$ ,  $P = 0.785$ ). For specific site information, see Supplementary Table 5.

### Extended Data Fig. 3 Robustness of differences in tropical montane forest AGC stocks among continents based on plot networks to differences in elevation.

**a**, Elevations of montane forests plots sampled in each continent. Violin plots show the distribution of data, with boxplots showing the median and interquartile range of elevation in each continent. **b**, Effect of removing submontane plots (800–1,000 m a.s.l.) and high elevation plots ( $>2,200$  m a.s.l., approximately the upper quartile of elevations for the African montane plot dataset) on AGC stocks in montane forests sampled by plot networks in each continent. Mean AGC stocks and

95% CIs are shown as estimated by models using all data, excluding plots 800–1,000 m and restricting plots to 1,000–2,200 m. Means for all plots differ from the analysis in Fig. 1 as literature plots without elevation data (plots in Colombia) were excluded from this analysis. Point symbols are proportional to the square-root plot area.  $N = 324$  plots.

**Extended Data Fig. 4 Relationship between AGC stocks and elevation for tropical montane forests in each continent based on plot networks.**

The dashed lines show relationships from a linear mixed-effects model of log-transformed AGC stocks as a function of elevation, continent and their interaction. Site was included as a random effect, and AGC stock–elevation relationships allowed to vary among sites. The lines show fitted slopes across sites. Neither the overall relationship between elevation and AGC stocks (slope  $-0.039$  [95% CI =  $-0.127$ – $0.057$ ],  $P = 0.420$ ) nor interactions between elevation and continent (Southeast Asia, change in slope =  $-0.074$  [ $-0.294$ – $0.149$ ],  $P = 0.503$ ; Neotropics, change in slope  $0.006$  [ $-0.132$ – $0.149$ ],  $P = 0.913$ ) are statistically significant.  $N = 324$  plots.

**Extended Data Fig. 5 Environmental drivers of AGC stocks across the AfriMont plot network.**

Coefficients are from a linear mixed-effects model with site as a random intercept. Results are following all-subsets regression and model averaging, in which variables that do not appear in well supported models are given coefficients of zero, leading to shrinkage in model coefficients. Statistically significant relationships ( $P < 0.05$ ) are indicated with asterisks. TPI refers to topographic position index (positive values indicate higher than surroundings and negative values indicate lower than surroundings).  $T_{\text{mean}}$ , annual mean temperature;  $T_{\text{seasonality}}$ , temperature seasonality;  $\text{Precip}_{\text{total}}$ , annual precipitation;  $\text{Precip}_{\text{seasonality}}$ , precipitation seasonality.

**Extended Data Fig. 6 Expected sampling effort if effort was distributed in proportion to the area of tropical montane forest biome in Africa.**

Data are summarized at  $1^{\circ}$  resolution. The upward triangles show grid cells where AfriMont sampling effort is more than double expected effort and the downward triangles show grid cells where AfriMont sampling effort is less than half expected effort. The circles denote AfriMont sampling effort being between half and double expected effort. The extent of the tropical montane forest biome was defined as

closed-canopy forests  $\geq 800$  m a.s.l. in December 2018, extracted from ref. [38](#) and clipped to ‘primary humid forest’ using ref. [39](#). This gridded map differs from Fig. [4](#) as numerous grids have very little tropical montane forest.

**Extended Data Fig. 7 Differences in the environmental conditions sampled by the AfriMont plot network and the tropical montane forest biome in Africa.**

The extent of the biome was defined as closed-canopy forests  $\geq 800$  m a.s.l. in December 2018, extracted from ref. [38](#) and clipped to ‘primary humid forest’ using ref. [39](#). Environmental variables for the biome were extracted at about 1-km resolution.

**Extended Data Fig. 8 Differences in AGC stocks in AfriMont plots located in montane forests with and without elephants.**

**a**, Differences across all plots. AGC stocks are statistically significantly lower in forests with elephants ( $t$ -test,  $t = 3.5$ , d.f. = 83.5,  $P = 0.001$ ). The thick line shows the median, and boxes cover the interquartile range (IQR). Values  $>1.5$  times IQR away from the IQR are shown by points. **b**, Differences in countries where elephants are present in at least one of the montane sites studied. The black squares show means in each country in forests with or without elephants and the solid lines denote statistically significant differences ( $t$ -tests,  $P < 0.05$ ). Elephant presence in 2019 was estimated by the co-authors (Supplementary Table [5](#)).

**Extended Data Fig. 9 Relationship between AGC stocks and Podocarpaceae.**

**a**, Relationship between AGC stocks and Podocarpaceae basal area across plots in the AfriMont network, expressed as a percentage of total plot basal area. These variables are not significantly correlated ( $r_s = 0.083$ ,  $n = 226$ ,  $P = 0.212$ ). **b**, Distribution of plots with at least 20% basal area of Podocarpaceae (black points) in relation to elevation and AGC stocks. AGC stocks are not significantly related to elevation or Podocarpaceae basal area (linear mixed effects model,  $P = 0.152$  and  $P = 0.132$ , respectively).

**Extended Data Fig. 10 Within-country variation in AGC stocks based on the AfriMont plot network.**

Error bars show means and 95% CIs estimated by linear mixed-effects models. Modelled means not shown for countries with fewer than five plots. Point size is proportional to plot area.

## Supplementary information

### Supplementary Information

This file contains Supplementary Tables S1 – S7.

### Peer Review File

## Rights and permissions

### Reprints and Permissions

## About this article

### Cite this article

Cuni-Sanchez, A., Sullivan, M.J.P., Platts, P.J. *et al.* High aboveground carbon stock of African tropical montane forests. *Nature* **596**, 536–542 (2021).  
<https://doi.org/10.1038/s41586-021-03728-4>

- Received: 20 December 2020
- Accepted: 14 June 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03728-4>

### African tropical montane forests store more carbon than was thought

- Nicolas Barbier

News & Views 25 Aug 2021

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

- Article
- Open Access
- [Published: 25 August 2021](#)

# Genome of a middle Holocene hunter-gatherer from Wallacea

- [Selina Carlhoff](#) [ORCID: orcid.org/0000-0001-9118-2839<sup>1,2 na1</sup>](#),
- [Akin Duli<sup>3 na1</sup>](#),
- [Kathrin Nägele](#) [ORCID: orcid.org/0000-0003-3861-8677<sup>1,2</sup>](#),
- [Muhammad Nur<sup>3</sup>](#),
- [Laurits Skov<sup>2</sup>](#),
- [Iwan Sumantri<sup>3</sup>](#),
- [Adhi Agus Oktaviana<sup>4,5</sup>](#),
- [Budianto Hakim<sup>6</sup>](#),
- [Basran Burhan](#) [ORCID: orcid.org/0000-0002-9172-5888<sup>7</sup>](#),
- [Fardi Ali Syahdar<sup>8</sup>](#),
- [David P. McGahan](#) [ORCID: orcid.org/0000-0002-3161-1512<sup>7</sup>](#),
- [David Bulbeck<sup>9</sup>](#),
- [Yinika L. Perston](#) [ORCID: orcid.org/0000-0002-8043-6466<sup>7</sup>](#),
- [Kim Newman](#) [ORCID: orcid.org/0000-0002-6840-1658<sup>7</sup>](#),
- [Andi Muhammad Saiful<sup>6</sup>](#),
- [Marlon Ririmasse<sup>4</sup>](#),
- [Stephen Chia<sup>10</sup>](#),
- [Hasanuddin<sup>6</sup>](#),
- [Dwia Aries Tina Pulubuhu<sup>11</sup>](#),
- [Suryatman<sup>6</sup>](#),
- [Supriadi<sup>3</sup>](#),
- [Choongwon Jeong](#) [ORCID: orcid.org/0000-0003-3049-2352<sup>12</sup>](#),
- [Benjamin M. Peter](#) [ORCID: orcid.org/0000-0003-2526-8081<sup>2</sup>](#),
- [Kay Prüfer](#) [ORCID: orcid.org/0000-0001-6242-3058<sup>1,2</sup>](#),
- [Adam Powell<sup>2</sup>](#),
- [Johannes Krause](#) [ORCID: orcid.org/0000-0001-9144-3920<sup>1,2 na2</sup>](#),
- [Cosimo Posth](#) [ORCID: orcid.org/0000-0002-8206-3907<sup>1,13,14 na2</sup>](#) &
- [Adam Brumm](#) [ORCID: orcid.org/0000-0002-2276-3258<sup>7 na2</sup>](#)

*Nature* volume 596, pages 543–547 (2021)

- 32k Accesses
- 980 Altmetric
- [Metrics details](#)

## Subjects

- [Archaeology](#)
- [Evolutionary genetics](#)
- [Genetic variation](#)
- [Population genetics](#)

## Abstract

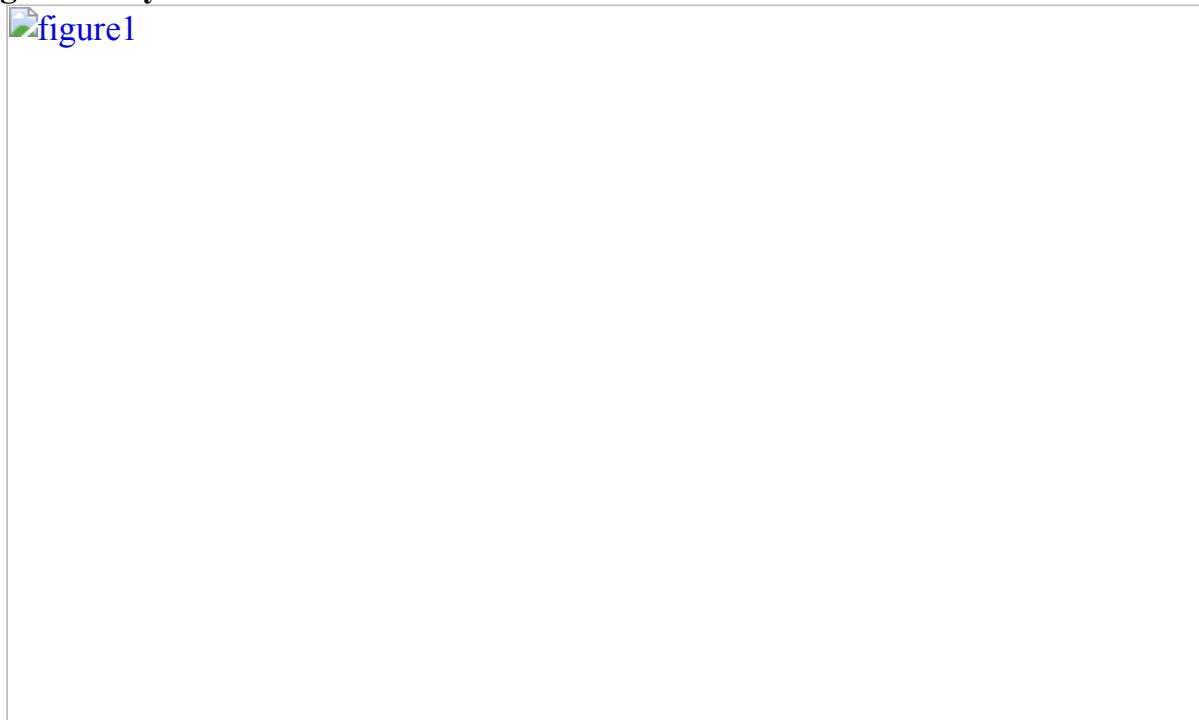
Much remains unknown about the population history of early modern humans in southeast Asia, where the archaeological record is sparse and the tropical climate is inimical to the preservation of ancient human DNA<sup>1</sup>. So far, only two low-coverage pre-Neolithic human genomes have been sequenced from this region. Both are from mainland Hòabìnhan hunter-gatherer sites: Pha Faen in Laos, dated to 7939–7751 calibrated years before present (yr cal bp; present taken as ad 1950), and Gua Cha in Malaysia (4.4–4.2 kyr cal bp)<sup>1</sup>. Here we report, to our knowledge, the first ancient human genome from Wallacea, the oceanic island zone between the Sunda Shelf (comprising mainland southeast Asia and the continental islands of western Indonesia) and Pleistocene Sahul (Australia–New Guinea). We extracted DNA from the petrous bone of a young female hunter-gatherer buried 7.3–7.2 kyr cal bp at the limestone cave of Leang Panninge<sup>2</sup> in South Sulawesi, Indonesia. Genetic analyses show that this pre-Neolithic forager, who is associated with the ‘Toalean’ technocomplex<sup>3,4</sup>, shares most genetic drift and morphological similarities with present-day Papuan and Indigenous Australian groups, yet represents a previously unknown divergent human lineage that branched off around the time of the split between these populations approximately 37,000 years ago<sup>5</sup>. We also describe Denisovan and deep Asian-related ancestries in the Leang Panninge genome, and infer their large-scale displacement from the region today.

[Download PDF](#)

## Main

Modern humans crossed through Wallacea (Fig. 1a) to Sahul<sup>5,6,7,8</sup> at least 50 thousand years ago (kya)<sup>9</sup>, and possibly by up to 65 kya<sup>10</sup>. Presently, however, the earliest archaeological evidence for our species in Wallacea dates to at least 45.5 kya for figurative art in Sulawesi<sup>11</sup>, and 47–43 kyr cal bp for a behavioural shift at Liang Bua (Flores, Indonesia)<sup>12</sup>. The oldest *Homo sapiens* skeletal remains date to 13 kya<sup>13</sup>. The route modern humans used to enter Sahul is not known<sup>14</sup>. Demographic models infer a population split between the ancestors of Oceanian and Eurasian groups approximately 58 kya, whereas Papuan and Aboriginal Australian groups separated around 37 kya<sup>5</sup>. Within this time interval, modern humans admixed multiple times with groups related to Denisovans<sup>15,16,17,18,19,20,21,22,23</sup>, and potentially other unknown hominins<sup>24,25</sup>. The genetic ancestry of the two Hòabìnhan-associated foragers from Pha Faen and Gua Cha<sup>1</sup> shows the highest similarity to modern Andamanese peoples. These ancient and present-day peoples lack substantial amounts of Denisovan-related ancestry, suggesting that the Hòabìnhan-associated and Onge-related lineage diverged before the main archaic human introgression events<sup>1</sup>. Current Wallacean individuals carry larger proportions of Denisovan-related ancestry, but at substantially lower frequencies than is the case in Papuan and Indigenous Australian individuals<sup>20</sup>. This is probably due to admixture with the East Asian Neolithic farmers ('Austronesian peoples') who arrived in Wallacea around 4 kya<sup>20,26</sup>.

**Fig. 1: Study site location.**



**a**, Sulawesi and Wallacea. The red rectangle indicates the region shown in **b**. **b**, Leang Panninge. The dotted line indicates Toalean site distribution.

# The Toalean burial from Leang Panninge

The most distinctive archaeological assemblages associated with Holocene hunter-gatherers in Wallacea belong to the Toalean technocomplex (8–1.5 kya)<sup>3,4,27,28</sup>. Found only in a 10,000 km<sup>2</sup> area of South Sulawesi<sup>3</sup> (Fig. 1b), Toalean cultural assemblages are generally characterized by backed microliths and small stone projectiles ('Maros points')<sup>4</sup> (Extended Data Fig. 1a–c). In 2015, excavations at Leang Panninge in the Mallawa district of Maros, South Sulawesi (Fig. 1b), uncovered the first relatively complete human burial from a secure Toalean context (Extended Data Figs. 1–5, [Supplementary Information](#)). The individual was interred in a flexed position<sup>29</sup> in a rich aceramic Toalean stratum. Exposed at a depth of around 190 cm, the burial has an inferred age of 7.3–7.2 kyr cal bp obtained from <sup>14</sup>C dating of a *Canarium* sp. seed (Extended Data Figs. 2, 3, Supplementary Table 1). Morphological characters indicate that this Toalean forager was a 17–18-year-old female with a broadly Australo-Melanesian affinity, although the morphology does not fall outside the range of recent Southeast Asian variation ([Supplementary Information](#)).

## Genomic analysis

We extracted ancient DNA from bone powder obtained from the petrous portion of the temporal bone of the Leang Panninge individual. After library preparation, we used a DNA hybridization capture approach to enrich for approximately 3 million single-nucleotide polymorphisms (SNPs) across the human genome (1240K and archaic admixture capture panels<sup>30</sup>) as well as for the entire mitochondrial genome (mtDNA capture<sup>31</sup>). We retrieved 263,207 SNPs on the 1240K panel, 299,047 SNPs on the archaic admixture panel and the almost complete mtDNA sequence. Authenticity of the analysed ancient DNA was confirmed by short average fragment length, elevated damage patterns towards the molecule ends, and low autosomal and mtDNA contamination estimates (Supplementary Fig. 1). We confirmed that the individual was genetically of female sex. Analysis of the polymorphisms present in the reconstructed mtDNA sequence suggests a deeply divergent placement within mtDNA haplogroup M (Supplementary Table 17, Supplementary Fig. 2).

We initiated our genomic investigation by principal component analyses (PCAs), comparing the Leang Panninge genome with present-day individuals from East Asia, southeast Asia and Near Oceania (comprising Indigenous Australia, Papua New Guinea and Bougainville) genotyped on the Human Origins SNP panel<sup>18,32,33,34</sup>. The newly generated genome and relevant published genomes from ancient individuals from eastern Eurasia were then projected on the PCA<sup>1,34,35,36,37,38</sup>. Leang Panninge falls into PCA space not occupied by any present-day or ancient individuals, but is broadly located between Indigenous Australian peoples and the Onge (Fig. 2a,

Extended Data Fig. 6).  $F_3$ -statistics<sup>33</sup> of the form  $f_3$  (Mbuti; Leang Panninge, X), where X is replaced with present-day Asian-Pacific groups, indicated that the new genome shares most genetic drift with Near Oceanian individuals (Fig. 2b). We confirmed these results with  $f_4$ -statistics<sup>33</sup>, suggesting similar affinity of Leang Panninge and Papuan individuals to present-day Asian individuals, despite Near Oceanian groups forming a clade to the exclusion of Leang Panninge (Extended Data Fig. 7a,b). All present-day groups from the region, with the exception of the Mamanwa and the Lebbo<sup>26</sup>, carry only a minor contribution of Papuan-related ancestry (Supplementary Fig. 4).

**Fig. 2: The Leang Panninge genome within the regional genetic context.**



**a**, PCA calculated on present-day individuals from eastern Eurasia and Near Oceania, projecting key ancient individuals from the region<sup>1,34,35,36,37,38</sup>. **b**, Shared genetic drift of present-day groups with the Leang Panninge individual, as calculated using  $f_3$  (Mbuti; Leang Panninge, X) mapped at the geographical position of the tested group. WGA, whole genome amplification. **c**, The amount of introgressed Denisovan sequence in fragments longer than 0.05 cM in present-day (Simons Genome Diversity Project) individuals and longer than 0.2 cM in ancient individuals (measured with admixfrog). Each bar represents the posterior mean estimate from a single genome and the whiskers indicate 2 s.d. (estimated from 200 samples from the posterior decoding).

To investigate the presence and distribution of genetic contributions attributable to Denisovan-related groups<sup>39</sup>, we calculated the statistic  $f_4$  (Mbuti, Denisova; Leang

Panninge, X), where X are groups from present-day Island Southeast Asia, Near Oceania and the Andamans, as well as ancient Asian-Pacific individuals<sup>1,37,38</sup>. Positive values calculated for Near Oceanian groups suggest higher proportions of Denisovan-related ancestry than the Leang Panninge individual ( $z$ -scores of  $>3.19$ ), while the Onge and the remaining ancient individuals returned negative values, indicating a lower proportion of Denisovan-related ancestry (Extended Data Fig. 7c, Supplementary Fig. 3). We also calculated  $f_4$ -ratio statistics to estimate the Denisovan proportion using SNPs from the 1240K capture panel and Han individuals from East Asia as a baseline<sup>18</sup>. We confirmed that Indigenous Australian and Papuan individuals carry a similar amount of Denisovan ancestry (approximately 2.9%)<sup>18,21,40</sup>, whereas the Leang Panninge individual has a lower value of approximately  $2.2 \pm 0.5\%$  (Supplementary Tables 18–20). The Denisovan admixture proportion in the Leang Panninge individual is higher than in the Hòabìnhian individuals from Pha Faen and Gua Cha<sup>1</sup>, suggesting that groups ancestral to hunter-gatherers from Wallacea and Sunda were involved in different introgression events with archaic hominins. In addition, we performed  $D$ -statistics on a set of SNPs designed to measure the contribution of archaic ancestry in modern humans (archaic admixture capture panel). The Leang Panninge individual shares fewer Denisovan-related alleles with Papuan individuals, but has more such alleles than most tested groups, including the Tianyuan individual from Late Pleistocene China<sup>38</sup>. Neanderthal allele sharing is similar across all tested present-day non-African groups (Supplementary Tables 21–23). Finally, we ran admixfrog<sup>41</sup> on the set of archaic admixture SNPs and measured 22.4 Mb ( $\pm 1.9$  Mb) of Denisovan-related ancestry in 33 fragments distributed across the Leang Panninge genome. This contribution accounts for around half of what is found in Papuan groups, but there is a significant correlation between the Denisovan fragments in the Leang Panninge genome and those in present-day Near Oceanian groups, suggesting shared introgression events (Fig. 2c, Extended Data Fig. 8, Supplementary Fig. 5).

To investigate whether the apparent PCA shift of Leang Panninge away from Near Oceanian groups is due to genetic drift alone, we performed a multidimensional scaling plot based on genetic similarities measured as  $1 - f_3$  (Mbuti; Leang Panninge, X). The multidimensional scaling positioning of the Leang Panninge individual recapitulates the PCA with an intermediate placement between Papuan and Asian individuals (Extended Data Fig. 9). We then used  $f_4$ -statistics and qpWave<sup>33</sup> to formally test for the presence of additional genetic sources in Leang Panninge other than the Papuan-related ancestry. This identified a marginal affinity towards ancient Asian genomes (Extended Data Fig. 7d), and a minimum of two streams of ancestry when Denisova<sup>37</sup> and/or ancient Asian groups<sup>1,37,38</sup> were included in the qpWave reference groups (Supplementary Table 24). On the basis of these results, we used qpAdm<sup>33</sup> to identify potential sources for an Asian-related ancestry in the genome

alongside the Papuan-related component (Supplementary Table 25). Using a rotating approach among different Asian groups<sup>1,37,38</sup>, we were able to model the Leang Panninge genome as a mixture between Papuan and Tianyuan ( $51 \pm 11\%$ ) or Onge ( $43 \pm 9\%$ ) (Fig. 3a, Supplementary Table 26). Further exploration with admixture graphs built in qpGraph<sup>33</sup> and TreeMix<sup>42</sup>, including present-day groups and relevant ancient individuals<sup>37,38,40,43</sup>, provided evidence for the presence of deep Asian ancestry (Fig. 3b,c, Supplementary Figs. 6–11). In TreeMix, the first admixture edge represents archaic introgression from Denisovan-related groups into the common ancestor of Leang Panninge and present-day Near Oceanian peoples. This is followed by an East Asian-related gene flow into Leang Panninge departing basally from the Qihe lineage, an early Neolithic genome from southeastern China<sup>37</sup> (Fig. 3b, Extended Data Fig. 10, Supplementary Fig. 6). The qpGraph analysis confirmed this branching pattern, with the Leang Panninge individual branching off from the Near Oceanian clade after the Denisovan gene flow, although with the most supported topology indicating around 50% of a basal East Asian component contributing to the Leang Panninge genome (Fig. 3c, Supplementary Figs. 7–11).

**Fig. 3: Admixture signals detected in the Leang Panninge genome.**



**a**, Admixture proportions modelling Leang Panninge as a combination of Papuan<sup>49</sup> and Tianyuan<sup>38</sup> or Onge<sup>49</sup> groups as estimated by qpAdm<sup>33</sup> using Mbuti, Denisovan<sup>39</sup>, Kostenki 14 (ref. 50) and ancient Asian individuals<sup>1,37</sup> as rotating reference groups (Supplementary Table 26). The error bars denote standard errors as calculated with block jackknife in the qpAdm software. **b, c**, Admixture graphs placing Leang Panninge

on the branch with the present-day Near Oceanian clade<sup>41</sup> and showing the admixture with a deep Asian-related ancestry in TreeMix<sup>42</sup> (**b**) (Extended Data Fig. 10, Supplementary Fig. 6) and qpGraph (**c**) (worst  $z$ -score of  $-2.194$ ; Supplementary Figs. 7–11)<sup>33,37,38,39,43</sup>. In **b**, ‘1’ and ‘2’ refer to the order in which the TreeMix software added ‘migration events’ (indicated by the arrows) to the graph. When plotting qpGraph results (**c**), the dotted arrows indicate admixture edges.

## Discussion

Genome-wide analyses of the Leang Panninge individual show that most genetic drift is shared with present-day groups from New Guinea and Aboriginal Australia (Fig. 2b, Extended Data Fig. 7a). However, this Toalean-associated genome represents a previously undescribed ancestry profile, one that branched off after Onge-related and Hòabìnhan-related lineages but around the time that Papuan and Indigenous Australian groups split (Fig. 3b, c, Extended Data Fig. 8, Supplementary Figs. 6–9). It is possible that this Toalean individual carries a local ancestry that was present in Sulawesi before the initial peopling of Sahul at least 50 kya<sup>9</sup>, although whether this population produced the Late Pleistocene rock art in the south of the island<sup>11,44,45</sup> is unknown.

The Toalean individual carries substantial Denisovan-related ancestry, probably sharing the archaic admixture event with present-day Near Oceanian groups (Fig. 2c, Extended Data Figs. 7c, 10, Supplementary Fig. 5, Supplementary Tables 21–23). This provides strong support for the main Denisovan-related gene flow happening before modern humans reached Sahul, making both Wallacea and Sunda equally likely locations for this archaic introgression event. However, previously published hunter-gatherer genomes from Sunda carry almost no Denisovan-related ancestry (Supplementary Tables 18–20), suggesting either a Hòabìnhan-related spread into southeast Asia after the aforementioned gene flow or that Wallacea was indeed the crucial meeting point between archaic and modern humans. The apparent presence of a long-established population of archaic hominins in southwestern Sulawesi<sup>46</sup> provides a possible source for the introgression event. Two previous studies have suggested that two deeply divergent Denisovan lineages admixed into the ancestors of Papuan individuals<sup>17,22</sup>, but our genomic data currently do not have enough resolution to distinguish among one or multiple introgression pulses.

The lower amount of Denisovan ancestry in the Leang Panninge individual than in Papuan and Indigenous Australian individuals could result from: (1) an additional admixture with Denisovan ancestry into the common ancestors of Near Oceanian groups, or (2) a dilution of the Denisovan-related ancestry in the Leang Panninge genome through admixture with lineages carrying less or no such ancestry. Our allele

frequency-based analyses do not support the first scenario (Supplementary Fig. 11), but they do favour the latter. The scarcity of pre-Neolithic genomes from across Asia prevents us from defining the exact source and admixture proportions of this gene flow event. It is noteworthy, however, that despite the reconstructed population trees (TreeMix and qpGraph) suggesting a genetic influence on middle Holocene Sulawesi from mainland East Asia, our qpAdm modelling cannot rule out a southeast Asian contribution from a group related to present-day Andamanese peoples (Fig. 3, Supplementary Figs. 6–11, Supplementary Table 26). This is consistent with a recent study that describes widespread admixtures across Asia between Onge-related and Tianyuan-related ancient populations<sup>47</sup>. However, the presence of this type of ancestry in a middle Holocene forager from Wallacea suggests that the Asian-related admixture could have taken place long before the expansion of Austronesian societies into the region.

We could not detect evidence for the Leang Panninge ancestry in any tested present-day groups (Supplementary Fig. 4). This could be owing to the overall limited proportion of Near Oceanian-related ancestry in Wallacea or large-scale genetic discontinuity between earlier hunter-gatherers and modern groups. The latter scenario would suggest that any genetic signal related to the Leang Panninge individual was obscured by later demographic processes, including the Austronesian expansion<sup>1,20,26,48</sup>. Higher coverage genetic data from present-day populations in Sulawesi, and additional Toalean ancient genomes, are needed to further investigate this unique ancestry profile and the genetic diversity of hunter-gatherers from Wallacea more generally.

## Methods

### Archaeology

Leang Panninge was first identified as a site with high archaeological potential during a 2013 survey by A.B., B.H. and B.B. Since this time, the limestone cave has been the focus of several excavations undertaken by different teams (Extended Data Fig. 2, [Supplementary Information](#)). The first, the excavation of a 1-m<sup>2</sup> test pit (labelled TP1), was conducted by Balai Arkeologi Sulawesi Selatan (Balar Sulsel) in 2014 (ref. 2). This was followed in 2015 by three excavations (of 1-m<sup>2</sup> test pits MLP/A.1'/13, MLP/A.2'/13 and MLP/B.3'/1) spread across the cave, including one just outside the mouth, by Balai Pelestarian Cagar Budaya (BPCB) Sulawesi Selatan. The purpose of these excavations was to assess the importance of the site (the resultant report concluded that it be listed on the BPCB cultural heritage database). Later the same year, Balar Sulsel returned in collaboration with Universitas Hasanuddin (UNHAS) and Universiti Sains Malaysia to excavate a trench in the northern end of the cave

(contiguous units S8T5 and S8T6) and in the central floor area (contiguous units S16T6 and S17T6). Part of a human skull in a burial context was discovered towards the end of this excavation. Owing to time and financial constraints, the burial was covered with plastic sheets and the trench backfilled to protect it for subsequent excavations. Balar Sulsel continued work further into the cave in 2016 (excavation unit S30T9). Squares S16T6 and S17T6 were reopened in 2018 to retrieve the human skeleton encountered at the base of the 2015 excavation. In 2019, this trench was extended towards the back of the cave (forming contiguous units S16T7 and S17T7) by a joint Indonesian–Australian team from Griffith University and Pusat Penelitian Arkeologi Nasional (ARKENAS), UNHAS and Balar Sulsel. The primary objectives of the 2019 fieldwork were to assess these adjacent deposits for other human skeletal remains, as well as to obtain samples of plant carbon and other materials with which to more precisely determine the age of the human burial first exposed in 2015. The 2019 investigations were conducted under a foreign research permit issued by Indonesia's State Ministry of Research and Technology (permit no.: 154/SIP/FRP/E5/Dit.KI/VII/2017). The previous, Indonesian-led investigations at Leang Panninge were carried out under the terms of formal notifications to conduct research (*Surat Pemberitahuan* or *Surat Penyampaian*) lodged with local government authorities at various levels of administration, from regency/municipality (*kabupaten*) to district (*kecamatan*) to village (*desa*).

The 2015 excavations were conducted in arbitrary 10-cm-thick spits and wet-sieved through a 3-mm mesh, to a depth of approximately 190 cm, at which point the human skeletal remains were encountered in the southwestern corner of the excavation (spits 19 and 20, layer 4). In 2019, deposits were excavated using the same method, only this time in 5-cm spits; consequently, spit names in S16T7 or S17T7 originate from a depth half that of a spit with the same number in S16T6 or S17T6 (for example, spit 18 is 170–180 cm in the first case or 85–90 cm in the latter; see Extended Data Fig. 3). As noted, the skeleton was recovered from the site in 2018. Owing to the fragility of the skeletal remains, visible elements comprising the skull and pelvic areas were first consolidated with a hardening solution (Paraloid B72 acryl resin) and then removed from the deposit en bloc (Extended Data Fig. 4e). The ‘skull block’ and ‘pelvic block’ were both submitted to computer tomography (CT) at a hospital facility in Makassar, Indonesia (Balai Pengamanan Fasilitas Kesehatan Makassar), using the following CT parameters: collimation: 0.625 mm; pitch: 1/0.625; milliamperes and kilovolts: left alone; kernel: bone; retro reconstruction: 0.3-mm interslice. After CT scanning, the two sediment blocks were excavated under laboratory conditions to remove the skeletal remains. The sediment block containing the skull consisted of an intact portion of the original grave fill located immediately adjacent to and below the cranium, mandible and dental elements. The thickest part of this sediment block measured approximately 100 mm. During the ‘skull block’ excavation, we recovered the right petrous portion of the human temporal bone and thereafter submitted it for

DNA analysis at the Max Planck Institute for the Science of Human History (MPI-SHH) in Jena, Germany. We also recovered stone artefacts and faunal remains, as well as a burnt *Canarium* sp. seed located a few centimetres from the main cluster of cranial bones ([Supplementary Information](#)). This seed yielded an accelerator mass spectrometry (AMS)  $^{14}\text{C}$  age of 7264–7165 yr cal bp (Wk-48639) (Supplementary Table 1).

## Morphological documentation

The Leang Panninge human remains (Supplementary Table 2) are stored at the Archaeology Laboratory of the Archaeology Department (Departemen Arkeologi Fakultas Ilmu Budaya) at UNHAS, Makassar, South Sulawesi, Indonesia. In 2019, D.B. reconstructed and described the human remains under the stewardship of M.N. and I.S. Joins were effected using Tarzan's Grip along with plasticine for missing bone. Skeletal weights were taken with a scale accurate to 1 g. Measurements were taken with a Kincrome electronic calliper accurate to 0.01 mm (generally rounded off to the closest tenth of a millimetre). Teeth were measured for their maximum mesiodistal and buccolingual diameters and also these diameters at the cementoenamel junction. The dental morphological features recorded were those of the Arizona State University system<sup>51</sup>, including reference to standard plaques illustrated in that work, and in ref. <sup>52</sup> for photographs of some other standard plaques. Other sources for recording measurements and anatomical characteristics are described in [Supplementary Information](#).

## Ancient DNA processing

Sampling, extraction, library preparation and indexing were performed in a dedicated clean room for ancient DNA at the MPI-SHH. We obtained bone powder from the right pars petrosa of the Leang Panninge individual by cutting along the margo superior and drilling near the cochlea<sup>53</sup>. DNA was extracted using a modified version of the ancient DNA protocol described in ref. <sup>54</sup>. From the extract, we built a double-stranded library after partial uracil-DNA glycosylase treatment<sup>55</sup> to reduce C>T transitions to the first two base pairs and a single-stranded library on an automated liquid handling system<sup>56</sup>. After double indexing with unique index combinations<sup>57</sup>, the libraries were shotgun-sequenced for a depth of approximately 4 million reads on an Illumina HiSeq 4000 at MPI-SHH using a 75-bp single-read configuration for initial quality assessments.

After further amplification, the libraries were hybridized in-solution to enrich for the complete mitogenome (mtDNA capture<sup>31</sup>) and twice for a targeted set of 2,986,592 SNPs across the human genome (two rounds of ‘1240K’ and ‘archaic ancestry’<sup>30</sup>

captures). The capture products were then sequenced on an Illumina HiSeq 4000 at MPI-SHH using a 75-bp single-read configuration. After AdapterRemoval as implemented in EAGER v.1.92.56<sup>58</sup>, the mtDNA-enriched reads were aligned to the mitochondrial reference genome (rCRS) and the reads from the genome-wide captures to the human reference genome (hg19) using a mapping quality filter of 30 for the circularmapper v.1.93.5 and BWA<sup>59</sup> aligner, respectively. Duplicates were removed with DeDup v.0.12.2 (<https://github.com/apeltzer/DeDup>). Contamination of the single-stranded sequences was assessed with AuthentiCT v.1.0<sup>60</sup>.

We reconstructed the mitochondrial consensus sequence and estimated mitochondrial contamination to  $2 \pm 1\%$  using schmutzi<sup>61</sup>. The mitochondrial haplogroup was ascertained with Haplofind<sup>62</sup>. After merging with published data using mafft v.7.305<sup>63</sup>, we constructed a maximum parsimony tree in MEGA X<sup>64</sup>. On the basis of the misincorporation pattern determined by mapDamage v.2.0.9 as implemented in EAGER v.1.92.56<sup>58</sup>, we trimmed 2 bp off the 1240K-captured double-stranded library data and genotyped the trimmed and untrimmed sequences individually for the 1240K panel using samtools v.1.3 (<https://github.com/samtools/samtools>) and pileupCaller v.1.4.0.2 (<https://github.com/stschiff/sequenceTools>), which randomly calls one allele per SNP site. The untrimmed and trimmed genotypes were then combined, retaining only transversions from the untrimmed genotype and transitions from the trimmed genotype to maximize information from the trimmed ends. The single-stranded library data were instead genotyped using the single-stranded mode of pileupCaller and the two genotypes merged using a custom script. The resulting coverage was suitable for population genetics analyses with 263,207 SNPs on the 1240K and 135,432 SNPs on the Human Origins panel (HO). We also genotyped single-stranded and double-stranded data individually after filtering with PMDtools v.0.6<sup>65</sup>.

## Population genetic analyses

PCAs were performed using smartpca with shrinkmode and lsqmode enabled<sup>66</sup>, calculating the principal components from present-day East and southeast Asian and Oceanian individuals genotyped on the Human Origins panel<sup>18,32,33,34</sup> and projecting all ancient genomes.

All  $f_3$ -statistics and  $f_4$ -statistics were calculated using qp3pop v.420 (inbreed: YES) and qpDstat v.721, respectively<sup>33</sup>. For  $f_3$ -statistics, we used East and southeast Asian and Oceanian groups from the Human Origins dataset to include more comparative populations, whereas for  $f_4$ -statistics, we used a more restricted dataset containing data from the Simons Genome Diversity Project (SGDP<sup>49</sup>) genotyped on the 1240K panel to maximize the number of overlapping SNPs with the Leang Panninge individual. The results of  $f_3$ -statistics were plotted in the geographical location of the test group

using ggplot2 v.3.3.3 in RStudio v.1.2.1335. To investigate the proportion of Denisovan-related ancestry ( $\alpha$ ), we calculated  $f_4$ -ratio statistics using qpF4Ratio<sup>18,20,33</sup>, admixfrog<sup>41</sup> and  $D$ -statistics<sup>33</sup> with a custom script. Using qpWave<sup>33</sup>, we investigated whether we could distinguish between the Papuan-like ancestry present in the Leang Panninge individual compared with present-day Papuan individuals. Admixture proportions were estimated with qpAdm (allsnps: YES)<sup>33</sup>. After file conversion with PLINK v.1.9<sup>67</sup>, we ran TreeMix v.1.12<sup>42</sup> setting the Denisovan genome<sup>39</sup> as the root and utilizing the parameters -k 150 and -global. Models were plotted using RColorBrewer v.1.1.2 in RStudio v.1.2.1335 and the fit was assessed by residual inspection after each additional migration edge was added. Admixture graphs with qpGraph<sup>33</sup> were constructed (outpop: NULL, useallsnps: YES, blgsize: 0.05, forcezmode: YES, lsqmode: YES, diag: 0.0001, bigiter: 15, hires: YES, lambdascale: 1, initmix: 1,000, inbreed: YES) by adding one group after the other, moving from archaic humans over present-day groups to ancient samples and testing all possible one-way and two-way mixtures using a custom script. The decision on which model was chosen to progress with the addition of another group was made based on the lowest worst  $z$ -score calculated for each proposed tree. Admixture time estimation was calculated with DATES v.753<sup>68</sup>.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

The raw and aligned sequences are available at the European Nucleotide Archive under the accession number [PRJEB43715](#).

## Code availability

The custom scripts used in this study are available on request from the corresponding authors.

## References

1. 1.

McColl, H. et al. The prehistoric peopling of Southeast Asia. *Science* **361**, 88–92 (2018).

2. 2.

Hasanuddin. Gua Panningnge di Mallawa, Maros: kajian tentang gua hunian berdasarkan artefak batu dan sisa fauna. *Naditira Widya* **11**, 81–96 (2017).

3. 3.

Bulbeck, D., Pasqua, M. & Di Lello, A. Culture history of the Toalean of South Sulawesi, Indonesia. *Asian Perspect.* **39**, 71–108 (2000).

4. 4.

Perston, Y. L. et al. A standardised classification scheme for the mid-Holocene Toalean artefacts of South Sulawesi, Indonesia. *PLoS ONE* **16**, e0251138 (2021).

5. 5.

Malaspinas, A.-S. et al. A genomic history of Aboriginal Australia. *Nature* **538**, 207–214 (2016).

6. 6.

Pugach, I., Delfin, F., Gunnarsdóttir, E., Kayser, M. & Stoneking, M. Genome-wide data substantiate Holocene gene flow from India to Australia. *Proc. Natl Acad. Sci. USA* **110**, 1803–1808 (2013).

7. 7.

Rasmussen, M. et al. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* **334**, 94–98 (2011).

8. 8.

Teixeira, J. C. & Cooper, A. Using hominin introgression to trace modern human dispersals. *Proc. Natl Acad. Sci. USA* **116**, 15327–15332 (2019).

9. 9.

O'Connell, J. F. et al. When did *Homo sapiens* first reach Southeast Asia and Sahul? *Proc. Natl Acad. Sci. USA* **115**, 8482–8490 (2018).

10. 10.

Clarkson, C. et al. Human occupation of northern Australia by 65,000 years ago. *Nature* **547**, 306–310 (2017).

11. 11.

Brumm, A. et al. Oldest cave art found in Sulawesi. *Sci. Adv.* **7**, eabd4648 (2021).

12. 12.

Sutikna, T. et al. The spatio-temporal distribution of archaeological and faunal finds at Liang Bua (Flores, Indonesia) in light of the revised chronology for *Homo floresiensis*. *J. Hum. Evol.* **124**, 52–74 (2018).

13. 13.

Samper Carro, S. C. et al. Somewhere beyond the sea: human cranial remains from the Lesser Sunda Islands (Alor Island, Indonesia) provide insights on Late Pleistocene peopling of Island Southeast Asia. *J. Hum. Evol.* **134**, 102638 (2019).

14. 14.

Bird, M. I. et al. Early human settlement of Sahul was not an accident. *Sci. Rep.* **9**, 8220 (2019).

15. 15.

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.e9 (2018).

16. 16.

GenomeAsia100K Consortium. The GenomeAsia 100K Project enables genetic discoveries across Asia. *Nature* **576**, 106–111 (2019).

17. 17.

Jacobs, G. S. et al. Multiple deeply divergent Denisovan ancestries in Papuans. *Cell* **177**, 1010–1021.e32 (2019).

18. 18.

Qin, P. & Stoneking, M. Denisovan ancestry in east Eurasian and Native American populations. *Mol. Biol. Evol.* **32**, 2665–2674 (2015).

19. 19.  
Reich, D. et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060 (2010).
20. 20.  
Reich, D. et al. Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am. J. Hum. Genet.* **89**, 516–528 (2011).
21. 21.  
Vernot, B. et al. Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* **352**, 235–239 (2016).
22. 22.  
Choin, J. et al. Genomic insights into population history and biological adaptation in Oceania. *Nature* **592**, 583–589 (2021).
23. 23.  
Teixeira, J. C. et al. Widespread Denisovan ancestry in Island Southeast Asia but no evidence of substantial super-archaic hominin admixture. *Nat. Ecol. Evol.* **5**, 616–624 (2021).
24. 24.  
Mondal, M., Bertranpetti, J. & Lao, O. Approximate Bayesian computation with deep learning supports a third archaic introgression in Asia and Oceania. *Nat. Commun.* **10**, 246 (2019).
25. 25.  
Mondal, M. et al. Genomic analysis of Andamanese provides insights into ancient human migration into Asia and adaptation. *Nat. Genet.* **48**, 1066–1070 (2016).
26. 26.  
Lipson, M. et al. Reconstructing Austronesian population history in Island Southeast Asia. *Nat. Commun.* **5**, 4689 (2014).
27. 27.

Suryatman, et al. Artefak batu preneolitik situs Leang Jarie: bukti teknologi Maros Point tertua di Kawasan budaya Toalean, Sulawesi Selatan. *AMERTA* **37**, 1–17 (2019).

28. 28.

Anggraeni, A., Simanjuntak, T., Bellwood, P. & Piper, P. Neolithic foundations in the Karama valley, West Sulawesi, Indonesia. *Antiquity* **88**, 740–756 (2014).

29. 29.

Pawlak, A., Crozier, R., Fuentes, R., Wood, R. & Piper, P. Burial traditions in early mid-Holocene Island Southeast Asia: new evidence from Bubog-1, Ilin Island, Mindoro Occidental. *Antiquity* **93**, 901–918 (2019).

30. 30.

Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).

31. 31.

Fu, Q. et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr. Biol.* **23**, 553–559 (2013).

32. 32.

Lazaridis, I. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).

33. 33.

Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).

34. 34.

Skoglund, P. et al. Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510–513 (2016).

35. 35.

Kanzawa-Kiriyama, H. et al. A partial nuclear genome of the Jomons who lived 3000 years ago in Fukushima, Japan. *J. Hum. Genet.* **62**, 213–221 (2017).

36. 36.

Sikora, M. et al. The population history of northeastern Siberia since the Pleistocene. *Nature* **570**, 182–188 (2019).

37. 37.

Yang, M. A. et al. Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* **369**, 282–288 (2020).

38. 38.

Yang, M. A. et al. 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. *Curr. Biol.* **27**, 3202–3208.e9 (2017).

39. 39.

Meyer, M. et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222–226 (2012).

40. 40.

Bergström, A. et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020).

41. 41.

Peter, B. M. 100,000 years of gene flow between Neandertals and Denisovans in the Altai mountains. Preprint at <https://doi.org/10.1101/2020.03.13.990523> (2020).

42. 42.

Pickrell, J. K. & Pritchard, J. K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).

43. 43.

Sikora, M. et al. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).

44. 44.

Aubert, M. et al. Pleistocene cave art from Sulawesi, Indonesia. *Nature* **514**, 223–227 (2014).

45. 45.

Aubert, M. et al. Earliest hunting scene in prehistoric art. *Nature* **576**, 442–445 (2019).

46. 46.

van den Bergh, G. D. et al. Earliest hominin occupation of Sulawesi, Indonesia. *Nature* **529**, 208–211 (2016).

47. 47.

Wang, C.-C. et al. Genomic insights into the formation of human populations in East Asia. *Nature* **591**, 413–419 (2021).

48. 48.

Bellwood, P. *First Islanders: Prehistory and Human Migration in Island Southeast Asia* (Wiley Blackwell, 2017).

49. 49.

Mallick, S. et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).

50. 50.

Fu, Q. et al. The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).

51. 51.

Scott, G. R. & Turner, C. G. II. *The Anthropology of Modern Human Teeth: Dental Morphology and its Variation in Recent Human Populations* (Cambridge Univ. Press, 1997).

52. 52.

Hillson, S. *Dental Anthropology* (Cambridge Univ. Press, 1996).

53. 53.

Pinhasi, R. et al. Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS ONE* **10**, e0129102 (2015).

54. 54.

Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).

55. 55.

Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B* **370**, 20130624 (2015).

56. 56.

Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat. Protoc.* **15**, 2279–2300 (2020).

57. 57.

Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2012).

58. 58.

Peltzer, A. et al. EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).

59. 59.

Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

60. 60.

Peyrégne, S. & Peter, B. M. AuthentiCT: a model of ancient DNA damage to estimate the proportion of present-day DNA contamination. *Genome Biol.* **21**, 246 (2020).

61. 61.

Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, 224 (2015).

62. 62.

Vianello, D. et al. HAPLOFIND: a new method for high-throughput mtDNA haplogroup assignment. *Hum. Mutat.* **34**, 1189–1194 (2013).

63. 63.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

64. 64.

Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

65. 65.

Skoglund, P. et al. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc. Natl Acad. Sci. USA* **111**, 2229–2234 (2014).

66. 66.

Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S. & Reich, D. Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441**, 1103–1108 (2006).

67. 67.

Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

68. 68.

Narasimhan, V. M. et al. The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).

## Acknowledgements

The 2019 research at Leang Panninge was funded by an Australian Research Council Future Fellowship awarded to A.B. (FT160100119), along with financial support from Griffith University. The research was authorized by the State Ministry of Research and Technology (RISTEK) and was conducted in collaboration with Pusat Penelitian Arkeologi Nasional (ARKENAS), Jakarta, Indonesia. Other key participating organizations in Indonesia include Departemen Arkeologi Fakultas Ilmu Budaya, Universitas Hasanuddin (UNHAS, Makassar), and Balai Arkeologi Sulawesi Selatan (Makassar). The following UNHAS archaeology students participated in the fieldwork: F. N. Shalawat, A. A. Qalam, K. M. Prayoga, M. Sura, M. N. Taufik and M. A. Oka. Local field assistants included Irwan, Amar and Hardin. The team assisting D.B.'s morphological analysis of the human skeletal remains included Fahkri (Balai Arkeologi Sulawesi Selatan), Cika and M. Hafdal. A. Crowther assisted with palaeobotanical identifications, and R. Wood advised on radiocarbon calibration. The 2015 excavation was undertaken by UNHAS in collaboration with the Centre for Global Archaeological Research, Universiti Sains Malaysia. We thank the computational and laboratory team at the Max Planck Institute for the Science of Human History (Jena), in particular C. Freund, R. Radzeviciute and A. Wissgott. Aspects of this research were also financially supported by S. Chia's Universiti Sains Malaysia Research University Grant (1001/PARKEO/870002), the ERC Starting Grant 'Waves' (ERC758967) and the Max Planck Society. The digital elevation models in Fig. 1a, b were created in ArcGIS (QGIS) from Shuttle Radar Topography Mission (SRTM 1) Arc-Second Global data courtesy of the US Geological Survey.

## Funding

Open access funding provided by Max Planck Society.

## Author information

### Author notes

1. These authors contributed equally: Selina Carlhoff, Akin Duli
2. These authors jointly supervised: Johannes Krause, Cosimo Posth, Adam Brumm

## Affiliations

1. Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany

Selina Carlhoff, Kathrin Nägele, Kay Prüfer, Johannes Krause & Cosimo Posth

2. Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Selina Carlhoff, Kathrin Nägele, Laurits Skov, Benjamin M. Peter, Kay Prüfer, Adam Powell & Johannes Krause

3. Departemen Arkeologi, Fakultas Ilmu Budaya, Universitas Hasanuddin, Makassar, Indonesia

Akin Duli, Muhammad Nur, Iwan Sumantri & Supriadi

4. Pusat Penelitian Arkeologi Nasional (ARKENAS), Jakarta, Indonesia

Adhi Agus Oktaviana & Marlon Ririmasse

5. Place, Evolution and Rock Art Heritage Unit, Griffith Centre for Social and Cultural Research, Griffith University, Gold Coast, Queensland, Australia

Adhi Agus Oktaviana

6. Balai Arkeologi Sulawesi Selatan, Makassar, Indonesia

Budianto Hakim, Andi Muhammad Saiful, Hasanuddin & Suryatman

7. Australian Research Centre for Human Evolution, Griffith University, Brisbane, Queensland, Australia

Basran Burhan, David P. McGahan, Yinika L. Perston, Kim Newman & Adam Brumm

8. Independent researcher, Makassar, Indonesia

Fardi Ali Syahdar

9. Archaeology and Natural History, School of Culture, History and Language, College of Asia and the Pacific, Australian National University, Canberra, Australian Capital Territory, Australia

David Bulbeck

10. Centre for Global Archaeological Research, Universiti Sains Malaysia, Penang, Malaysia

Stephen Chia

11. Departemen Sosiologi, Fakultas Ilmu Sosial, Universitas Hasanuddin, Makassar, Indonesia

Dwia Aries Tina Pulubuhu

12. School of Biological Sciences, Seoul National University, Seoul, Republic of Korea

Choongwon Jeong

13. Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of Tübingen, Tübingen, Germany

Cosimo Posth

14. Senckenberg Centre for Human Evolution and Palaeoenvironment, University of Tübingen, Tübingen, Germany

Cosimo Posth

## Contributions

A.B., J.K., C.P. and A.D. conceived and led the study. The 2015 excavation was led by A.D. in collaboration with S. Chia and involved key contributions from M.N., I.S., Supriadi and D.A.T.P. A.B. directed the 2019 excavation, additional analysis and conceptualization in collaboration with A.A.O., M.R., B.H. and the senior Universitas Hasanuddin team (A.D., M.N. and I.S.). F.A.S. supervised the 2019 fieldwork activities with B.B., who recovered the skeletal remains from the site in 2018 with A.M.S. D.P.M. conducted the laboratory excavation of the human skeletal remains and provided key in-field and post-excavation observations and analyses. D.B. described and analysed the human skeletal remains. Y.L.P. and Suryatman analysed the stone artefacts. K. Newman conducted key post-excavation analyses, data processing and figure preparation with B.B. Hasanuddin conducted key early excavations at the site. S. Carlhoff generated genomic data. S. Carlhoff, L.S., B.M.P. and K.P. performed genetic analyses with critical input from C.P., K. Nägele, C.J. and J.K. S. Carlhoff, K. Nägele, C.P. and A.B. wrote the manuscript with critical input from J.K. and the remaining authors.

## Corresponding authors

Correspondence to [Johannes Krause](#) or [Cosimo Posth](#) or [Adam Brumm](#).

# Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

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

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

## Extended data figures and tables

### [Extended Data Fig. 1 Toalean artefact types and excavations at Leang Panninge.](#)

**a, b**, Typical middle-to-late Holocene artefact types from Toalean assemblages in South Sulawesi; pressure-flaked stone ‘Maros points’ (both artefacts are from Leang Pajae, Maros) (**a**) and backed microliths (top: Leang Pajae; bottom: Leang Bulu Bettue, Maros) (**b**). **c**, Osseous points (left: from the layer 4 Toalean human burial, Leang Panninge; right: Leang Rakkoe, Maros). **d–i**, Excavation at Leang Panninge in 2019. Excavation of a trench placed adjacent to the rear wall in the west-facing cave entrance (**d, e**): viewed from northwest to southeast (note that the name of the cave site is incorrectly transcribed in this locally produced banner) (**d**); viewed from north to south (**e**); northwest to southeast (**f**); detail of the excavation work (**g**). Excavated trench faces at the end of the 2019 field season, following soil column-sampling (**h, i**): trench viewed from northeast to southwest, showing the intersection between the western walls of excavation squares S16T6 and S17T6 and southern walls of S17T6 and S17T7 (**h**); trench viewed from north to south, with the main stratigraphic section facing the viewer showing the northern walls of S16T6 and S16T7 (**i**). Image credits: Y. Perston (**a–c**); Leang Panninge research team (**d–i**).

### [Extended Data Fig. 2 Site plan, stratigraphy and dating evidence at Leang Panninge.](#)

**a–c**, Plan views of the cave site showing the locations of previous excavation units and squares S16T6 and S17T6, where the Toalean burial was exposed in 2015.

Excavations in the west-facing entrance (which is 10-m high and 28-m wide) yielded dense Toalean archaeological assemblages. Light grey, limestone bedrock; dark grey, speleothem column/pillar; grey with x, stalagmite. Contour levels are in centimetres. Acronyms and initialisms denoting Indonesian archaeological institutions and heritage departments are as follows: BALAR, Balai Arkeologi Sulawesi Selatan; BPCB, Balai Pelestarian Cagar Budaya Makassar; UNHAS, Universitas Hasanuddin. The 2015 trench was extended in 2019, forming squares S16T7 and S17T7. **d**, Section of the cave mouth. **e**, Stratigraphy and dating. L1, layer 1; L2, layer 2, and so on. Black dots denote the 3D-plotted positions of stone artefacts, faunal items (bone, teeth and shell), palaeobotanical remains, shells and other findings excavated in 2019. Eight distinct archaeological layers of human occupation are evident: the uppermost cultural stratum, layer 1 (approximately 1.5 kyr cal bp), contains ‘Neolithic’ pottery and Toalean backed microliths; layers 2–4 comprise aceramic Toalean occupation deposits. Maros points pre-date backed microliths and appear in the earliest Toalean level (see also ref. [27](#)); layer 5 (approximately 9.4–8.7 kyr cal bp) and underlying strata (layers 6–8) are pre-Toalean. The deepest Toalean horizon, layer 4, a silty clay (7.9–5.1 kyr cal bp), contained the region’s only definite Toalean-associated burial. The burial was revealed in 2015 within the contiguous squares S16T6 and S17T6. In 2019, a 2 × 1 m trench excavated adjacent to these squares exposed undisturbed strata to a depth of 3 m without reaching bedrock (Extended Data Figs. [2](#), [3](#)). Calibrated AMS  $^{14}\text{C}$  ages (95% probability) are shown on the right ([Supplementary Information](#) contains details on dating methods). Dated materials comprised samples of wood charcoal ( $N = 8$ ) and charred *Canarium* sp. seeds ( $N = 3$ ). The age inversion in layer 4 suggests a degree of post-depositional mixing that does not affect the inferred burial age. **f**, Toalean burial from layer 4 (3D image generated using PhotoScan). The individual was interred in a flexed position and covered with large cobbles (denoted 1–5). **g**, Fragmented cranium. **h**, **i**, Maros points recovered with the human remains. An atypical chert Maros point (**h**) was found just below the cranium. A classic chert Maros point (**i**) with a hollow base and serrated margins (tip is missing due to breakage) was found approximately 40 mm below the greater sciatic notch. Image credits: Leang Panninge research team (**a**–**g**); Y. Perston (**h**, **i**).

### [Extended Data Fig. 3 Stratigraphic section at Leang Panninge.](#)

Data distributions presented in the tables on the left, and calibrated AMS radiocarbon ( $^{14}\text{C}$ ) dates (shown projected onto the stratigraphic profile), are based on the results of the 2019 excavations. Calibrated AMS  $^{14}\text{C}$  ages are reported at 95% probability ([Supplementary Information](#) contains details on dating methods). The Toalean human skeleton was exposed in 2015 by a Universitas Hasanuddin (UNHAS) team in collaboration with Balar Sulsel and Universiti Sains Malaysia (see [Supplementary Information](#)).

## Extended Data Fig. 4 Toalean human burial from Leang Panninge.

**a**, Single adult inhumation. The 17–18-year-old female was buried in a flexed position and partially covered with several large cobbles. **b**, Detail of the fragment clusters comprising the skull area. The skeleton is fragmentary (Supplementary Table 2), but the skull, although crushed post-mortem, is well represented by the mandible and mandibular dentition, maxilla and maxillary dentition, frontal bone, temporals, right parietal and occipital bone. The cranium was positioned between two large cobbles (denoted 1 and 2). **c**, Detail of the burial feature; cobble #1 was placed over one of the hands, cobble #4 was positioned over the left foot, and cobble #5 covered the pelvic region. **d**, Semi-complete pelvis and left foot (plantar surface facing up). **e**, Sediment block containing the cranium, following removal from the archaeological trench and before excavation in the laboratory. Image credits: Leang Panninge research team (a–g).

## Extended Data Fig. 5 Toalean stone tool types from the Toalean burial and associated contexts at Leang Panninge.

**a, b**, Position of the chert Maros point (shown inset in a) found below the right pelvis during excavation of the sediment block containing the semi-complete pelvis. The red arrows in a and b indicate the position of the Maros point amidst stone-flaking debris and other cultural artefacts (at the time these photographs were taken, the bones of the pelvic area had already been removed). **c**, Position of the non-classic chert Maros point (shown inset, original location highlighted by the blue arrow) found below the skull amidst other stone artefacts and cultural remains (human teeth are still visible at the top of the sediment block under excavation). **d**, Maros point exposed during the excavation. **e–h**, Examples of Maros points from layer 4, above the level of the Toalean human burial: square S16T6, spit 18 (Find #62) (e); square S16T6, spit 18 (Find #77) (f); square S17T6, spit 18 (Find #54) (g); square S17T6, spit 16 (Find #95) (h). **i–k**, Examples of Maros points from spits 19 and 20 within squares S16T6 and S17T6, where the Toalean human burial was located. **l–p**, Examples of Maros points from other excavated squares at Leang Panninge. Image credits: Leang Panninge research team (a–p).

## Extended Data Fig. 6 Principal component analyses.

Calculated on present-day groups<sup>18,32,33</sup>, projecting all ancient genomes<sup>1,34,35,36,37,38</sup> including either all available SNPs or only damaged DNA fragments after PMD filtering of the single-stranded and double-stranded libraries of the Leang Panninge individual. **a**, Principal components calculated on individuals from East and southeast Asia where Papuan, Indigenous Australian and Nasioi (from Bougainville) groups are

projected. **b**, Principal components calculated on individuals from East, southeast Asia and Near Oceania.

**Extended Data Fig. 7 The position of the Leang Panninge genome within the regional population genetic history.**

Tested with  $f_4$ -statistics, where X denotes the tested group. **a**, The affinity of East Asian, South Asian, southeast Asian and Oceanian groups<sup>49</sup> to Leang Panninge and Papuan individuals. **b**, Affinities between present-day Near Oceanian groups and the Leang Panninge genome<sup>49</sup>. **c**, The affinity of the Leang Panninge genome to the Denisovan genome compared to present-day and ancient individuals from the region<sup>26,36,37,38</sup>. **d**, The affinity of the Leang Panninge genome to ancient genomes from the region<sup>26,36,37,38</sup> in comparison to Papuan groups<sup>49</sup>. Data are presented as exact  $f_4$ -values  $\pm$  one and three standard errors indicated as dark and light grey lines, respectively; statistics with z-scores above  $|3|$  are displayed in green.

**Extended Data Fig. 8 Permutation test on the overlap of introgressed Denisovan sequence in the Leang Panninge individual and present-day groups.**

After 500 permutations where the Denisovan sequence is distributed randomly across the genome. The dotted vertical line is the observed overlap in the non-permuted data. The empirical p-value indicates how many permutations show more overlap than the observed overlap in the non-permuted data.

**Extended Data Fig. 9 Multidimensional scaling plot.**

Visualizing the pairwise difference between Leang Panninge and East and southeast Asian and Near Oceanian groups<sup>18,32,33,34</sup> as calculated from  $1 - f_3$  (Mbuti; Leang Panninge, X) on the Human Origins dataset.

**Extended Data Fig. 10 The genetic relationship of Leang Panninge with other Asia-Pacific groups.**

Data for other Asia-Pacific groups from refs. <sup>37,38,39,43,49</sup>. **a–d**, Maximum likelihood trees after incrementally adding one migration event (from zero to four) as inferred by TreeMix<sup>42</sup>; corresponding residuals of individual models are presented on the right (Supplementary Fig. 6).

# Supplementary information

## Supplementary Information

This file contains supplementary text, supplementary tables 1 – 27, supplementary figures 1 – 11 and supplementary references.

## Reporting Summary

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## Reprints and Permissions

## About this article

### Cite this article

Carlhoff, S., Duli, A., Nägele, K. *et al.* Genome of a middle Holocene hunter-gatherer from Wallacea. *Nature* **596**, 543–547 (2021). <https://doi.org/10.1038/s41586-021-03823-6>

- Received: 27 November 2020
- Accepted: 13 July 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03823-6>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03823-6>

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

- Article
- Open Access
- [Published: 04 August 2021](#)

# Fair algorithms for selecting citizens' assemblies

- [Bailey Flanigan](#) ORCID: [orcid.org/0000-0002-7805-1989<sup>1</sup>](https://orcid.org/0000-0002-7805-1989),
- [Paul Gölz](#) ORCID: [orcid.org/0000-0002-8101-6818<sup>1</sup>](https://orcid.org/0000-0002-8101-6818),
- [Anupam Gupta<sup>1</sup>](#),
- [Brett Hennig](#) ORCID: [orcid.org/0000-0001-8958-0440<sup>2</sup>](https://orcid.org/0000-0001-8958-0440) &
- [Ariel D. Procaccia](#) ORCID: [orcid.org/0000-0002-8774-5827<sup>3</sup>](https://orcid.org/0000-0002-8774-5827)

*Nature* volume 596, pages 548–552 (2021)

- 5036 Accesses
- 138 Altmetric
- [Metrics details](#)

## Subjects

- [Computer science](#)
- [Economics](#)

## Abstract

Globally, there has been a recent surge in ‘citizens’ assemblies’<sup>1</sup>, which are a form of civic participation in which a panel of randomly selected constituents contributes to questions of policy. The random process for selecting this panel should satisfy two properties. First, it must produce a panel that is representative of the population. Second, in the spirit of democratic equality, individuals would ideally be selected to serve on this panel with equal probability<sup>2,3</sup>. However, in practice these desiderata are in tension owing to differential participation rates across subpopulations<sup>4,5</sup>. Here we apply ideas from fair division to develop selection algorithms that satisfy the two desiderata simultaneously to the greatest possible extent: our selection algorithms

choose representative panels while selecting individuals with probabilities as close to equal as mathematically possible, for many metrics of ‘closeness to equality’. Our implementation of one such algorithm has already been used to select more than 40 citizens’ assemblies around the world. As we demonstrate using data from ten citizens’ assemblies, adopting our algorithm over a benchmark representing the previous state of the art leads to substantially fairer selection probabilities. By contributing a fairer, more principled and deployable algorithm, our work puts the practice of sortition on firmer foundations. Moreover, our work establishes citizens’ assemblies as a domain in which insights from the field of fair division can lead to high-impact applications.

[Download PDF](#)

## Main

In representative democracies, political representatives are usually selected by election. However, over the past 35 years, an alternative selection method has been gaining traction among political scientists<sup>2,6,7</sup> and practitioners<sup>1,8,9,10</sup>: ‘sortition’, which is the random selection of representatives from the population. The chosen representatives form a panel—usually known as a citizens’ assembly—that convenes to deliberate on a policy question. (Such panels also go by other names; our work applies to all panels in the broader category of ‘deliberative minipublics’<sup>11</sup>.) Citizens’ assemblies are now being administered by more than 40 organizations in over 25 countries<sup>12</sup>; one of these organizations—the Sortition Foundation in the UK—recruited 29 panels in 2020. Although many citizens’ assemblies are initiated by civil-society organizations, they are also increasingly being commissioned by public authorities on municipal, regional, national and supranational levels<sup>1</sup>. Notably, since 2019, two Belgian regional parliaments have internally established permanent sortition bodies<sup>13,14</sup>. The growing use of citizens’ assemblies by governments is giving the decisions of these assemblies a more direct path to affecting policy. For example, two recent citizens’ assemblies commissioned by the national legislature of Ireland led to the legalization of same-sex marriage and abortion<sup>15</sup>.

Ideally, a citizens’ assembly selected using sortition acts as a microcosm of society: its participants are representative of the population, and thus its deliberation simulates the entire population convening ‘under conditions where it can really consider competing arguments and get its questions answered from different points of view’<sup>16</sup>. However, whether this goal is realized in practice depends on exactly how assembly members are chosen.

Panel selection is generally done in two stages: first, thousands of randomly chosen constituents are invited to participate, a subset of whom opt into a ‘pool’ of volunteers.

Then, a panel of prespecified size is randomly chosen from this pool using some fixed procedure, which we term a ‘selection algorithm’. As the final and most complex component of the selection process, the selection algorithm has great power in deciding who will be chosen to represent the population. In this Article, we introduce selection algorithms that preserve the key desirable property pursued by existing algorithms, while more fairly distributing the sought-after opportunity<sup>17,18,19,20</sup> of being a representative.

To our knowledge, all of the selection algorithms previously used in practice (Supplementary Information section 12) aim to satisfy one particular property, known as ‘descriptive representation’ (that the panel should reflect the composition of the population)<sup>16</sup>. Unfortunately, the pool from which the panel is chosen tends to be far from representative. Specifically, the pool tends to overrepresent groups with members who are on average more likely to accept an invitation to participate, such as the group ‘college graduates’. To ensure descriptive representation despite the biases of the pool, selection algorithms require that the panels they output satisfy upper and lower ‘quotas’ on a set of specified features, which are roughly proportional to the population rate of each feature (for example, quotas might require that a 40-person panel contain between 19 and 21 women). These quotas are generally imposed on feature categories delineated by gender, age, education level and other attributes that are relevant to the policy issue at hand. In Supplementary Information section 3, we demonstrate that quota constraints of this form are more general than those that are achievable via ‘stratified sampling’, which is a technique that is often used for drawing representative samples.

Selection algorithms that pre-date this work focused only on satisfying quotas, leaving unaddressed a second property that is also central to sortition: that all individuals should have an equal chance of being chosen for the panel. Several political theorists present equality of selection probabilities as a central advantage of sortition, and stress its role in promoting ideals such as equality of opportunity<sup>2,21</sup>, democratic equality<sup>16,21,22,23</sup> and allocative justice<sup>23,24</sup>. Engelstad, who introduced an influential model of the benefits of sortition, argues that this form of equality constitutes ‘[t]he strongest normative argument in favour of sortition’<sup>25</sup> (for more details on desiderata from political theory, see Supplementary Information section 4). In addition to political theorists, major practitioner groups have also advocated for equal selection probabilities<sup>4,26</sup>. However, these practitioners face the fundamental hurdle that, in practice, the quotas almost always necessitate selecting people with somewhat unequal probabilities, as individuals from groups that are underrepresented in the pool must be chosen with disproportionately high probabilities to satisfy the quotas. Two previous papers<sup>27,28</sup> have suggested mathematical models in which selection algorithms can reconcile equal selection probabilities with representativeness, but both of these

studies make assumptions that are incompatible with current practice (Supplementary Information section 5).

Although it is generally impossible to achieve perfectly equal probabilities, the reasons to strive for equality also motivate a more gradual version of this goal: making probabilities as equal as possible, subject to the quotas. We refer to this goal as ‘maximal fairness’. We find that our benchmark (a selection algorithm representing the previous state of the art) falls far short of this goal, giving volunteers markedly unequal probabilities across several real-world instances. This algorithm even consistently selects some types of volunteer with near-zero probability, and thus excludes them in practice from the chance to serve. We further show that, in these instances, it is possible to give all volunteers a probability of well above zero while satisfying the quotas, demonstrating that the level of inequality produced by the benchmark is avoidable.

In this Article, we close the gaps we have identified, both in theory and in practice. We first introduce not only a selection algorithm that achieves maximal fairness, but also a more general algorithmic framework for producing such algorithms. Motivated by the multitude of possible ways to quantify the fairness of an allocation of selection probabilities, our framework gives a maximally fair selection algorithm for any measure of fairness with a particular functional form. Notably, such measures include the most prominent measures from the literature on fair division<sup>29,30</sup>, and we show that these well-established metrics can be applied to our setting by casting the problem of assigning selection probabilities as one of fair resource allocation (Supplementary Information section 9). Then, to bring this innovation into practice, we implement a deployable selection algorithm that is maximally fair according to one specific measure of fairness. We evaluate this algorithm and find that it is substantially fairer than the benchmark on several real-world datasets and by multiple fairness measures. Our algorithm is now in use by a growing number of sortition organizations around the world, making it one of only a few<sup>31,32,33,34</sup> deployed applications of fair division.

## Algorithmic framework

### Definitions

We begin by introducing necessary terminology, which we illustrate with an example in Supplementary Information section 1. We refer to the input to a selection algorithm—a pool of size  $n$ , a set of quotas and the desired panel size  $k$ —as an ‘instance’ of the panel selection problem. Given an instance, a selection algorithm randomly selects a ‘panel’, which is a quota-compliant set of  $k$  pool members. We define the ‘output distribution’ of the algorithm for an instance as the distribution that specifies the probabilities with which the algorithm outputs each possible panel. Then, the

‘selection probability’ of a pool member is the probability that they are on a panel randomly drawn from the output distribution. We refer to the mapping from pool members to their selection probabilities as the ‘probability allocation’, which we aim to make as fair as possible. Finally, a ‘fairness measure’ is a function that maps a probability allocation to a fairness ‘score’ (for example, the geometric mean of probabilities, of which higher values correspond to greater fairness). An algorithm is described as ‘optimal’ with respect to a fairness measure if, for any instance, the fairness of the probability allocation of the algorithm is at least as high as that of any other algorithm.

## Formulating the optimization task

To inform our approach, we first analysed algorithms that pre-dated our own. Those algorithms that we have seen in use all have the same high-level structure: they select individuals for the panel one-by-one, and in each step randomly choose whom to add next from among those who—according to a myopic heuristic—seem unlikely to produce a quota violation later. As finding a quota-compliant panel is an algorithmically hard problem (Supplementary Information section 6), it is already an achievement that these simple algorithms find any panel in most practical instances. However, owing to their focus on finding any panel at all, these algorithms do not tightly control which panel they output or, more precisely, their output distribution (the probabilities with which they output different panels). Because the output distribution of an algorithm directly determines its probability allocation, the probability allocations of existing algorithms are also uncontrolled, which leaves room for them to be highly unfair.

In contrast to these existing algorithms, which have output distributions that arise implicitly from a sequence of myopic steps, the algorithms in our framework (1) explicitly compute a maximally fair output distribution and then (2) sample from that distribution to select the final panel (Fig. 1). Crucially, the maximal fairness of the output distribution found in the first step makes our algorithms optimal. To see why, note that the behaviour of any selection algorithm on a given instance is described by some output distribution; thus, as our algorithm finds the fairest possible output distribution, it is always at least as fair as any other algorithm.

**Fig. 1: Algorithm optimizing a fairness measure  $F$ .**

---

 **figure1**

Step (1): construct a maximally fair output distribution  $\mathcal{D}$  over an optimal portfolio  $\mathcal{P}$  of quota-compliant panels (denoted by coloured boxes), which is done by iteratively building an optimal portfolio of panels and computing the fairest distribution over that portfolio. Step (2): sample the distribution to select a final panel.

As step (2) of our selection algorithm is simply a random draw, we have reduced the problem of finding an optimal selection algorithm to the optimization problem in step (1)—finding a maximally fair distribution over panels. To fully specify our algorithm, it remains only to solve this optimization problem.

## Solving the optimization task

A priori, it might seem that computing a maximally fair distribution requires constructing all possible panels, because achieving optimal fairness might necessitate assigning non-zero probability to all of them. However, such an approach would be impracticable, as the number of panels in most instances is intractably large.

Fortunately, because we measure fairness according to individual selection probabilities only, there must exist an ‘optimal portfolio’—a set of panels over which there exists a maximally fair distribution—containing few panels (by Carathéodory’s theorem, as discussed in Supplementary Information section 7). This result brings a practical algorithm within reach, and shapes the goal of our algorithm: to find an optimal portfolio while constructing as few panels as possible.

We accomplish this goal using an algorithmic technique known as ‘column generation’, where, in our case, the ‘columns’ being generated correspond to panels (a formal description is provided in Supplementary Information section 8). As shown in Fig. 1, our algorithms find an optimal portfolio by iteratively building a portfolio of panels  $\mathcal{P}$ . In each iteration, a panel is chosen to be added to  $\mathcal{P}$  via the following two steps: (1a) finding the optimal distribution  $\mathcal{D}$  over only the panels currently in  $\mathcal{P}$  and (1b) adding a panel to  $\mathcal{P}$  that—on the basis of the gradient of the fairness measure—will move the portfolio furthest towards optimality. This second subtask makes use of

integer linear programming, which we use to generate quota-compliant panels despite the theoretical hardness of the problem. Eventually, the panel with the most promising gradient will already be in  $\mathcal{P}$ , in which case  $\mathcal{P}$  is provably optimal, and  $\mathcal{D}$  must be a maximally fair distribution. In practice, we observe that this procedure terminates after few iterations.

Our techniques extend column generation methods that are typically applied to linear programs, allowing them to be used to solve a large set of convex programs (Supplementary Information section 8.1). This extension allows our framework to be used with a wide range of fairness measures—essentially any for which the fairest distribution over a portfolio can be found via convex programming. Supported measures include those most prominent in the fair division literature: egalitarian welfare<sup>35</sup>, Nash welfare<sup>30</sup>, Gini inequality<sup>36,37</sup> and the Atkinson indices<sup>37,38</sup>. Our algorithmic approach also has the benefit of easily extending to organization-specific constraints beyond quotas; for example, practitioners can prevent multiple members of the same household from appearing on the same panel. Owing to its generality, our framework even applies to domains outside of sortition, such as the allocation of classrooms to charter schools<sup>39</sup> or kidney exchange<sup>40</sup> (Supplementary Information section 8.2).

## Deployable selection algorithm

To bring fair panel selection into practice, we developed an efficient implementation of a specific maximally fair selection algorithm, which we call LEXIMIN (defined in Supplementary Information section 10). LEXIMIN optimizes the well-established fairness measure leximin<sup>30,39,41</sup>, which is sensitive to the very lowest selection probabilities. In particular, leximin is optimized by maximizing the lowest selection probability, and then breaking ties between solutions in favour of probability allocations with highest second-lowest probability, and so on. This choice of fairness measure is motivated by the fact that—as we show here and in Supplementary Information section 13—LEGACY (the algorithm used by the Sortition Foundation before their adoption of LEXIMIN) gives some pool members a near-zero probability when much more equal probabilities are possible. This type of unfairness is especially pressing because if it consistently affected pool members with particular combinations of features, these individuals and their distinct perspectives would be ‘systematically excluded from participation’<sup>42</sup>, which runs counter to a key promise of random selection.

To increase the accessibility of LEXIMIN, we have made its implementation available through an existing open-source panel selection tool<sup>43</sup> and on <https://panelot.org/><sup>44</sup>, a website on which anyone can run the algorithm without installation. LEXIMIN has since been deployed by several organizations, including Cascadia (USA), the Danish

Board of Technology (Denmark), Nexus (Germany), of by for\* (USA), Particitiz (Belgium) and the Sortition Foundation (UK). As of June 2021, the Sortition Foundation alone has already used LEXIMIN to select more than 40 panels.

We measure the effect of adopting LEXIMIN over pre-existing algorithms by comparing its fairness to that of LEGACY (described in Supplementary Information section 11). We chose LEGACY as a benchmark because it was widely used before this work, is similar to several other selection algorithms used in practice (Supplementary Information section 12) and is the only existing algorithm we found that was fully specified by an official implementation. We compare LEXIMIN and LEGACY on ten datasets from real-world panels and with respect to several fairness measures, including the minimum probability (Table 1), the Gini coefficient and the geometric mean. This analysis shows that LEXIMIN is fairer in all examined instances, and substantially so in nine out of ten.

**Table 1 List of instances on which algorithms were evaluated**

## Effect of adopting LEXIMIN over LEGACY

We compare the fairness of LEXIMIN and LEGACY using datasets from ten citizens' assemblies, which were organized by six different sortition organizations in Europe and North America. As Table 1 shows, our instances are diverse in panel size (range of 20–170, median of 37.5) and number of quota categories (range of 4–8). On consumer hardware, the run-time of our algorithm is well within the time available in practice.

Out of concern for low selection probabilities, we first compare the minimum selection probabilities given by LEGACY and LEXIMIN, summarized in Table 1. Notably, in all instances except for ‘mass’ (an outlier in that its quotas only mildly restrict the fraction of panels that are feasible), LEGACY chooses some pool members with probability close to zero. We can furthermore identify combinations of features that lead to low selection probabilities by LEGACY across all instances (as described in ‘Individuals rarely selected by LEGACY’ in the Methods), raising the concern that LEGACY may in fact systematically exclude some groups from participation. By contrast, LEXIMIN selects no individual nearly so infrequently, with minimum selection probabilities ranging from 26% to 65% (median of 49%) of  $k/n$ —the ‘ideal’ probability individuals would receive in the absence of quotas.

One might wonder whether this increased minimum probability achieved by LEXIMIN affects only a few pool members who are most disadvantaged by LEGACY. This is not the case: as shown in Fig. 2 (shaded boxes), between 13% and 56% of pool members (median of 46%) across instances receive probability from LEGACY lower than the minimum given to anyone by LEXIMIN (Extended Data Table 2). Thus, even

the first stage of LEXIMIN alone (that is, maximizing the minimum probability) provides a sizable section of the pool with more equitable access to the panel.

**Fig. 2: Selection probabilities.**

 figure2



Selection probabilities given by LEGACY and LEXIMIN to the bottom 60% of pool members on six representative instances, in which pool members are ordered in order of increasing selection probability given by the respective algorithms. Shaded boxes denote the range of pool members with a selection probability given by LEGACY that is lower than the minimum probability given by LEXIMIN. LEGACY probabilities are estimated over 10,000 randomly sampled panels and are indicated with 99% confidence intervals (as described in ‘Statistics’ in the Methods). Green dotted lines show the equalized probability ( $k/n$ ). Extended Data Figs. 1,2 show corresponding graphs for the remaining instances and up to the 100th percentile.

We have so far compared LEGACY and LEXIMIN over only the lower end of selection probabilities, as this is the range in which LEXIMIN prioritizes being fair. However, even considering the entire range of selection probabilities, we find that LEXIMIN is quantifiably fairer than LEGACY on all instances by two established metrics of fairness, namely the Gini coefficient and the geometric mean (Extended Data Table 1). For example, across instances (excluding the instance mass), LEXIMIN decreases the Gini coefficient—a standard measure of inequality—by between 5 and 16 percentage points (median of 12; negligible improvement for mass). Notably, the 16-point improvement in the Gini coefficient achieved by LEXIMIN on the instance ‘obf’ (from 59% to 43%) approximately reflects the gap between relative income inequality in Namibia (59% in 2015) and the USA (41% in 2018)<sup>45</sup>.

## Discussion

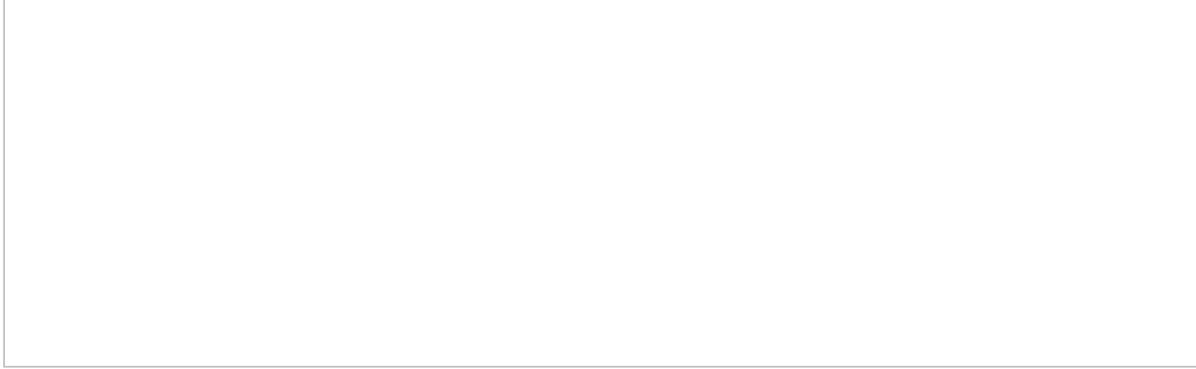
As the recommendations made by citizens’ assemblies increasingly affect public decision-making, the urgency that selection algorithms distribute this power fairly across constituents also grows. We have made substantial progress on this front: the

optimality of our algorithmic framework conclusively resolves the search for fair algorithms for a broad class of fairness measures, and the deployment of LEXIMIN puts an end to some pool members being virtually never selected in practice.

Beyond these immediate benefits to fairness, the exchange of ideas we have initiated between practitioners and theorists presents continuing opportunities to improve panel selection in areas such as transparency. For example, for an assembly in Michigan, we assisted of by for\* in selecting their panel using a live lottery in which participants could easily observe the probabilities with which each pool member was selected. Such lotteries represent an advance over the transparency possible with previous selection algorithms. In this instance, we found that the output distribution of LEXIMIN could be transformed into a simple lottery without a meaningful loss of fairness (Fig. 3). Further mathematical work is needed to show that this transformation can in general preserve strong fairness guarantees.

**Fig. 3: Using LEXIMIN output to select a panel via a live uniform lottery.**

 figure3



**a**, To construct the lottery, the output distribution was transformed into a uniform distribution over 1,000 panels (numbered 000–999). **b**, During the lottery, the three digits that determined the final panel were drawn from lottery machines, making each panel observably selected with equal probability. **c**, The personalized interface (screenshot taken simultaneously with **b**) showed each pool member the number of panels out of 1,000 that they were on, allowing them to verify their own selection probabilities and those of others. Screen capture credit, of by for\*.

The Organisation for Economic Co-operation and Development describes citizens' assemblies as part of a broader democratic movement to 'give citizens a more direct role in [...] shaping the public decisions that affect them'<sup>1</sup>. By bringing mathematical structure, increased fairness and greater transparency to the practice of sortition, research in this area promises to put practical sortition on firmer foundations, and to promote the mission of citizens' assemblies to give everyday people a greater voice.

# Methods

## Theoretical results

The mathematical definitions and proofs supporting this Article can be found in the [Supplementary Information](#). In Supplementary Information section [2](#), we formally define our model of the panel selection problem. In Supplementary Information section [6](#), we prove that, under widely accepted assumptions in complexity theory, panel selection algorithms cannot run in polynomial time, which justifies that our algorithms aim for acceptable running times on observed panel instances rather than for theoretical runtime guarantees. In Supplementary Information section [7](#), we show that Carathéodory's theorem implies the existence of small optimal portfolios, which motivates our use of column generation. Supplementary Information section [8](#) describes the algorithmic ideas behind our algorithmic framework and its applicability to domains outside of sortition, formally defines the framework and when it can be applied, and proves its termination and correctness. In Supplementary Information section [9](#), we cast the problem of panel selection into the language of fair division, which allows us to apply a range of fairness measures from the literature. We also show how each of these fairness measures can be optimized using our framework. In Supplementary Information section [10](#), we describe our algorithm LEXIMIN and prove its correctness. In Supplementary Information section [11](#), we describe the benchmark LEGACY. In Supplementary Information section [13](#), we construct a family of instances in which LEGACY is highly unfair even though the instances allow one to select all agents with equal probability. Finally, in Supplementary Information section [15](#), we analyse panel selection from an axiomatic perspective and describe why we found this approach to be less fruitful than the optimization approach we adopted in this Article.

## Individuals rarely selected by LEGACY

The empirical results in Table [1](#) demonstrate that, in most instances, LEGACY selects some pool members with very low probability. However, in any given citizens' assembly, this does not automatically imply that these individuals had low probability of serving on the panel. Indeed, if such an individual would have been selected by LEGACY with higher probability in most other pools that could have formed (as a result of other sets of agents being randomly invited alongside this individual), then the individual might still have had a substantial overall probability of serving on the citizens' assembly.

In this section, we show how our data suggest that this is not the case, and that some people do in fact seem to have very low likelihood overall of ending up on the panel when LEGACY is used. We make this case by demonstrating two separate points.

First, we show that, across instances, LEGACY tends to give very low selection probabilities to agents who have many features that are overrepresented in the observed pool relative to the quotas. Second, we discuss why it is likely that, across possible pools for the same citizens' assembly, it is usually the same agents who have many overrepresented features. These two points, taken together, suggest that agents who have many overrepresented features in the pools we observe are rarely selected by LEGACY overall.

### **Relationship between overrepresentation of features and selection probability**

To measure the relationship between the level of overrepresentation of an agent's features and that agent's selection probability by LEGACY, we first construct a simple indicator called the 'ratio product', which measures the level of overrepresentation of a given agent's set of features in the pool. The ratio product is composed of, for each of the features of an agent, the ratio between the fraction of this feature in the pool and the fraction of the quotas of the feature (specifically, the mean of lower and upper quota) in the panel. That is, if we denote the set of pool members with a feature  $f$  by  $N_f$  and if we denote the lower and upper quotas of the feature by  $\ell_f$  and  $u_f$ , respectively, then the ratio product of an agent  $i$  is defined as:

$$\$ \$ \mathop{\prod} \limits_{f=1}^N \frac{N_{f,i}}{(\ell_f + u_f)/2k} . \$ \$$$

Given that the quotas are typically set in proportion to the share of the feature in the population, we say that agents with a high ratio product have many overrepresented features. Using this indicator, we find that there is a clear negative relationship in all instances between the ratio product of an individual and their selection probability by LEGACY (Extended Data Fig. 3). Most importantly, as this trend would suggest, we find that the pool members with the largest ratio products consistently have some of the lowest selection probabilities.

### **The same agents probably have many overrepresented features across most possible pools**

Recall that we define an instance with respect to a single pool. However, this observed pool is only one among several hypothetical pools that could have resulted from the random process of sending out invitation letters. We define the ratio product of an agent with respect to a single instance and, therefore, a single observed pool. Then, if a different hypothetical pool (including that agent) had instead been drawn during the invitation process, the ratio product of the same agent with respect to that pool would probably be different, depending on which constituents were invited to join the pool alongside them. As the quotas and the target panel size  $k$  would be the same for

all these hypothetical instances, the differences in ratio product would be due to different values of  $|N_f|$ , for all features  $f$  of the agent. Here,  $|N_f|$ —a random variable, the value of which is determined during the random invitation process—essentially follows a hypergeometric distribution, because it is simply the number of invitations sent to constituents who both have feature  $f$  and are willing to participate. Consequentially, all  $|N_f|$  are well-concentrated, from which it follows that the ratio product of an individual should not vary much across all hypothetical pools containing them. The ratio product should be especially concentrated when all of an individual's features tend to be overrepresented, and thus all factors of the ratio product are large.

## Interpretation of results

The analysis so far suggests that LEGACY selects individuals with many overrepresented features with low probability. Even so, one might consider the possibility that these individuals are more likely to join the pool if invited (given that they are overrepresented in the pool), and that, therefore, their lower selection probability by LEGACY in the panel-selection stage is outweighed by their higher probability of entering the pool in the pool-formation stage. This raises the question of whether the low selection probabilities given to these individuals by LEGACY are necessarily inconsistent with a scenario in which the probabilities of people going from population to panel (their ‘end-to-end’ probabilities<sup>17</sup>) are actually equal.

A back-of-the-envelope calculation suggests that this is not the case—that, in fact, the end-to-end probabilities are probably far from equal when using LEGACY. Across instances, the median ratio between the average selection probability  $k/n$  and (the upper confidence bound on) the minimum selection probability given by LEGACY is larger than 100. If the selection probability of an individual conditioned on appearing in some pool is indeed 100 times lower than that of an ‘average’ citizen, the individual would have to enter the pool 100 times more frequently than this average citizen to serve on the panel with equal end-to-end probability. Given that average response rates are typically between 2 and 5%, someone opting into the pool 100 times more frequently than an average citizen is simply not possible.

Although we have demonstrated that LEGACY underrepresents a specific group (agents with many overrepresented features), we do not have reason to believe that LEGACY would exclude groups defined by intersections of few features (for example, ‘young women’ or ‘conservatives with a university degree’ are the intersection of two features). In Supplementary Information section 14, we investigate the representation of such groups for one instance, ‘sf(e)’. There, we find that LEGACY and LEXIMIN represent intersectional groups to similar degrees of accuracy (Extended Data Fig. 4), explore factors determining the representation of an intersectional group and describe

how the accuracy of intersectional representation could be improved using our algorithmic framework.

## Instance-data preprocessing

At the request of practitioners, we pseudonymize the features of each dataset. This does not affect the analysis, as both LEGACY and LEXIMIN are agnostic to this information.

For data from Healthy Democracy (instance ‘hd’), of by for\* (instance ‘obf’) and MASS LBP (instance ‘mass’), and for the instance ‘sf(e)’ from the Sortition Foundation, respondent data and quotas were taken without modification. For privacy reasons, pool members with non-binary gender in the instances ‘sf(a)’ to ‘sf(d)’ were randomly assigned female or male gender with equal probability. In two of these instances (‘sf(a)’ and ‘sf(d)’), the originally used quotas were not recorded in the data, but we reconstructed them according to the procedures of the Sortition Foundation for constructing quotas from the population fractions. The panel from the Center for Climate Assemblies (instance ‘cca’) did not formally use upper and lower quotas; instead, exact target values for each feature were given (which could not simultaneously be satisfied) as well as a priority order over which targets were more important than others. We set quotas by identifying the minimal relaxation to the lowest-priority target that could be satisfied. For the Nexus instance (instance ‘nexus’), the region of one pool member was missing and inferred from their city of residence. Because Nexus only used lower quotas, the upper quotas of each feature were set to the difference between  $k$  and the sum of lower quotas of all other features of the same category. Such a change does not influence the output distribution of either LEGACY or LEXIMIN but makes the ratio product defined in ‘Individuals rarely selected by LEGACY’ above more meaningful. Because Nexus permitted  $k$  to range between 170 and 175, we chose 170 to make their lower quotas as tight as possible.

## Statistics

The selection probabilities of LEXIMIN are not empirical estimates, but rather exact numbers generated by the algorithm, computed from its output distribution.

By contrast, the selection probabilities given to each agent by LEGACY (as used in the numbers in the text and tables) refer to the fraction of 10,000 sampled panels in which the agent appears (in which each sample is from a single run of LEGACY on the same instance).

In Fig. 2, Extended Data Figs. 1,2, when plotting the line representing LEGACY, agents are sorted along the  $x$  axis in order of this empirical estimate of their selection

probability by LEGACY, and this is the selection probability given on the  $y$  axis. As, for each agent, the number of panels on which they appear across runs of LEGACY is distributed as a binomial variable with 10,000 trials and unknown success probability, we indicate Jeffreys' intervals for each of these success probabilities (that is, selection probabilities) with 99% confidence<sup>46</sup>. These are confidence intervals on the selection probability of a specific agent, not on the selection probability of a specific percentile of the agents.

In addition to reporting two-sided 99% confidence intervals on each agents' selection probability by LEGACY, in Table 1, we report a 99% confidence upper bound on the minimum selection given to any agent by LEGACY per instance. We cannot simply set this upper bound equal to the smallest upper end of the two-sided confidence interval of any agent as computed above because out of these many confidence intervals, some are likely to lie entirely below the true selection probability of the respective agent. Instead, we compute the upper bound on the minimum probability using the confidence interval for a single agent, by running two independent sets of 10,000 samples: In the first set of samples (the one discussed two paragraphs prior), we identify a single agent who was least frequently chosen to the panel in this set; then, we count how often this specific agent is selected across the second set of samples and calculate an upper bound based on a one-sided Jeffreys' interval as follows: if the specific agent was selected in  $s$  out of the 10,000 panels, the confidence bound is the 99th percentile of the distribution  $\text{beta}(1/2 + s, 1/2 + 10,000 - s)$ . (The bound would be 1 if  $s = 10,000$ , but this does not happen in any of the instances.) With 99% confidence, this is an upper bound on the selection probability of the specific agent, and thus also an upper bound with 99% confidence on the minimum selection probability.

As the magnitudes of the two-sided confidence intervals in Fig. 2 and Extended Data Figs. 1,2 show, the empirical estimates we get of the selection probabilities of agents by LEGACY are likely to be close to their true values. Moreover, two of the three statistics we report are not very sensitive to sampling errors: For Gini inequality, additive errors in the estimate of selection probabilities translate into additive errors in the Gini coefficient; and, when we report the number of agents whose selection probability by LEGACY lies under the minimum selection probability of LEXIMIN, Fig. 2 and Extended Data Figs. 1,2 show that the confidence intervals of most agents lie either below or above this threshold. Therefore, our analysis of LEGACY selection probabilities should not be substantially affected by the fact that we can only use empirical estimates of selection probabilities rather than the ground-truth selection probabilities themselves. The one exception is the geometric mean, for which the error in estimating small selection probabilities can severely affect the measure. In particular, in all instances in which one individual appeared in 0 out of 10,000 sampled panels, the geometric mean of empirical selection probabilities would be 0. Thus,

when computing the geometric mean for LEGACY in Extended Data Table 1 and in the body of the Article, we erred on the side of being generous to LEGACY by setting the selection probabilities of these individuals to 1/10,000 instead of 0.

The running times of LEXIMIN were measured on a 2017 Macbook Pro with a 3.1-GHz dual-core Intel i5 processor. Although the running time should not depend on random decisions in the algorithm, the running time of calls to the optimization library Gurobi depends on how the operating system schedules different threads. Reported times are medians of three runs, and are rounded to the nearest second if below 60 s, or to the nearest minute otherwise.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

The panel datasets analysed in this Article are not publicly available owing to the potential for identifying specific panels or participants. We cannot share the dataset nexus owing to agreements between Nexus and their upstream data sources. All other datasets are available from P.G. for research purposes only. Any publication of results based on these data are subject to the permission of the organizations supplying the data. For cca and hd data, publication does not require permission.

## Code availability

An implementation of our selection algorithm LEXIMIN as well as all code required to reproduce the empirical results of this Article are available at <https://github.com/pgoelz/citizensassemblies-replication>.

## References

1. 1.

OECD. *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave* (OECD Publishing, 2020).

2. 2.

Carson, L. & Martin, B. *Random Selection in Politics* (Praeger, 1999).

3. 3.

Leydet, D. Which conception of political equality do deliberative mini-publics promote? *Eur. J. Polit. Theory* **18**, 349–370 (2019).

4. 4.

MASS LBP. How to run a civic lottery: designing fair selection mechanisms for deliberative public processes. <https://www.masslbp.com/s/civiclotteryguide.pdf> (2017).

5. 5.

newDemocracy Foundation and United Nations Democracy Fund. *Enabling National Initiatives to Take Democracy Beyond Elections* (newDemocracy Foundation, 2018).

6. 6.

Dahl, R. A. *After the Revolution? Authority in a Good Society* (Yale Univ. Press, 1970).

7. 7.

Delannoi, G. & Dowlen, O. *Sortition: Theory and Practice* (Imprint Academic, 2010).

8. 8.

Fishkin, J. S. & Luskin, R. C. Experimenting with a democratic ideal: deliberative polling and public opinion. *Acta Polit.* **40**, 284–298 (2005).

9. 9.

McKay, S. & MacLeod, P. MASS LBP and long-form deliberation in Canada. *Democr. Theory* **5**, 108–113 (2018).

10. 10.

Participedia [https://participedia.net/search?  
selectedCategory=case&recruitment\\_method=random,stratified](https://participedia.net/search?selectedCategory=case&recruitment_method=random,stratified) (2021).

11. 11.

Smith, G. & Setälä, M. in *The Oxford Handbook of Deliberative Democracy* (eds Bächtiger, A. et al.) 299–314 (Oxford Univ. Press, 2018).

12. 12.

Democracy R&D. About the network. <https://democracyrd.org/about/> (2019).

13. 13.

Niessen, C. & Reuchamps, M. *Designing a Permanent Deliberative Citizens' Assembly*. Centre for Deliberative Democracy and Global Governance Working Paper Series 6 (2019).

14. 14.

Reuchamps, M. Belgium's experiment in permanent forms of deliberative democracy. *Constitution Net* <http://constitutionnet.org/news/belgiums-experiment-permanent-forms-deliberative-democracy> (17 January 2020).

15. 15.

The Irish Citizen's Assembly Project. Irish citizens' assembly (2019–2020). <http://www.citizenassembly.ie/work/> (2019).

16. 16.

Fishkin, J. S. *Democracy When the People Are Thinking: Revitalizing Our Politics Through Public Deliberation* (Oxford Univ. Press, 2018).

17. 17.

Jacquet, V. The role and the future of deliberative mini-publics: a citizen perspective. *Polit. Stud.* **67**, 639–657 (2019).

18. 18.

Neblo, M. A., Esterling, K. M., Kennedy, R. P., Lazer, D. M. & Sokhey, A. E. Who wants to deliberate—and why? *Am. Polit. Sci. Rev.* **104**, 566–583 (2010).

19. 19.

Of by for\*. Final report of the citizens' panel on COVID-19. <https://joinofbyfor.org/wp-content/uploads/2020/11/Final-Report-of-the-Citizens-Panel-on-COVID-19.pdf> (2020).

20. 20.

Of by for\*. Democratic lottery – the citizens' panel on COVID-19. *Vimeo* <https://vimeo.com/458304880#t=17m59s> (2020).

21. 21.

Parker, J. M. *Randomness and Legitimacy in Selecting Democratic Representatives* (Univ. of Texas at Austin, 2011).

22. 22.

Fishkin, J. S. *When the People Speak: Deliberative Democracy and Public Consultation* (Oxford Univ. Press, 2009).

23. 23.

Stone, P. Sortition, voting, and democratic equality. *Crit. Rev. Int. Soc. Polit. Philos.* **19**, 339–356 (2016).

24. 24.

Stone, P. *The Luck of the Draw: The Role of Lotteries in Decision Making* (Oxford Univ. Press 2011).

25. 25.

Engelstad, F. The assignment of political office by lot. *Soc. Sci. Inf.* **28**, 23–50 (1989).

26. 26.

Allianz Vielfältige Demokratie & Bertelsmann Foundation. Citizens' participation using sortition.

[http://aei.pitt.edu/102678/1/181102\\_Citizens\\_Participation\\_Using\\_Sortition\\_mb - Copy.pdf](http://aei.pitt.edu/102678/1/181102_Citizens_Participation_Using_Sortition_mb - Copy.pdf) (2018).

27. 27.

Benadè, G., Götz, P. & Procaccia, A. D. No stratification without representation. In *Proc. ACM Conf. Econ. Comp.* (eds Karlin, A. et al.) 281–314 (Association for Computing Machinery, 2019).

28. 28.

Flanigan, B., Götz, P., Gupta, A. & Procaccia, A. D. Neutralizing self-selection bias in sampling for sortition. In *Adv. in Neural Inform. Process. Syst.* 33 (eds Larochelle, H. et al.) 6528–6539 (Curran Associates, 2020).

29. 29.

Brams, S. J. & Taylor, A. D. *Fair Division: From Cake-Cutting to Dispute Resolution* (Cambridge Univ. Press, 1996).

30. 30.

Moulin, H. *Fair Division and Collective Welfare* (MIT Press, 2004).

31. 31.

Fair Outcomes. Fair Proposals. <https://www.fairproposals.com/> (accessed 2021).

32. 32.

Sun, A. Divide your rent fairly. *New York Times* <https://www.nytimes.com/interactive/2014/science/rent-division-calculator.html> (28 April 2014).

33. 33.

Goldman, J. & Procaccia, A. D. Spliddit: unleashing fair division algorithms. *ACM SIGecom Exchanges* **13**, 41–46 (2015).

34. 34.

Budish, E., Cachon, G. P., Kessler, J. B. & Othman, A. Course match: a large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Oper. Res.* **65**, 314–336 (2017).

35. 35.

Endriss, U. Lecture notes on fair division. Preprint at <https://arxiv.org/abs/1806.04234> (2010).

36. 36.

Lesca, J. & Perny, P. LP solvable models for multiagent fair allocation problems. In *Proc. 19th European Conf. Artificial Intelligence* (eds Coelho, H. et al.) 393–398 (IOS, 2010).

37. 37.

Endriss, U. Reduction of economic inequality in combinatorial domains. In *Proc. 2013 Internat'l Conf. Autonom. Agents Multiagent Syst.* (eds Gini, M. L. et al.) 175–182 (International Foundation for Autonomous Agents and Multiagent Systems, 2013).

38. 38.

Schneckenburger, S., Dorn, B. & Endriss, U. The Atkinson inequality index in multiagent resource allocation. In *Proc. 16th Internat'l Conf. Autonom. Agents Multiagent Syst.* (eds Larson, K. et al.) 272–280 (International Foundation for Autonomous Agents and Multiagent Systems, 2017).

39. 39.

Kurokawa, D., Procaccia, A. D. & Shah, N. Leximin allocations in the real world. *ACM TEAC* **6**, 1–24 (2018).

40. 40.

Roth, A. E., Sönmez, T. & Ünver, M. U. Pairwise kidney exchange. *J. Econ. Theory* **125**, 151–188 (2005).

41. 41.

Bogomolnaia, A. & Moulin, H. Random matching under dichotomous preferences. *Econometrica* **72**, 257–279 (2004).

42. 42.

Smith, G. *Democratic Innovations: Designing Institutions for Citizen Participation* (Cambridge Univ. Press, 2009).

43. 43.

Hennig, B. & Götz, P. StratifySelect.  
<https://github.com/sortitionfoundation/stratification-app> (2020).

44. 44.

Flanigan, B., Götz, P., Gupta, A., Procaccia, A. D. & Rusak, G. Panelot.  
<http://www.panelot.org/> (2020).

45. 45.

The World Bank Group. GINI index (World Bank estimate).  
<https://data.worldbank.org/indicator/SI.POV.GINI> (2021).

46. 46.

Brown, L. D., Cai, T. T. & DasGupta, A. Interval estimation for a binomial proportion. *Stat. Sci.* **16**, 101–117 (2001).

## Acknowledgements

We thank all the practitioners who have provided data and/or shared their insights with us, including G. Muller and M. Chang (on behalf of Cascadia Consulting Group), A. Gąsiorowska and M. Gerwin (on behalf of the Center for Climate Assemblies), D. Schechter (on behalf of Democracy R&D), Y. Dejaeghere (on behalf of G1000), J. Birkenhäger (on behalf of IFOK), C. Ellis (on behalf of MASS LBP), C. von Blanckenburg (on behalf of Nexus), A. Cronkright and G. Zisiadis (on behalf of by for\*), P. Verpoort (on behalf of the Sortition Foundation) and S. Pek (on behalf of the University of Victoria); B. Cook, A. Kazachkov, F. Kilinc-Karzan and B. Shepherd for technical discussions; and Y. Dejaeghere, W. Flanigan, J. Gastil, M. Gray, T. Lee, L. Leopold, E. Vitercik and M. Wang for comments on the text. This work was supported by the Office of Naval Research under grant N00014-20-1-2488 (A.D.P.); by the National Science Foundation under grants CCF-1907820 (A.G.), CCF-1955785 (A.G.), CCF-2006953 (A.G.), CCF-2007080 (A.D.P.) and IIS-2024287 (P.G. and A.D.P.); by a Fannie and John Hertz Foundation Fellowship (B.F.), a JPMorgan Chase AI Research Fellowship (P.G.) and a National Science Foundation Graduate Research Fellowship (B.F.). The funders had no role in study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

## Author information

### Affiliations

1. Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, USA  
Bailey Flanigan, Paul Götz & Anupam Gupta
2. Sortition Foundation, Cambridge, UK  
Brett Hennig

3. School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

Ariel D. Procaccia

## Contributions

All authors contributed to the problem formulation, and B.F., P.G., A.G. and A.D.P. did the technical (theoretical) work. B.F., P.G. and B.H. procured the data. B.F. and P.G. conceived the experiments, and P.G. implemented the experiments and the algorithm used by practitioners. B.F. and P.G. took the lead on writing the paper and supplementary materials, and all authors contributed to the editing process.

## Corresponding authors

Correspondence to [Bailey Flanigan](#) or [Paul Gölz](#) or [Ariel D. Procaccia](#).

## Ethics declarations

## Competing interests

B.H. is the founder and co-director of the Sortition Foundation.

## Additional information

**Peer review information** *Nature* thanks Edith Elkind and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

### [Extended Data Fig. 1 Selection probabilities for remaining instances.](#)

Selection probabilities given by LEGACY and LEXIMIN to the bottom 60% of pool members on the 4 instances that are not shown in Fig. 2. Pool members are ordered across the  $x$  axis in order of increasing probability given by the respective algorithms. Shaded boxes denote the range of pool members with a selection probability given by

LEGACY that is lower than the minimum probability given by LEXIMIN. LEGACY probabilities are estimated over 10,000 random panels and are indicated with 99% confidence intervals (as described in ‘Statistics’ in the Methods). Green dotted lines show the equalized probability ( $k/n$ ).

### **Extended Data Fig. 2 Selection probabilities up to the 100th percentile.**

Selection probabilities given by LEGACY and LEXIMIN on all ten instances. Pool members are ordered across the  $x$  axis in order of increasing probability given by the respective algorithms. In contrast to Fig. 2 and Extended Data Fig. 1, this graph shows the full range of selection probabilities (up to the 100th percentile). Shaded boxes denote the range of pool members with a selection probability given by LEGACY that is lower than the minimum probability given by LEXIMIN. LEGACY probabilities are estimated over 10,000 random panels and are indicated with 99% confidence intervals (as described in ‘Statistics’ in the Methods). Green dotted lines show the equalized probability ( $k/n$ ).

### **Extended Data Fig. 3 Overrepresentation and LEGACY selection probabilities.**

Relationship between how overrepresented the features of an agent are and how likely they are to be chosen by the LEGACY algorithm. The level of overrepresentation is quantified as the ratio product (as described in ‘Individuals rarely selected by LEGACY’ in the Methods); agents further to the right are more overrepresented. Across instances, pool members with high ratio product are consistently selected with very low probabilities.

### **Extended Data Fig. 4 Representation of feature intersections.**

For all intersections of two features on the instance sf(e), how far the expected number of group members selected by LEGACY or LEXIMIN differs from the proportional share in the population is shown. Although many intersectional groups are represented close to accurately, some groups are over- and underrepresented by more than 15 percentage points by either algorithm. Which groups get over- and underrepresented is highly correlated between both algorithms. Panel shares are computed for a pool of size 1,727, and population shares are based on a survey with 1,915 respondents after cleaning.

**Extended Data Table 1 Gini coefficient and geometric mean of LEGACY and LEXIMIN**

**Extended Data Table 2 Share below LEXIMIN minimum probability**

# Supplementary information

## Supplementary Information

This file contains Supplementary Sections 1-15 – see Contents page for details.

## Reporting Summary

## Peer Review File

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## Reprints and Permissions

## About this article

### Cite this article

Flanigan, B., Götz, P., Gupta, A. *et al.* Fair algorithms for selecting citizens' assemblies. *Nature* **596**, 548–552 (2021). <https://doi.org/10.1038/s41586-021-03788-6>

- Received: 15 January 2021
- Accepted: 30 June 2021
- Published: 04 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03788-6>

## A bridge across the democracy-expertise divide

- Mark E. Warren

News & Views 04 Aug 2021

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03788-6>

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

- Article
- Open Access
- [Published: 11 August 2021](#)

# Oxytocin neurons enable social transmission of maternal behaviour

- [Ioana Carcea](#) [ORCID: orcid.org/0000-0002-4459-9321](#)<sup>1,2,3,4,5</sup>,
- [Naomi López Caraballo](#)<sup>1,2,3,4</sup>,
- [Bianca J. Marlin](#)<sup>1,2,3,4</sup> nAff13,
- [Rumi Ooyama](#)<sup>1,2,5,6</sup>,
- [Justin S. Riceberg](#)<sup>7</sup>,
- [Joyce M. Mendoza Navarro](#)<sup>1,2,3,4</sup>,
- [Maya Opendak](#)<sup>2,4,8,9</sup>,
- [Veronica E. Diaz](#)<sup>1,2,3,4</sup>,
- [Luisa Schuster](#)<sup>1,2,3,4</sup>,
- [Maria I. Alvarado Torres](#)<sup>1,2,3,4</sup>,
- [Harper Lethin](#)<sup>1,2,3,4</sup>,
- [Daniel Ramos](#)<sup>1,2,3,4</sup>,
- [Jessica Minder](#)<sup>1,2,3,4</sup>,
- [Sebastian L. Mendoza](#)<sup>10</sup>,
- [Chloe J. Bair-Marshall](#)<sup>1,2,3,4</sup>,
- [Grace H. Samadjopoulos](#)<sup>5</sup>,
- [Shizu Hidema](#)<sup>11</sup>,
- [Annegret Falkner](#) [ORCID: orcid.org/0000-0001-9756-7887](#)<sup>2,4,12</sup> nAff14,
- [Dayu Lin](#) [ORCID: orcid.org/0000-0003-2006-0791](#)<sup>2,4,12</sup>,
- [Adam Mar](#)<sup>2,4</sup>,
- [Youssef Z. Wadghiri](#)<sup>10</sup>,
- [Katsuhiko Nishimori](#)<sup>11</sup>,
- [Takefumi Kikusui](#)<sup>6</sup>,
- [Kazutaka Mogi](#)<sup>6</sup>,
- [Regina M. Sullivan](#)<sup>2,4,8,9</sup> &
- [Robert C. Froemke](#) [ORCID: orcid.org/0000-0002-1230-6811](#)<sup>1,2,3,4</sup>

*Nature* volume 596, pages 553–557 (2021)

- 15k Accesses
- 815 Altmetric
- [Metrics details](#)

## Subjects

- [Neural circuits](#)
- [Social behaviour](#)

## Abstract

Maternal care, including by non-biological parents, is important for offspring survival<sup>1,2,3,4,5,6,7,8</sup>. Oxytocin<sup>1,2,9,10,11,12,13,14,15</sup>, which is released by the hypothalamic paraventricular nucleus (PVN), is a critical maternal hormone. In mice, oxytocin enables neuroplasticity in the auditory cortex for maternal recognition of pup distress<sup>15</sup>. However, it is unclear how initial parental experience promotes hypothalamic signalling and cortical plasticity for reliable maternal care. Here we continuously monitored the behaviour of female virgin mice co-housed with an experienced mother and litter. This documentary approach was synchronized with neural recordings from the virgin PVN, including oxytocin neurons. These cells were activated as virgins were enlisted in maternal care by experienced mothers, who shepherded virgins into the nest and demonstrated pup retrieval. Virgins visually observed maternal retrieval, which activated PVN oxytocin neurons and promoted alloparenting. Thus rodents can acquire maternal behaviour by social transmission, providing a mechanism for adapting the brains of adult caregivers to infant needs via endogenous oxytocin.

[Download PDF](#)

## Introduction

Social interactions, such as pair bond formation and child rearing, are fundamental aspects of animal and human behaviour<sup>1,2,3,4</sup>. Parental care is especially important, and is therefore believed to be at least partially innate or induced after mating. However, maternal behaviour can also be acquired from experience. In primates, including humans, non-biological parents can learn to care for children after instruction or observation of experienced caretakers and infants<sup>1,2,3,4,5,6,7,8</sup>. However, it is unclear whether alloparenting can be learned from experience in other species, and what neural mechanisms might underlie maternal learning.

The evolutionarily ancient hormone oxytocin is an important molecular signal for maternal behaviour<sup>1,2,9,10</sup>. In mammals, oxytocin release from the hypothalamus is associated with childbirth and lactation<sup>10,11</sup>. Oxytocin also acts in the brain to increase the salience of social information and enable alloparenting in mice<sup>1,9,12,13,14</sup>. Pup-naive virgin female mice initially ignore neonates and ultrasonic distress calls emitted by isolated pups<sup>15</sup>. However, within days of co-housing with experienced mothers (dams) and litters, most virgin females begin expressing alloparenting behaviours, including retrieving pups to the nest. Oxytocin accelerates the onset of pup retrieval, promoting plasticity in virgin auditory cortex for recognizing pup calls<sup>15</sup>. However, little is known about when PVN neurons in nulliparous, non-lactating adults are activated to promote alloparenting behaviour.

## Continuous behaviour monitoring over days

Emergence of pup retrieval in co-housed virgin females provides an opportunity to monitor neural activity during interactions with adults and infants. To examine the behavioural events and neural activity patterns that lead to maternal behaviour in virgin female mice, we built a system for days-long combined behavioural and neural activity monitoring in cages of mother, litter and co-housed virgin female mice (Fig. [1a](#)). This system consists of an overhead camera imaging homecages with visible (daytime) and infrared light (nighttime), synchronized with ultrasonic audio recordings and amplifiers for neural recordings. We constructed ethograms<sup>16</sup> and quantified the frequency and duration of specific behaviours (for example, spontaneous pup retrieval and time in nest; Fig. [1b](#), Extended Data Fig. [1](#), Supplementary Videos [1](#), [2](#)). We examined daily pup retrieval offline to determine when virgins began reliably responding to infant calls (Fig. [1b,c](#)).

**Fig. 1: Dams shepherd virgins to nest.**

---

 **figure1**

---

**a**, System for continuous monitoring of behaviour and neural activity. **b**, Ethogram showing activities of co-housed dam and virgin mice over a period of four days. **c**, Co-housing with dam and pups led to earlier retrieval by virgins. Left, individual retrieval rates. Right, mean retrieval probability. **d**, Day of retrieval onset was earlier in D+

virgins. **e**, Time in nest for dams and co-housed virgin mice. **f**, Correlation between virgin nest entry during day 1 and virgin retrieval at the end of day 1. **g**, Illustration of shepherding behaviour; arrows, movement direction; dashed circle, nest area. **h**, Probabilities of dam starting (yellow) and ending (orange) chasing of a co-housed virgin, relative to nest position (circle indicates the nest, radius approximately 10 cm). **i**, Distances from nest of dam→virgin chases. **j**, Frequency of shepherding events (grey, individual dyads; red, daily averages across cages) was more frequent than dam→virgin chases in the absence of pups (0.2, indicated by dashed line). **k**, Correlation between shepherding during day 1 and virgin retrieval at day 1 end. Data are mean ± s.e.m.; \* $P < 0.05$ , \*\* $P < 0.01$ .

We found that the presence of the mother accelerated the onset of alloparenting in co-housed virgins. Virgins co-housed with dams and litters (D+ virgins,  $n = 14$ ) began reliably retrieving (in more than 1 out of 10 trials) on day two of co-housing, earlier than virgins co-housed with pups but without dams (D− virgins,  $n = 10$ ) (Fig. 1c,d, Extended Data Fig. 2a,b, Supplementary Videos 3,4;  $P < 0.0001$ ). Pup retrieval was not initially performed by virgins but was reliably performed by dams regardless of litter size (Extended Data Fig. 2c). Virgins also began to spend time in the nest with pups (Fig. 1b,e;  $P < 0.0001$ ), and the more virgins entered the nest, the more likely it was that retrieval occurred on day 1 (Fig. 1f;  $r = 0.66$ ,  $P = 0.009$ ).

## PVN activity when dams shepherd virgins

To determine how dams influenced virgin alloparenting, we analysed four days of video per cage and observed some unusual interactions between dams and virgins. First, mothers attempted to keep virgins within the nest with pups. If the virgin left the nest, the mother would escort or ‘shepherd’ her back (Fig. 1g–j, Supplementary Videos 5–7). Shepherding behaviour emerged over hours to days (Extended Data Fig. 3a) and seemed different from maternal aggression; rather than preventing contact with pups, dams encouraged virgins to enter and remain in the nest. Shepherding was unlikely to be dams ‘retrieving’ virgins as this rarely occurred without pups (Extended Data Fig. 3b,c). Shepherding occurred hundreds of times during co-housing (Fig. 1j, Extended Data Fig. 1a), and its frequency correlated with virgin retrieval at day 1 (Fig. 1k;  $r = 0.75$ ,  $P = 0.001$ ).

We hypothesized that co-housing led to virgin alloparenting via oxytocin. We used designer receptors exclusively activated by designer drugs coupled to inhibitory G-protein receptors (DREADDi) to silence PVN oxytocin neurons (OT-PVN cells; Extended Data Fig. 4a), and found that retrieval onset in naive co-housed virgins was delayed (Fig. 2a). However, silencing of OT-PVN cells had little effect after virgins began retrieving (Extended Data Fig. 4b), similar to primiparous dams<sup>17</sup>.

**Fig. 2: Shepherding and nest entry activate virgin PVN/OT-PVN neurons.**

 figure2

**a**, Left, DREADDi–mCherry in OT-PVN cells. Right, CNO treatment during initial co-housing reduced retrieval in DREADDi-expressing mice. **b**, Optically tagged OT-PVN neurons in vivo. Left,  $\mu$ CT image used to localize electrodes. Middle, ChETA–eGFP expression in OT-PVN neurons. Right, raster plot of photo-tagged OT-PVN neuron; inset, waveform; blue bar, light pulse. **c**, **d**, Raster plot (**c**) and before–after plot (**d**) of PVN single unit from co-housed virgin during shepherding (shading). **e**, Change in spiking rate during shepherding for individual units. Filled bars indicate significant changes in spiking rate. **f**, Units with significant rate changes during shepherding (red) fired less during virgin→dam chases (grey). **g**, Simultaneously recorded PVN cells including OT-PVN cell (u3) during nest entry event. Units are numbered u1–u9. **h**, u3 spiking during nest entry ( $n = 140$  events). **i**, Firing rate during nest entry. **j**, Spiking rate when virgin entered nest after shepherding (red) or voluntarily (grey). Data are mean  $\pm$  s.e.m.; \* $P < 0.05$ , \*\* $P < 0.01$ .

We investigated the interactions that activate virgin PVN and OT-PVN cells. We performed single-unit recordings from virgin PVN during co-housing, including optically identified OT-PVN units expressing channelrhodopsin 2 (Fig. 2b, Extended Data Figs. 5a, 6a, b;  $n = 17$  virgins, 541 PVN units, 21 OT-PVN units), localizing electrodes with co-registered micro-computed tomography ( $\mu$ CT) and MRI (Fig. 2b, Extended Data Fig. 5b). Some units were recorded offline during retrieval testing or other dam–pup interactions (Extended Data Fig. 6c–e). Virgin PVN and OT-PVN units were activated by shepherding or initial nest entry, with less activation from chasing dams, entry into an empty nest or pup calls (Fig. 2c–j, Extended Data Figs. 6c–e, 7a–c).

## Social transmission of maternal behaviour

We also found that virgin alloparenting and PVN activity increased after spontaneous pup retrieval by dams. About once every two hours, dams dropped and then retrieved pups, sometimes bringing pups to the virgin (Fig. 3a, Extended Data Figs. 1, 8a, b, Supplementary Videos 1, 8–10). Remarkably, retrieval by dams evoked responses in virgin PVN (Fig. 3b, c), although the virgin was not directly involved. Some PVN and OT-PVN units responded to both shepherding and dam retrieval, but population responses were heterogeneous (Fig. 3d). We analysed cross-correlations across simultaneously recorded units, and found that pairs containing OT-PVN cells were more correlated during shepherding, nest entry and dam retrieval (Extended Data Fig. 7d, e). These data suggest that diverse experiences activate oxytocin neurons, ensuring that new parents become attuned to infant needs.

**Fig. 3: Observation of pup retrieval.**

---

 **figure3**

---

**a**, Illustration of spontaneous retrieval by dam; arrows, movement direction; dashed lines, nest area; red circle, dropped pup. **b**, Activities of simultaneously recorded virgin PVN cells including OT-PVN cell (blue) during spontaneous retrieval episode by dam. Green, pup dropped; grey, pup retrieval. Top, frequency of emitted sounds. **c**, Spiking rate change in virgins during spontaneous retrievals by dams. **d**, Heterogeneous responses ( $n = 35$  (PVN),  $n = 9$  (OT-PVN) units) during nest entry,

shepherding, and maternal retrieval. Top, OT-PVN; bottom, PVN. **e**, Pup retrieval by D– virgins. **f**, Retrieval by D– virgins retrieval was higher after exposure to dam retrieval with or without a transparent barrier compared with an opaque barrier or with OXTR-KO virgins. **g**, Cumulative distributions (Kaplan–Meier) for retrieval onset. **h**, Example virgin PVN unit responding to both observation of maternal retrieval and self-performance of retrieval (time 0, pup retrieved). **i**, Change in virgin PVN firing rate for units recorded during observation of dam retrievals (blue) and self-performance (white). **j**, PVN unit activity increased during observation in learners ( $n = 65$  units,  $N = 3$  virgins) but not non-learners ( $n = 103$  units,  $N = 5$  virgins).  $*P < 0.05$ ,  $**P < 0.01$ .

Additionally, these results indicate that virgins watch dams engage in maternal care, accelerating alloparenting via social transmission to observing virgins. We hypothesized that alloparenting could also be enhanced in non-co-housed virgins observing pup retrieval by dams. In separate non-co-housed virgins, for four consecutive days we performed ten retrieval tests with the dam (forty dam retrievals total) with the virgin present. Retrieval was assessed in virgins approximately 30 min later. Onset of retrieval in virgins increased after observed dam retrievals (Fig. 3e–g, No barrier: 11 out of 15 virgins retrieving after 40 maternal demonstrations;  $P = 0.0002$ , Fisher’s exact test), but independently of pup behaviours (Extended Data Fig. 8c).

We tested whether this transmission of retrieval behaviour involved physical contact by inserting a transparent barrier between other observer pup-naive virgins and the retrieving dams. Most virgins behind a transparent barrier learned to retrieve (Fig. 3f, g, Supplementary Video 11; transparent: 12 out of 19 virgins retrieving,  $P < 0.0001$ ). Visual input was required, as most non-co-housed virgins behind an opaque barrier did not begin retrieving (Fig. 3f,g, Supplementary Video 12; opaque: 2 out of 18 virgins retrieving,  $P = 0.2$ ). Oxytocin signalling was also important, as oxytocin receptor (OXTR)-knockout (KO) virgins observing maternal retrievals through a transparent barrier did not begin retrieving (Fig. 3f,g, OXTR-KO: 3 out of 10 virgins retrieving,  $P = 0.1$ ). The retrieval-learning rate of non-co-housed virgins was slower than that if co-housed virgins (Fig. 1c, Extended Data Fig. 8d), perhaps owing to heterogeneity of PVN and OT-PVN responses to maternal interactions, such that the variety of experiences occurring during co-housing more effectively activates the central oxytocin system.

We performed monosynaptic rabies-based retrograde tracing to identify visual inputs onto oxytocin neurons, and found projections from medial superficial layers of ipsilateral superior colliculus (sSC→PVN; Extended Data Fig. 9a,b), involved in visual processing<sup>18</sup>. Optogenetic sSC→PVN stimulation could replace visual experience in non-co-housed virgins exposed to maternal retrievals behind an opaque barrier (ChR2-Venus: 4 out of 5 virgins retrieving; control: 1 out of 6 virgins

retrieving;  $P = 0.03$ , Mantel–Cox test; Extended Data Fig. 9c, d). Interactions between virgins and dams or pups in the nest might recruit somatosensory or arousal-based PVN inputs, whereas display of maternal behaviour by dams may lead to collicular visual responses. These results suggest a circuit model by which sensory inputs from social experience converge to activate oxytocin neurons and possibly other hypothalamic cell types to enable parental care (Extended Data Fig. 9e).

Retrievals by dams activated PVN neurons in non-co-housed virgins during observational testing, even though the virgins were not interacting with the dam or pups (Fig. 3h–j). Several of the same cells active during observation also fired during subsequent performance of pup retrieval (Fig. 3h, i; left, 15 out of 65 PVN neurons increased activity during dam retrieval; right, 15 out of 41 of these units increased activity during self-performance). Increased PVN activity was higher on first day versus the last day of observation (Fig. 3j, left), and was elevated only in those non-co-housed virgins that eventually began retrieving (Fig. 3j, right;  $n = 3$  learners, 5 non-learners;  $P = 0.0001$ ). Thus, there is a shared hypothalamic cell population that is sensitive to both watching and performing retrieval, reminiscent of mirror or ‘social’ cells in other brain structures<sup>19,20</sup>. This suggests that maternal behaviours performed by the dam are socially transmitted to other females, sometimes visually, via recruitment of virgin PVN neurons including oxytocin cells.

## PVN activity and cortical plasticity

One important neural adaptation with motherhood occurs in the left auditory cortex, which is modulated by oxytocin to amplify responses to pup distress calls to promote retrieval<sup>15,21</sup>. Our results show that although virgin PVN did not respond directly to pup calls, virgin PVN and OT-PVN neurons were recruited by specific experiences during co-housing when abandoned pups would emit distress calls (Extended Data Fig. 6a, b). This might naturally pair distress calls with PVN activation and oxytocin signalling in target brain areas of virgin mice, including regions important for processing vocalizations such as left auditory cortex.

To examine how virgin PVN activation relates to cortical plasticity and eventual pup retrieval, we took recordings from virgin left auditory cortex and PVN using single-unit recordings and fibre photometry. One group of virgins was implanted with tetrodes in left auditory cortex, co-housed with experienced dams and pups for four days, and tested for retrieval on each day. As expected, single-unit responses to played-back pup calls increased during co-housing up to retrieval onset for each mouse (Extended Data Fig. 10a). Similar results were obtained in the second group of wild-type co-housed virgins, with GCaMP6s expression and fibre photometry in left auditory cortex (Fig. 4a–c). We found a progressive increase in photometric signal before the day of retrieval onset (Fig. 4b, c;  $P = 0.003$ ,  $n = 3$ ). Decreased responses 24

h after retrieval onset might reflect further refinement of cortical activity, perhaps via homeostatic mechanisms or inhibitory plasticity<sup>21</sup>, to prevent over-excitability once call-sensitive ensembles have formed.

**Fig. 4: PVN modulation and plasticity of left auditory cortex.**



**a**, Example photometry recording from virgin female left auditory cortex (AC). **b**, Change in photometry signal during co-housing in example animal. **c**, Summary of photometry data in AC. **d**, Schematic of simultaneous auditory cortex photometry and PVN single-unit recording during retrieval observation with a transparent barrier. **e**, Example single-trial auditory cortex photometry (insets, pup vocalizations) and PVN spiking before (left) and after (right) virgin began retrieving. Time is zero at dam pup retrieval. **f**, Summary of single-trial responses during dam retrievals. **g**, Schematic of photometry from virgin PVN projections to left AC during transparent-barrier observations. **h**, Photometric responses of auditory cortex from virgin females during observation. Time is zero at dam pup retrieval. **i**, Single-trial (grey) and per-animal responses (red). **j**, Effect of OTA or saline infusion in non-co-housed virgin left auditory cortex before dam retrieval observation. **k**, Optogenetic stimulation (opto stim) of left auditory cortex OT-ChETA fibres during opaque-barrier testing increased virgin retrieval. Data are mean  $\pm$  s.e.m.; \* $P < 0.05$ .

We next tested how virgin PVN activity related to emergence of cortical pup call responses in non-co-housed virgins. We took advantage of the trial-based structure of retrieval testing and observational testing, and performed simultaneous single-unit recordings from left PVN with photometry from left auditory cortex, together with audio recordings of pup vocalizations during dam pup retrieval (Fig. 4d–f). On some trials during observational testing, virgin PVN spiking transiently increased concomitantly with enhanced cortical responses to pup calls (Fig. 4e). Even before virgins began retrieving pups, single-trial PVN activity correlated with pup call-evoked responses in cortex (Fig. 4f;  $r = 0.5$ ,  $P = 0.00001$ ,  $n = 4$ ). We hypothesize that several episodes of this natural pup call pairing could persistently enhance cortical pup call responses. In this way, the newly parental auditory cortex becomes much more sensitive to pup distress, allowing emergence of rapid behavioural responses.

We tested whether onset of pup retrieval required signalling from virgin PVN to left auditory cortex, performing photometry of PVN→auditory cortex projections during observation of dam retrieval through the transparent barrier (Fig. 4g–i; Extended Data Fig. 10b). In virgins that began retrieving, PVN→auditory cortex projections were activated during observation of dam retrieval (Fig. 4h,i, learners,  $P = 0.02$ ). By contrast, PVN→auditory cortex projections were not significantly activated in virgins that did not learn after observation (Fig. 4h,i, non-learners,  $P = 0.5$ ).

Finally, we investigated whether PVN→auditory cortex input involved oxytocin signalling. Infusing a selective oxytocin receptor antagonist (OTA) into virgin left auditory cortex before transparent-barrier observational testing prevented virgins from learning to retrieve (Fig. 4j, saline, 7 out of 9 virgins retrieving; OTA, 2 out of 10 virgins retrieving;  $P = 0.03$ ). Optogenetic stimulation of OT-PVN→auditory cortex projections during observational testing with the opaque barrier enhanced retrieval

learning (Fig. 4k, ChETA-EYFP, 4 out of 4 virgins; EYFP, 2 out of 4 virgins; Mantel–Cox test,  $P=0.008$ ).

## Discussion

Most parental behaviours are likely to result from both innate and acquired components owing to their rapid onset, reliable performance and importance for species survival<sup>2,4,22,23,24,25,26</sup>. Here we examined how mouse alloparenting behaviour might be acquired by social transmission. We constructed an integrated system that combined continuous days-long videography with audio and neural activity monitoring, for recording documentary videos and subsequent behavioural discovery. We found that experienced mothers shepherd virgins towards pups in the nest but not an empty nest, ensuring that pups are protected by at least one co-caregiver. Virgins could also observe dams retrieving pups after displacement from the nest. This social learning depended on visual input, and we identified direct projections from visual layers of superior colliculus onto PVN oxytocin neurons. Oxytocin may act to amplify neural representations of pup cues such as distress calls in areas important for learned aspects of maternal care, including left auditory cortex. These valuable alloparenting experiences acquired during co-housing are likely to improve the quality of early maternal care when the virgins themselves have litters, and could occur throughout early life in animals living communally in the wild. Socially transmitted parental behaviour would then be useful for helping to ensure the survival and success of offspring in consociating species such as rodents and humans.

## Methods

All procedures were approved under NYU School of Medicine IACUC protocols.

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized. The behavioural raters were blinded to allocation during co-housing and outcome assessment.

### Behaviour

#### Co-housing

Pup-naive C57Bl/6 virgin female mice were bred and raised at NYU School of Medicine and kept isolated from dams and pups until used for these studies when approximately eight weeks old. For experiments where viral injections were

performed, we first allowed two weeks for viral expression before animals were used in experiments. Dams were initially pre-screened to ensure they behaved maternally, meaning that they retrieved pups and built nests; about 1% of dams did not retrieve pups and these animals were not used for co-housing. Naive virgins were initially pre-screened for retrieval or pup mauling before co-housing; around 5% of the naive virgins retrieved at least one pup or mauled pups during pre-screening and these mice were excluded from subsequent behavioural studies.

Co-housing of a virgin female with a mother and litter was conducted for 4–6 consecutive days in  $80 \times 40 \times 50$  cm plastic home cages. The floor was covered with abundant bedding material, food pellets and a pack of hydrogel for hydration placed in a corner of the bin and refreshed daily. Nesting material was also placed in the cage. We first placed the dam and her postnatal day 1 (P1) litter in the cage. After the dam was acclimatized for ~30 min, we introduced the virgin female with a tail mark for identification. Well-being of the adult mice and pups was monitored at least twice a day. A surveillance infrared camera system (Blackrock Microsystems) was positioned ~100 cm above the home cage to capture the entire surface. An ultrasonic microphone (Avisoft) was placed in the corner of the cage, ~10 cm above the nest. Two initial cages had a second camera placed on the side but these videos were not analysed for these experiments. For studies of spontaneous pup retrieval by dams and the influence of co-housing, some dams were singly housed with their litter but not with other adults.

In cases where co-housing was done only between a virgin female and pups (Fig. 1c), the pups were returned to the donor mother every 12 h (for at least 48 h) and immediately replaced with new pups. This was done to ensure that they stay alive and healthy despite not being fed during co-housing with the virgin. The procedure was repeated throughout the duration of the co-housing<sup>27</sup>.

### Pup retrieval testing

This test was used for the initial screening of dams and virgin female mice. In addition, outside of the spontaneous home cage behaviours, we specifically monitored pup retrieval every 24 h by the virgin females. We placed the female mouse to be tested in a behavioural arena ( $38 \times 30 \times 15$  cm) containing bedding and nesting material; the female was alone, without contact with other animals. Each animal was given 20 min to acclimatize before each testing session began. The entire litter (ranging from 3 to 7 P1–4 pups) were grouped in a corner of the arena and covered with nesting material, and the adult female given an additional 2 min of acclimatization (pup group size did not affect retrieval behaviour; Extended Data Fig. 2c). One pup was removed from the nest and placed in an opposite corner of the arena. The experimental female was given 2 min per trial to retrieve the displaced pup and

return it back to the nest; if the displaced pup was not retrieved within 2 min, the pup was returned to the nest and the trial was scored as a failure. If the pup was successfully retrieved, the time to retrieval was recorded and the trial was scored as a success. Another pup was then taken out of the nest, placed away from the nest (varying the position of the isolated pup relative to the nest from trial to trial), and the next trial was begun. After ten trials, pups were placed back into their home cage with their dam. We used an ultrasonic microphone (Avisoft) to verify that isolated pups vocalized during testing.

We reported probability of retrieving out of ten trials. Reliable retrieval was defined as having at least two out of ten successful trials. We used two-way ANOVA and Sidak's multiple-comparison test corrections to compare probability of retrieving in each group over days, and Student's *t*-test to compare the day of retrieval onset for each group.

### **Video and audio analysis**

Video and audio recordings were synchronized with the neuronal recordings, and then analysed with Adobe Audition and Avisoft. For video recordings we used the BORIS suite for scoring of behavioural observations. Three separate teams of independent scorers (two scorers from the Sullivan laboratory, three scorers from the Carcea laboratory and four scorers from the Froemke laboratory) were trained in a similar way on how to identify relevant individual and social behaviours during co-housing, and then scored the videos blind to the conditions. The results from each raster were compared and compiled, and results from each lab were cross-validated. Nest entry was considered the moment when the head of the animal entered the nest. Nest exit was considered the time when the rear of the animal left the nest. We used two-way ANOVA and Sidak's multiple-comparison test to compare pup retrieval rates and time in nest across days for each group.

Any event in which the dam chased the virgin towards the nest was identified as a shepherding event (that is, where distance from start to nest was greater than distance from end to nest). To determine the distance from nest during shepherding, we measured the distance from the bottom left corner of the cage to the position of the snout of the mouse, and to the position of the nest center. We then calculated distance from the virgin to nest. In cases of physical contact, start of shepherding was considered to be the moment when the dam made contact with the virgin, and the end of shepherding was the moment when the virgin stopped running. In some cases (especially later into co-housing), we noticed that virgins started running as soon as they noticed the dam approaching; in those cases, the start of shepherding was considered to be the moment when the virgins started running after the dam's approach. For Fig. 1*i*, we used paired *t*-tests to compare distance from start of

shepherding to nest with the distance from end of shepherding to nest. For Fig. 1j, we used one-sample Student's *t*-tests to determine if the daily frequency of shepherding was higher than 0.2 events per h (which was the average rate of dam–virgin chases in absence of pups). Audio recordings were processed in Adobe Audition, and isolation or distress calls were distinguished from adult calls and wriggling calls on the basis of the characteristic statistics (bout rate of 4–8 Hz and frequencies of 40–90 kHz).

## Observation of experienced retrievers

We first confirmed that virgins did not retrieve and dams retrieved at 100% at baseline. The exposures were done in standard behavioural arena ( $38 \times 30 \times 15$  cm). The virgin and dam were acclimatized for 20 min, then the nest with pups was transferred to this arena. After another 5–10 min, we manually isolated one pup at a time so that the dam would retrieve the pup back into the nest. We repeated this for ten times per session. In the experiments where either a transparent or an opaque divided the cage, the two adult animals were acclimatized on opposite sides of the barriers. After exposure, the adult animals were separated and the virgins were tested for pup retrieval 30 min later, as described above. As the preparation for testing and the acclimatization to the testing cage also took 30 min, this amounted to a total 60-min interval between virgin observation and testing of responses to isolated pups. The exposure was repeated for four sessions (one per day). A virgin that retrieved at least once during the four days of observation was considered as having acquired pup retrieval behaviour. We used chi-square exact tests to compare retrieval between conditions: wild-type mice with no barrier, wild-type mice with transparent barrier, wild-type mice with opaque barrier, and OXTR-KO virgins with transparent barrier.

## Surgery

### Viral injections

To test the effects of DREADDs, stereotaxic viral injections were performed in *Oxy-IRES-Cre* mice<sup>15</sup>. Mice were anaesthetized with 0.7–2.5% isoflurane (adjusted on the basis of scored reflexes and breathing rate during surgery), placed into a stereotaxic apparatus (Kopf), and bilateral craniotomies performed over PVN (from bregma: 0.72 mm posterior, 0.12 mm lateral). Injections were performed at a depth of 5.0 mm with a 5  $\mu$ l Hamilton syringe and a 33 gauge needle. Cre-inducible AAV2 hSyn::DIO-hM4D(Gi)-mCherry (University of North Carolina Viral Core) virus was injected into PVN at 0.1  $\mu$ l min<sup>-1</sup> for a final injection volume of 1.2–1.5  $\mu$ l. The craniotomy was sealed with a silicone elastomer (World Precision Instruments), the skin sutured. Animals were used for experiments after two weeks to allow for viral expression.

For monosynaptic retrograde tracing of inputs on OT-PVN neurons, we injected 0.5  $\mu$ l of a helper virus, AAV8-DIO-TVA-2A-oG (Salk Institute), in the left PVN of *Oxy-IRES-Cre mice*. Two weeks later, we injected 0.2  $\mu$ l of a pseudotyped rabies virus, EnvA G-deleted rabies-mCherry (Salk Institute) at the same location. After another two weeks, we perfused the animals and collected the brains for histology.

For optogenetic stimulation of superficial superior colliculus (sSC)→PVN projections, wild-type virgins were injected in the left sSC (from bregma: −3.08 mm posterior, −0.27 mm lateral, −1.9 mm ventral) with 0.5  $\mu$ l of either AAV9-pACAGW-ChR2-Venus or control AAV9-CAG-Venus virus. We then implanted a 400  $\mu$ m optic fibre in the left PVN.

For fibre photometry from auditory cortex, we performed viral injections into the left auditory cortex (~1.7 mm anterior from the occipital suture, 0.5 mm lateral from the temporal ridge, 1 mm ventral from pia), using a similar procedure. These coordinates corresponded with the ones previously published from our laboratory<sup>15</sup>: 2.9 posterior, 4.0 lateral (left) from bregma. We also oriented using local markers consisting in the branching patterns of the rhinal vein and of the middle cerebral artery, to target the auditory cortex as previously described<sup>28</sup>. We injected 1  $\mu$ l of AAV1 Syn::GCaMP6s (Addgene) at a titre of  $1 \times 10^{13}$  vg ml<sup>−1</sup> in the auditory cortex of wild-type mice. Following the injection, we implanted a 400  $\mu$ m optical fibre (ThorLabs) just above the auditory cortex, or inserted in the superficial layers of the cortex (200–300  $\mu$ m below the pial surface). For photometry from PVN→auditory cortex projection neurons, we injected 50 nl of AAVrg-hSyn1-GCaMP6s-P2A-nls-dTomato virus (Addgene) at three locations within the left auditory cortex. The virus titre was  $1.31 \times 10^{12}$  vg ml<sup>−1</sup>, and we injected at 10 nl min<sup>−1</sup>.

For optogenetically manipulating OT-PVN→auditory cortex projections, we injected the left PVN of Oxy-IRES-Cre virgins with 0.5  $\mu$ l of either AAV9-Ef1a-DIO-ChETA-EYFP or of AAV9-pCAG-FLEX-EYFP-WPRE. We then implanted a 400- $\mu$ m optic fibre in the left auditory cortex.

## Microdrive implantations

For in vivo single-unit electrophysiology we implanted microdrives either in the left PVN (Figs. 2–4) or the left auditory cortex (Fig. 4). We built microdrives using the parts and instructions for 4-tetrode Versadrives (Neuralynx), adapting these instructions for two bundles each made up of eight 12.5- $\mu$ m Nichrome wires. The day of the implantation, the wires were cleaned and gold plated to achieve impedances <500 k $\Omega$ . After the virgin female mice were anaesthetized with isoflurane, a craniotomy (1.5–2 mm in diameter) was performed above the target structure, and two additional small craniotomies were performed in the occipital bone and the right

parietal bone for insertion of bone screws. The ground and reference wires of the microdrives were soldered separately to these two bone screws. The dura was removed at the desired implantation site and the electrode bundles were slowly lowered to ~500  $\mu\text{m}$  above the target brain structure (4 mm ventral from pia for PVN). For recordings from the auditory cortex, we first acutely recorded multiunit activity with tungsten electrodes to localize auditory cortex during implantation procedure. Auditory cortex was identified with pure tones (60 dB SPL, 7–79 kHz, 50 ms, 1 ms cosine on–off ramps) delivered in pseudo-random sequence at 0.5–1 Hz. The craniotomy was covered with mineral oil and silicone elastomer, and the microdrive was secured to the skull using dental cement (C&B Metabond).

For optically identified recordings from OT-PVN units, we used *Oxt-IRES-Cre* mice bred into a C57BL/6 background. Prior to the microdrive implantation, we injected the PVN of these animals with 1–2  $\mu\text{l}$  of AAV1-CAG::DIO-ChR2 at a titre of  $1 \times 10^{13}$  vg  $\text{ml}^{-1}$  (in one mouse, we injected AAV2-EEF1::DIO-ChETA at a titre of  $1 \times 10^{13}$  vg  $\text{ml}^{-1}$ ). Then, after implanting and cementing the microdrive, we rotated the head of the animal at a 45°–50° angle around the anterior–posterior axis, and another craniotomy was done at 0.72 mm posterior and 3.5 mm left from bregma. A 400  $\mu\text{m}$  optical fibre for delivering blue light was slowly lowered at this position for 4.5 mm below the brain surface. For dual recordings, we implanted a microdrive in left PVN, injected the GCaMP virus (as described above) and implanted an optic fibre in the left auditory cortex during the same surgery.

## Optical fibre implantation

For fibre photometry we used 400- $\mu\text{m}$ -diameter 0.48NA optic fibres inserted into 2.5-mm-diameter ceramic ferrules (Doric Lenses). For recordings of PVN→auditory cortex, we used 1.25-mm metal ferrules, with fibres 1 mm long (below ferrule) for recordings in the auditory cortex and 5 mm long for recordings in PVN. They were implanted during the same surgery as the viral injections, and cemented to the skull as described above.

## Recordings

### Single-unit recordings during behaviour

Neuronal recordings were performed at 30 kHz sampling rate using the Cereplex  $\mu$  headstage, a digital hub and a neuronal signal processor (Blackrock Microsystems). The recordings were synchronized with a video recording system (Neuromotive, Blackrock). Before the start of the experiment the electrode bundles were lowered into the target structure and then advanced daily by ~70  $\mu\text{m}$ . For PVN recordings, optical

tagging tests were performed at the end of each co-housing day prior to lowering the electrodes for unit identification on the following day. Optical identification was performed in a separate clean cage in absence of pups or dam. The optical fibre was connected to a blue laser triggered using an analogue output from the recording system. The laser pulses were 5 ms in duration and delivered at either 2 or 5 Hz. Light intensity was controlled by adjusting the output of the laser, and at least three different intensities were tested in each mouse. The optical identification procedure was <5 min each day, and we used lower frequency (2 Hz) stimulation to minimize the likelihood of functional changes induced by optogenetic stimulation. To allow free movement of the implanted mice, we used a pulley combined with a fibre-optic and electrical rotary joint-commutator (Doric Lenses).

### Spike sorting

Neurons were isolated offline using Cerebrus and BOSS software (Blackrock Microsystems), as previously described<sup>29</sup>. A notch filter was applied to eliminate the line noise, then a high-pass filter of 250 Hz was applied, and a threshold multiplier was used to extract spike from signal energy between 1–5 kHz. We then used principal component analysis to extract spike features and to manually cluster them in different units. We eliminated spikes that violated a refractory period of 2–3 ms, and spikes that were simultaneously recorded in 10 or more channels (as potential artifacts). For tetrode recordings, we used a similar procedure in BOSS (Blackrock) that allowed triangulation of waveforms from the four tetrode channels. For identifying oxytocin neurons in optically tagged recordings, we aligned neuronal activity to the onset of the blue light pulse<sup>30</sup>. Neurons that reliably fired ( $\geq 70\%$  of trials) at short latency ( $\leq 4$  ms) after the onset of blue light were selected as being oxytocin neurons.

### Spike train analysis

To analyse changes in firing patterns, we aligned the spike trains to the onset of the behavioural episode and calculated the firing rate for the duration of the behaviour. In the case of ‘nest entry’, we used a 40-s cut-off, in order to minimize the impact from other behaviours that occur in the nest over longer time periods (for example, grooming, sleeping and nest building); this cut-off was chosen as most nest dwell times were  $\leq 40$  s (Extended Data Fig. 7b). For each behavioural episode we also calculated the corresponding baseline firing rate, either for an interval similar in duration to the behavioural episode (in the case of nest entry), or for 10 s before the start of the behavioural episode (for shepherding and dam retrievals). The change in firing rate was calculated as:

$(\text{Spiking}_{\text{behaviour}} - \text{Spiking}_{\text{baseline}}) \times 100 / (\text{Spiking}_{\text{behaviour}} + \text{Spiking}_{\text{baseline}})$ , where Spiking<sub>behaviour</sub> was firing rate during the behavioural event, and Spiking<sub>baseline</sub> was

firing rate during baseline. To determine whether these changes in firing rate were statistically significant, we shifted the behaviour intervals by a random value between  $-500$  and  $500$  s, and then calculated new firing rates and modulation percentages for 1,000 random shifts. All cells for which the modulation of activity in the unshifted data was greater than modulation in 950 shuffles were considered to be potentially significantly activated<sup>31</sup>. Cells for which the change in firing rate was smaller than the change in 950 shuffles were considered to be potentially significantly suppressed. For each unit, we computed the  $P$ -value from the permutation analysis and then applied a false-discovery rate correction for the multiple comparisons performed on each data subset (all OT-PVN or all PVN neurons recorded in a day), to identify neurons with significant changes in firing rate. The threshold for  $q$ -value was set 0.05, that is, 5% accepted false-discovery rate. To determine if the fraction of activated neurons is meaningful, we then applied a one-sided Fisher's exact test (with the null hypothesis of zero neurons being activated) with the threshold for alpha set at 0.05 (ref. <sup>32</sup>). To compare fractions of OT-PVN and PVN neurons activated by a specific behaviour, we applied a two-sided Fisher's exact test, and calculated the effect size as the relative risk: the probability of neurons being activated in the OT-PVN group/the probability of neurons being activated in the PVN group, using the Koopman asymptotic score for calculating 95% confidence intervals. To investigate how PVN activity correlated with activity in the auditory cortex, a population average change in PVN firing was calculated for each maternal-retrieval trial of the observational learning procedure.

To quantify unit activity during behavioural episodes, we aligned the neuronal data to the onset of the behaviour, and binned the spiking data in 250-ms bins for an interval between  $-20$  s and  $+40$  s ( $-10$  s to  $+20$  s in the case of shepherding episodes). We then calculated the  $z$ -score value for each bin in the chosen time window; for display purposes, we used bounds of  $-3$  (minimum  $z$ -score) and  $+3$  (maximum  $z$ -score). For responses to pup call playback, spike trains from single-unit recordings in PVN or auditory cortex were aligned to the onset of the played-back vocalization. The average firing rate during the pup call (1-s interval) was normalized to the firing rate during baseline (1 s preceding pup call onset).

To measure the cross-correlation for simultaneously recorded OT-PVN and PVN single-units during co-housing (Extended Data Fig. <sup>7d,e</sup>), we calculated the normalized zero-lag correlation of spike counts (in 10-ms bins) across all pairs of units<sup>33</sup>, separately for each of the three behavioural conditions (shepherding, nest entry and maternal retrievals). We then averaged  $z$ -transformed  $r$  values across all events of each behaviour, to obtain a single cross-correlation value for each pair and each behaviour. Higher correlation  $r$  values reflect higher synchrony within the pair.

## Fibre photometry

To perform fibre photometry, we connected the optical fibre implanted in the auditory cortex and housed in a ceramic ferrule to a custom-built photometry rig<sup>34</sup>. A 400-Hz sinusoidal blue light was delivered via the optical fibre from an LED (30 µW) for GCaMP6s excitation. We collected the emitted (green) light via the same optical fibre, used a dichroic mirror and appropriate filters to direct emitted light to a femtowatt silicon photoreceiver (Newport), and recorded using a real-time processor (RX8, TDT). The analogue readout was then low-pass filtered at 20 Hz. The intensity of blue light excitation was adjusted to produce similar baseline fluorescence levels across sessions in the same mouse. The sound processor used for delivering the pup calls (RZ6, TDT) was synchronized with the fibre photometry system.

For investigating changes in neuronal activity of the left auditory cortex throughout co-housing, we recorded responses to six different played-back pup calls every 24 h. To avoid possible changes in the magnitude of calcium transients with changes in the position of the mouse relative to the speaker, in some recordings we restrained the mouse in a mesh cup, at the same distance from the speaker. Data from restrained or unrestrained mice showed similar trends, and were thus pooled together. To determine the evoked response recorded with fibre photometry, we calculated the  $\Delta F/F$  for each pup call interval.

## Perturbations

### Chemogenetic suppression of oxytocin neurons during co-housing

*Oxy-IRES-Cre* virgin female mice expressing in OT-PVN cells either DREADD coupled to inhibitory G-protein (DREADDi, also known as hM4D(Gi)) and fused with mCherry, or mCherry alone (as control), were co-housed with mothers and pups as described above. In both experimental (*Oxy-IRES-Cre* virgins expressing DREADDi-mCherry) and control (*Oxy-IRES-Cre* virgins expressing mCherry) cohorts, clozapine-N-oxide (CNO) was administered in the drinking water at a concentration of 25 mg l<sup>-1</sup> (ref. 35). Five sugar pellets were added in each bottle to obscure the taste of CNO. At this concentration, for the average mouse body weight (~30 g) and average daily water consumption in mice (~6 ml), the average CNO consumption was ~5 mg kg<sup>-1</sup> per day.

Three out of seven mice in the mCherry<sup>+</sup> CNO cohort had DREADDi virus expressed outside of the PVN due to mis-targeting. These mice were also treated with CNO; as their behaviour was similar to the mCherry-expressing mice treated with CNO—we combined data from these groups together in Fig. 2a. For each animal, the retrieval probability for the first two days of co-housing was averaged. For Extended Data Fig. 4b, retrieval testing was done in animals that stably retrieved at baseline and as such were not subjected to the co-housing procedure.

## Optogenetic stimulation

For optogenetics, we connected the optical fibre implanted either in PVN or in the left auditory cortex of the virgin mouse to a blue light laser (OptoEngine). Light pulses (5 ms) were delivered at 20 Hz (for Extended Data Fig 9; ref. 21) or 30 Hz (for Fig. 4; ref. 15), while the dam performed 10 retrieval trials across an opaque barrier from the virgin. To minimize potential adaptation or synaptic depression, stimulation was performed for  $\leq 5$  min, enough time for dams to complete 10 retrievals. After 30 min, the non-co-housed virgin was tested for retrieval. This procedure was repeated once a day for four consecutive days.

## Drug infusion

To determine the role of cortical oxytocin receptor activation during observation of pup retrieval, we implanted cannulas in the left auditory cortex of naive virgin females<sup>15</sup>. After mice recovered from surgery, we placed them in the observational chamber, and then infused either saline or the oxytocin receptor antagonist OTA (1  $\mu$ M solution in saline) through the cannula. We injected a volume of 1.5  $\mu$ l at a rate of 1  $\mu$ l min<sup>-1</sup>. We removed the internal from the cannula five minutes after the end of injection to allow the solution to diffuse, and then allowed the virgins to acclimatize to the observational chamber for another 10 min. After this, we exposed virgins to maternal retrievals across the transparent barrier as described above. Drug conditions were compared with a survival log-rank analysis to determine effects on retrieval onset across groups.

## Immunohistology

To verify viral expression at the end of the experiments, animals were fixed with transcardiac perfusions of 4% paraformaldehyde. Brains were removed and further preserved in a paraformaldehyde solution for 1–2 h at 4 °C. Afterwards, brains were sequentially cryopreserved in 15% and then 30% sucrose solution, embedded in OCT solution and sectioned (30  $\mu$ m) with a cryostat then mounted on positively charged slides. Immunohistochemistry was performed on the mounted sections as previously described<sup>29,36</sup>. Sections were blocked in 5% goat serum solution for 1 h at room temperature or for 14 h at 4 °C. A solution of the appropriate primary antibodies was diluted in 1% goat serum and 0.01% Triton X solution and then applied for 24 h at 4 °C. We used a rabbit anti-oxytocin antibody (EMD Millipore, 1:500), mouse anti-oxytocin antibody (a gift from H. Gainer at national Institutes for Health), a chicken anti-GFP (Aves, 1:500) and a chicken anti-mCherry (Abcam, 1:1,000) antibody. The sections were washed in PBS solution, and a solution of fluorophore-conjugated secondary antibodies applied for 1.5 h at room temperature. All secondary antibodies

were from Jackson Immunoresearch and used at 1:200. Slides were examined and imaged using a Carl Zeiss LSM 700 confocal microscope with four solid-state lasers (405/444, 488, 555 and 639 nm) and appropriate filter sets. For imaging sections co-stained with multiple antibodies, we used short-pass 555 nm (Alexa Fluor 488), short-pass 640 nm (Alexa Fluor 555), and long-pass 640 nm (Alexa Fluor 647) photomultiplier tubes.

## Micro-computed tomography imaging

The localization of the implanted electrodes was assessed *in vivo* using  $\mu$ CT scans in post-implanted mice, followed by co-registration with an online digital MRI mouse brain atlas. The  $\mu$ CT datasets were acquired using the  $\mu$ CT module of a MultiModality hybrid micro-Positron Emission Tomography– $\mu$ CT Inveon Scanner (Siemens Medical Solutions). The Inveon scanner is equipped with a 165 mm  $\times$  165 mm x-ray camera and a variable-focus tungsten anode x-ray source operating with a focal spot size of less than 50  $\mu$ m. The scan consisted of a 20-min whole-head acquisition over an axial field of view of 22 mm and a transaxial of 88 mm with a resolution of 21.7  $\mu$ m pixels binned to 43.4  $\mu$ m. 440 projections were acquired using a 1-mm aluminum filter, a voltage of 80 kV, and a current of 500  $\mu$ A. The datasets were reconstructed using the Feldkamp algorithm<sup>37</sup>.

The hybrid scanner was equipped with a M2M Biovet (Cleveland) module used to monitor continuously vital signs. All mice were monitored continuously throughout the scanning session via a respiration sensor pad (SIMS Graseby). The imaging scan consisted of initially placing each mouse in an induction chamber using 3–5% isofluorane exposure for 2–3-min until the onset of anesthesia. The mouse was then subsequently positioned laterally along the bed palate over a thermistor heating pad in which 1.0% to 1.5% isofluorane was administered via a 90° angled nose cone throughout the scan. The head of each subject was judiciously oriented perpendicular to the axis of the mouse body so that the extracranial part of the implanted electrode could be easily kept away from the field of view of the  $\mu$ CT image acquisition. Importantly, the large extracranial metal components and dental cement of the implant can cause beam hardening that can appear as cupping, streaks, dark bands or flare in the  $\mu$ CT<sup>38,39,40</sup>. To this effect, the head positioning helped reduce the risks of image artifacts that could be induced by the implant along the path of the X-ray beam.

Unlike MRI,  $\mu$ -CT imaging can be performed on subjects with metal implants. However, lack of the soft tissue contrast of the  $\mu$ CT limited its usefulness to provide the needed brain anatomical detail in order to verify the electrode correct localization. Our approach combined registration of post-implant  $\mu$ CT with an existing online MRI brain atlas for adult C57Bl/6 mice from the Mouse Imaging Centre: (<https://wiki.mouseimaging.ca/display/MICePub/Mouse+Brain+Atlases>).

The three-dimensional MRI mouse brain atlas was established by acquiring 40 individual ex vivo mice using T2-weighted sequence on a 7-Tesla scanner. All the data were averaged and resulted into a  $40\text{ }\mu\text{m}$  isotropic resolution dataset detailed in Dorr et al. (ref. <sup>41</sup>). The hypothalamic PVN was manually segmented and color-coded, with the guidance of the P56 coronal Allen mouse histology brain atlas<sup>42</sup>. This region was set as the target of reference. A rigid co-registration between  $\mu\text{CT}$  and the modified MRI atlas images was systematically performed using a commercial software Amira (Thermo Fisher Scientific). Both datasets were then overlaid for visual analysis and to determine the sub-millimetric localization of the electrode tip.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

Data and statistical tests are available in the supplementary data sheet provided with this manuscript. The video data generated during this study are available in Databrary (<https://nyu.databrary.org/volume/1333/slot/55145/>). Other data are available in Zenodo (<https://zenodo.org/record/5039601#.YNvaYEwpCE5>) and DANDI (<https://gui.dandiarchive.org/#/dandiset/000114>). The NYU Data Catalog record will point to the most updated locations of this dataset (<https://datacatalog.med.nyu.edu/dataset/10449?tak=160dbb2281acf5>). [Source data](#) are provided with this paper.

## Code availability

The Matlab code is available at github (<https://github.com/ic283/Carcea-et-al.-2021>).

## References

1. 1.

Insel, T. R. & Young, L. J. The neurobiology of attachment. *Nat. Rev. Neurosci.* **2**, 129–136 (2001).

2. 2.

Dulac, C., O'Connell, L. A. & Wu, Z. Neural control of maternal and paternal behaviors. *Science* **345**, 765–770 (2014).

3. 3.

Feldman, R. The neurobiology of human attachments. *Trends Cogn. Sci.* **21**, 80–99 (2017).

4. 4.

Opendak, M., Gould, E. & Sullivan, R. Early life adversity during the infant sensitive period for attachment: programming of behavioral neurobiology of threat processing and social behavior. *Dev. Cogn. Neurosci.* **25**, 145–159 (2017).

5. 5.

Wilson, E. O. *Sociobiology: The New Synthesis* (Harvard Univ. Press, 1975).

6. 6.

Moehlman, P. D. Jackal helpers and pup survival. *Nature* **277**, 382–383 (1979).

7. 7.

Lee, P. C. Allomothering among African elephants. *Animal Behavior* **35**, 278–291 (1987).

8. 8.

Hotelling, B. A. Newborn capabilities: parent teaching is a necessity. *J. Perinat. Educ.* **13**, 43–49 (2004).

9. 9.

Insel, T. R. & Shapiro, L. E. Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *Proc. Natl Acad. Sci. USA* **89**, 5981–5985 (1992).

10. 10.

Nishimori, K. et al. Oxytocin is required for nursing but is not essential for parturition or reproductive behavior. *Proc. Natl Acad. Sci. USA* **93**, 11699–11704 (1996).

11. 11.

Johnson, Z. V. et al. Oxytocin receptors modulate a social salience neural network in male prairie voles. *Horm. Behav.* **87**, 16–24 (2017).

12. 12.

Burkett, J. P. et al. Oxytocin-dependent consolation behavior in rodents. *Science* **22**, 375–378 (2016).

13. 13.

Kendrick, K. M. et al. Neural control of maternal behavior and olfactory recognition of offspring. *Brain Res. Bull.* **44**, 383–395 (1997).

14. 14.

Okabe, S. et al. Pup exposure facilitates retrieving behavior via the oxytocin neural system in female mice. *Psychoneuroendocrinology* **79**, 20–30 (2017).

15. 15.

Marlin, B. J. et al. Oxytocin enables maternal behavior by balancing cortical inhibition. *Nature* **520**, 499–504 (2015).

16. 16.

Dhawale, A. K. et al. Automated long-term recording and analysis of neural activity in behaving animals. *eLife* **6**, e27702 (2017).

17. 17.

Rich, M. E. et al. Impairments in the initiation of maternal behavior in oxytocin receptor knockout mice. *Plos ONE* **9**, 0098839 (2014).

18. 18.

Ito, S., Feldheim, D. A. & Litke, A. M. Segregation of visual response properties in the mouse superior colliculus and their modulation during locomotion. *J. Neurosci.* **37**, 8428–8443 (2017).

19. 19.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. Understanding motor events: a neurophysiological study. *Exp. Brain Res.* **91**, 176–180 (1992).

20. 20.
- Bonini, L. & Ferrari, P. F. Evolution of mirror systems: a simple mechanism for complex cognitive functions. *Ann. NY Acad. Sci.* **1225**, 166–175 (2011).
21. 21.
- Schiavo, J. K. et al. Innate and plastic mechanisms for maternal behavior in auditory cortex. *Nature* **587**, 426–431 (2020).
22. 22.
- Yang, T. & Shah, N. M. Molecular and neural control of sexually dimorphic social behaviors. *Curr. Opin. Neurobiol.* **38**, 89–95 (2016).
23. 23.
- Yoshihara, C., Numan, M. & Kuroda, K. Oxytocin and parental behaviors. *Curr. Top. Behav. Neurosci.* **35**, 119–153 (2018).
24. 24.
- Lehrman, D. S. A critique of Konrad Lorenz's theory of instinctive behavior. *Quart. Rev. Biol.* **28**, 4 (1953).
25. 25.
- Scott, N., Prigge, M., Yizhar, O. & Kimchi, T. A sexually dimorphic hypothalamic circuit controls maternal care and oxytocin secretion. *Nature* **525**, 519–522 (2015).
26. 26.
- Remedios, R. et al. Social behaviour shapes hypothalamic neural ensemble representations of conspecific sex. *Nature* **550**, 388–392 (2017)
27. 27.
- Fleming, A. S. & Rosenblatt, J. S. Maternal behavior in the virgin and lactating rat. *J. Comp. Physiol. Psychol.* **86**, 957–972 (1974).
28. 28.

Hackett, T. A., Barkat, T. R., O'Brien, B. M., Hensch, T. K. & Polley, D. B. Linking topography to tonotopy in the mouse auditory thalamocortical system. *J. Neurosci.* **31**, 2983–2995 (2011).

29. 29.

Cacea, I., Insanally, M. N. & Froemke, R. C. Dynamics of auditory cortical activity during behavioral engagement and auditory perception. *Nat. Commun.* **8**, 14412 (2017).

30. 30.

Hangya, B., Ranade, S. P., Lorenc, M. & Kepecs, A. Central cholinergic neurons are rapidly recruited by reinforcement feedback. *Cell* **162**, 1155–1168 (2015).

31. 31.

Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N. & Harvey, C. D. Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell* **170**, 986–999 (2017).

32. 32.

Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).

33. 33.

Shapiro, M. L. & Ferbinteanu, J. Relative spike timing in pairs of hippocampal neurons distinguishes the beginning and end of journeys. *Proc. Natl Acad. Sci. USA* **103**, 4287–4292 (2006).

34. 34.

Falkner, A. L., Grosenick, L., Davidson, T. J., Deisseroth, K. & Lin, D. Hypothalamic control of male aggression-seeking behavior. *Nat. Neurosci.* **19**, 596–604 (2016).

35. 35.

Zhan, J., Komal, R., Keenan, W. T., Hattar, S. & Fernandez, D. C. Non-invasive strategies for chronic manipulation of DREADD-controlled neuronal activity. *J. Vis. Exp.* **150**, 59439 (2019)

36. 36.

Carcea, I. et al. Maturation of cortical circuits requires semaphorin 7A. *Proc. Natl Acad. Sci. USA* **111**, 13978–13983 (2014).

37. 37.

Feldkamp, L. A., Davis, L. C. & Kress, J. W. Practical cone-beam algorithm. *J. Opt. Soc. Am. A* **1**, 612–619 (1984).

38. 38.

Barrett, J. F. & Keat, N. Artifacts in CT: recognition and avoidance. *Radiographics* **24**, 1679–1691 (2004).

39. 39.

Duerinckx, A. J. & Macovski, A. Polychromatic streak artifacts in computed tomography images. *J. Comput. Assist. Tomogr.* **2**, 481–487 (1978).

40. 40.

Joseph, P. M. & Spital, R. D. A method for correcting bone induced artifacts in computed tomography scanners. *J. Comput. Assist. Tomogr.* **2**, 100–108 (1978).

41. 41.

Dorr, A. E., Lerch, J. P., Spring, S., Kabani, N. & Henkelman, R. M. High resolution three-dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice. *NeuroImage* **42**, 60–69 (2008).

42. 42.

Dong, H. W. *The Allen Reference Atlas: A Digital Color Brain Atlas of the C57Bl/6J Male Mouse* (Wiley, 2008).

## Acknowledgements

We thank G. Buzsaki, E. Glennon, V. Grinevich, B. Hangya, K. Kuchibhotla, M. A. Long, J. Minder, J. K. Schiavo, R. Tremblay, R. W. Tsien and D. Vallentin for comments, discussions and technical assistance; and C. A. Loomis and NYU School of Medicine Histology Core for assistance with anatomical studies. Oxytocin-IRES-Cre mice were obtained from The Jackson Laboratory. Oxytocin receptor knockout

mice were obtained from R. W. Tsien. S. E. Ross created artwork in Figs. 1a, 3e, 4d, g, j, k, Extended Data Fig. 9e. This work was funded by NICHD (HD088411 to R.C.F. and R.M.S.), NIDCD (DC12557 to R.C.F.), the BRAIN Initiative (NS107616 to R.C.F., A.M., and D.L.), NIMH (K99/R00 MH106744 to I.C.; F32 MH112232 to M.O., and T32 MH019524 to B.J.M.), the Strategic Program for Brain Sciences from Japan (AMED 16K15698 to S.H. and K.N.), a McKnight Scholarship (R.C.F.), a Pew Scholarship (R.C.F.), a Howard Hughes Medical Institute Faculty Scholarship (R.C.F.); and NARSAD Young Investigator awards (I.C. and M.O.). In vivo imaging was performed under the DART Preclinical Imaging Core partially funded by the NYU Laura and Isaac Perlmutter Cancer Center Support Grant, NIH/NCI P30CA016087. The Center for Advanced Imaging Innovation and Research (CAI2R, [www.cai2r.net](http://www.cai2r.net)) at NYU School of Medicine is supported by NIH/NIBIB P41 EB017183.

## Author information

### Author notes

1. Bianca J. Marlin

Present address: Department of Psychology, Zuckerman Neuroscience Institute, Columbia University, New York, NY, USA

2. Annegret Falkner

Present address: Princeton Neuroscience Institute, Princeton, NJ, USA

## Affiliations

1. Skirball Institute for Biomolecular Medicine, New York University School of Medicine, New York, NY, USA

Ioana Carcea, Naomi López Carballo, Bianca J. Marlin, Rumi Ooyama, Joyce M. Mendoza Navarro, Veronica E. Diaz, Luisa Schuster, Maria I. Alvarado Torres, Harper Lethin, Daniel Ramos, Jessica Minder, Chloe J. Bair-Marshall & Robert C. Froemke

2. Neuroscience Institute, New York University School of Medicine, New York, NY, USA

Ioana Carcea, Naomi López Carballo, Bianca J. Marlin, Rumi Ooyama, Joyce M. Mendoza Navarro, Maya Opendak, Veronica E. Diaz, Luisa Schuster, Maria I. Alvarado Torres, Harper Lethin, Daniel Ramos, Jessica Minder, Chloe J. Bair-

Marshall, Annegret Falkner, Dayu Lin, Adam Mar, Regina M. Sullivan & Robert C. Froemke

3. Department of Otolaryngology, New York University School of Medicine, New York, NY, USA

Ioana Carcea, Naomi López Caraballo, Bianca J. Marlin, Joyce M. Mendoza Navarro, Veronica E. Diaz, Luisa Schuster, Maria I. Alvarado Torres, Harper Lethin, Daniel Ramos, Jessica Minder, Chloe J. Bair-Marshall & Robert C. Froemke

4. Department of Neuroscience and Physiology, New York University School of Medicine, New York, NY, USA

Ioana Carcea, Naomi López Caraballo, Bianca J. Marlin, Joyce M. Mendoza Navarro, Maya Opendak, Veronica E. Diaz, Luisa Schuster, Maria I. Alvarado Torres, Harper Lethin, Daniel Ramos, Jessica Minder, Chloe J. Bair-Marshall, Annegret Falkner, Dayu Lin, Adam Mar, Regina M. Sullivan & Robert C. Froemke

5. Department of Pharmacology, Physiology and Neuroscience, Brain Health Institute, Rutgers, The State University of New Jersey, New Jersey, USA

Ioana Carcea, Rumi Ooyama & Grace H. Samadjopoulos

6. Department of Animal Science and Biotechnology, Azabu University, Sagamihara, Kanagawa, Japan

Rumi Ooyama, Takefumi Kikusui & Kazutaka Mogi

7. Department of Neuroscience, Mount Sinai School of Medicine, New York, NY, USA

Justin S. Riceberg

8. Emotional Brain Institute, Nathan Kline Institute for Psychiatric Research, Orangeburg, NY, USA

Maya Opendak & Regina M. Sullivan

9. Department of Child and Adolescent Psychiatry, New York University School of Medicine, New York, NY, USA

Maya Opendak & Regina M. Sullivan

10. Center for Advanced Imaging Innovation and Research, Bernard and Irene Schwartz Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine, New York, NY, USA

Sebastian L. Mendoza & Youssef Z. Wadghiri

11. Department of Obesity and Internal Inflammation, Fukushima Medical University, Fukushima, Japan

Shizu Hidema & Katsuhiko Nishimori

12. Department of Psychiatry, New York University School of Medicine, New York, NY, USA

Annegret Falkner & Dayu Lin

## Contributions

I.C., N.L.C., R.O., J.M.M.N., J.S.R., M.O., M.I.A.T., H.L., J.M., A.M., planned and conducted behavioural experiments; I.C., R.O., M.I.A.T., D.R., G.H.S., M.O., R.M.S. planned and conducted analysis of video recordings; I.C., N.L.C., R.O., B.J.M., V.E.D., L.S., C.J.B.M., A.F., D.L., planned and conducted neuronal and photometry recording experiments; I.C. and J.S.R. conducted analysis of neuronal recordings; S.L.M. and Y.Z.W performed and analysed CT and MRI imaging; I.C., R.O., M.I.A.T., performed histology and microscopy. I.C. and R.C.F. conceived the study and wrote the paper. All authors discussed results and edited the manuscript.

## Corresponding authors

Correspondence to [Ioana Carcea](#) or [Robert C. Froemke](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Zoe Donaldson, Aaron Wilber and Stephen Shea for their contribution to the peer review of this work.

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

## Extended data figures and tables

### Extended Data Fig. 1 Behavioural ethograms from 3 pairs of animals.

Ethograms were constructed from visual analysis of continuous four-day videography by two to four independent observers. Dashed vertical lines indicate end of each day.

### Extended Data Fig. 2 Pup retrieval over animals.

**a**, Retrieval probabilities of all virgins co-housed with just pups (D-, left) or with experienced dam and pups (D+, right). **b**, Summary of virgin retrieval probabilities over days, color indicates retrieval success each day. **c**, Dam retrieval probability uniformly high across dams with different litter sizes (N=3 dams,  $p>0.9$ , Kruskal-Wallis test). Error bars, SEM.

### Extended Data Fig. 3 Shepherding behaviour.

**a**, Average latency for dams to shepherd virgin to the nest (time 0 is virgin nest exit). Latency decreased over days of co-housing (day 1:  $155.9\pm34$  s; day 2:  $136.1\pm17.94$  s; day 3:  $75.7\pm5.5$  s; day 4:  $93.9\pm8.5$  s; n=676 events from N=9 dam-virgin pairs,  $p<0.0001$ , one-way ANOVA, Holm-Sidak's multiple-comparison test). Red lines, medians. Error bars, SEM. **b**, Left, ethograms from two separate virgins, one co-housed just with dam without pups (top), the other co-housed with dam and pups (bottom). Note differences between time in nest and amount of shepherding when pups were present. Right, summary of shepherding events per hour when pups were removed vs when pups were present ('No pups':  $0.2\pm0.1$  events/hour; 'Pups':  $3.5\pm0.9$  events/hour, N=4,  $p=0.03$ , Student's two-tailed paired t-test). Error bars, SEM. **c**, Locations of start (yellow) and end (orange) positions of dam→virgin chasing events without pups in the cage (N=4 dyads), relative to nest center (nest radius was ~10 cm, gray).  $46.1\pm10.4\%$  of dam→virgin chases ended in the nest, indicating that the frequency of these chases was significantly reduced but the direction of the virgins was comparable with or without pups (all chasing events ending in nest,  $p=0.9$ , two-sided Fisher's exact test).

### Extended Data Fig. 4 Controls for DREADDi-based impairment of oxytocin system during acquisition of pup retrieval behaviour.

**a**, Recordings in female mouse PVN brain slices. Left, current-clamp whole-cell recording from OT-PVN neuron expressing DREADDi-mCherry before (top) and after (bottom) CNO bath administration. Right, summary of chemogenetic suppression of OT-PVN neuronal firing in brain slices ('hM4D(Di)-mCherry' baseline firing rate before CNO:  $8.8 \pm 1.7$  Hz, after CNO:  $0.6 \pm 0.1$  Hz, n=7 cells from N=4 mice; 'mCherry' baseline firing rate before CNO:  $4.8 \pm 0.6$  Hz, after CNO:  $5.0 \pm 0.9$  Hz, n=8 cells from N=3 mice; two-way ANOVA with Sidak's correction for multiple comparisons, p=0.0002 between baseline and CNO for hM4D(Di)-mCherry, p=0.9 between baseline and CNO for mCherry). Error bars, SEM. **b**, CNO in retrieving females that expressed DREADDi did not affect retrieval rates (baseline retrieval:  $76.0 \pm 9.2\%$ , N=5; CNO:  $72.0 \pm 18.8\%$ , N=5; saline:  $92.0 \pm 8.0\%$ , N=5; repeated measures ANOVA, p=0.332). Error bars, SEM.

### Extended Data Fig. 5 Confirmation of opsin expression and targeting for optically tagged OT-PVN recordings.

**a**, Example immunostained tissue sections of PVN showing co-labeling of oxytocin peptide (red) and ChETA channelrhodopsin 2 conjugated to EGFP (green) together with DAPI labeling for cell density (blue). Insets show 2x magnified images to highlight the co-localization of red and green staining (arrows) in PVN cells. 84.8±7.2% of cells (N=6 animals) expressing ChETA also expressed oxytocin peptide. **b**, Verification of electrode locations in PVN with combined MRI/μCT. μCT imaging was performed after implantation and subsequently co-registered with a mouse MRI atlas to localize electrode bundles within PVN, as well as the separate fibre optic implanted for optogenetic identification of oxytocin neurons. Electrodes were then lowered at the end of each day of co-housing. Yellow/orange, μCT imaging of electrodes in situ. Images shown are from different animals.

### Extended Data Fig. 6 PVN neurons in adult female virgin mice responded more robustly during interactions with dams than with pups.

**a**, Example of OT-PVN neuron instantaneous firing (top), synchronized with homecage behaviour (gray bar marks a nest entry episode), and with pup vocalizations (bottom), illustrates the different data streams recorded with our system. **b**, Example spectrograms of pup vocalizations and distress calls in home cage during shepherding and nest entry. **c**, Normalized (z-scored) spiking for 248 PVN neurons recorded during 5-minute interaction periods with pup (left) or a dam (right). Recordings sorted by peak activity. **d**, Proportion of activated (purple; z-score > 1.645, i.e. p < 0.05) and suppressed (gray, z-score < -1.645) PVN cells. During the 5-minute interaction periods, interactions with dam activated more virgin PVN cells than interactions with pups (with dam: 25/248, with pup: 7/248, p=0.001, two-sided Fisher's exact test). After dam interactions, PVN cells continued to fire more than after interactions with

pup (with dam: 30/248, with pup: 12/248,  $p=0.005$ ). **e**, Social sniffing between the virgin and dam activated PVN neurons (9/43 neurons significantly activated while the virgin was sniffed,  $p=0.002$ ; 11/43 neurons activated while the virgin sniffed the other adult,  $p=0.0005$ ). A similar fraction of PVN neurons was activated during digging (8/43,  $p=0.005$ ). Only a few PVN neurons were activated by entering an empty nest without pups (2/43,  $p=0.4$ ). Dark bars, significantly modulated neurons. The proportion of PVN neurons activated during empty nest entry was smaller than the proportion of PVN neurons activated while being sniffed (two-sided Fisher's exact test,  $p=0.04$ ), during sniffing ( $p=0.01$ ), and after entering a nest with pups ( $p=0.001$ ). Fraction of cells responding during empty nest entry was also smaller than the proportion of OT-PVN neurons activated during shepherding (0.03), dam retrieval ( $p=0.01$ ), and after entering a nest with pups ( $p=0.01$ ).

### **Extended Data Fig. 7 PVN unit activity related to specific interactions with dams or pups.**

**a**, Patterns of OT-PVN (top) and PVN (bottom) cell activity during shepherding (left; 5/21 OT-PVN neurons, Fisher's exact test,  $p=0.02$ ; 9/97 PVN neurons,  $p=0.001$ ), nest entry (middle; 6/21 OT-PVN neurons,  $p=0.01$ ; 33/126 PVN neurons,  $p=0.0001$ ; shepherding vs voluntary entry,  $p=0.005$ , Wilcoxon signed-rank test) and dam retrieval (right; 5/17 OT-PVN neurons,  $p=0.02$ ; 10/69 PVN neurons,  $p=0.0007$ ). Recordings were aligned to start of interaction, and sorted by peak activity. **b**, Distribution of durations of time in nest for virgins after entry events. ~70% of all nest dwell times are  $\leq 40$  s. **c**, Left, example PSTH for PVN unit during play-back of pup distress calls (sonogram on top). Right, summary of PVN activity during pup call presentation. Almost all cells in virgin PVN were unresponsive, with only 1/71 PVN units having significant responses to pup vocalizations. **d**, Examples of simultaneously recorded unit spike trains aligned to onset of individual shepherding, nest entry, and dam retrieval events during co-housing. Blue, OT-PVN units. Yellow, periods of coincidental activity within 10 ms. **e**, Summary of average zero-lag pair correlations for each pair of simultaneously recorded units during each co-housing behaviour. Neuronal pairs including OT-PVN neurons (OT-PVN:OT-PVN and OT-PVN:PVN pairs) were more strongly correlated than PVN:PVN pairs (shepherding,  $n=530$  pairs,  $p<0.0001$ ; nest entry,  $n=776$  pairs,  $p<0.0001$ ; dam retrieval,  $n=278$ ,  $p<0.0001$ ; ANOVA and Tukey's multiple-comparison test). Error bars, SEM.

### **Extended Data Fig. 8 Many virgins seem to learn retrieval from maternal demonstration.**

**a**, Frequency of self-generated retrievals by dams. Top, example ethogram depicting two separate spontaneous retrievals by the dam (blue). Bottom, summary of individual dam spontaneous hourly pup retrieval rate per day (day one:  $0.6\pm0.2$  events/hour, day

two:  $0.3 \pm 0.1$  events/hour, day three:  $0.4 \pm 0.2$  events/hour, day four:  $0.4 \pm 0.2$  events/hour, N=14). Lines, individual dam behaviour. Filled circles, daily average across animals. Dashed line, average rate of spontaneous retrieval for singly housed dams and litters without other adults ( $0.2 \pm 0.1$  events/hour, N=5). Error bars, SEM. **b**, Virgin PVN responses when mother brought pups to virgin. For these animals, three separate pup delivery episodes occurred in a twenty-minute period after the second day of co-housing. On each of those occurrences, this PVN unit recorded in the virgin responded with increased spiking for several seconds until the dam returned to retrieve the pup herself. Top, raster of each of these three trials. Bottom, summary PSTH. Inset, spike waveforms; scale: 0.4 ms, 0.5 mV. **c**, Pup activity for non-co-housed virgins tested across transparent barrier, separated by virgins that began retrieving after observation ('L') and virgins that did not begin retrieving after observation ('NL'). Left, pup movement (p=0.64, Student's unpaired two-tailed t-test, N=16 animals: 3 learners, 13 non-learners). Middle, number of pup vocalizations (p=0.07). Right, pup call mean frequency (p=0.77). Error bars, SEM. **d**, Onset of retrieval behaviour was faster in co-housed virgins relative to non-co-housed virgins exposed to retrievals behind transparent barrier or with no barrier (Mantel-Cox test, p=0.004, N=48 mice total). \*\*, p<0.01.

### Extended Data Fig. 9 Superior colliculus input to PVN enhances acquisition of pup retrieval.

**a**, In *Oxy-IRES-Cre* virgins, we injected an AAV helper virus into PVN, followed two weeks later by injecting a rabies virus that carried mCherry fluorescent protein. Left, retrograde monosynaptic input tracing to OT-PVN neurons. Right, SC anatomy (AQ, aqueduct; PAG, periaqueductal gray; dSC, deep SC layers). **b**, We examined visual structures for mCherry expression (areas providing direct projections to OT-PVN cells). Immunohistochemistry showing mCherry+ cells (red) in the medial superficial layers of ipsilateral SC monosynaptically projecting to OT-PVN neurons. Blue, DAPI stain for cell number. Replicates from N=2 animals. **c**, Top, schematic of experimental design for optogenetic stimulation in PVN of fibres from ipsilateral sSC expressing 'ChR2-Venus' or 'Venus', during the 'opaque barrier' condition ('Dam retrieving' and 'Virgin retrieving', 10 pup retrieval trials per session; blue, optogenetic stimulation occurring during dam retrieval). Bottom left, individual performance on pup retrieval testing after each observation session. Bottom right, summary of behavioural data (Kaplan-Meier). **d**, Immunohistochemistry showing fibres from sSC (green, arrows) innervating PVN including near OT-PVN neurons (insert). Red, oxytocin neurons; green, Venus; blue, DAPI. 3V, third ventricle. Replicates from N=3 animals. **e**, Schematic of overall hypothesis connecting behavioural interactions between dam and virgin during co-housing to activation of PVN and OT-PVN neurons (red arrow).

## **Extended Data Fig. 10 Oxytocinergic modulation from PVN promotes plasticity of AC pup call responses.**

**a**, Emergence of pup call responses in left auditory cortex preceded retrieval onset; from 71 units recorded in five animals, 40 units in two animals were maintained throughout co-housing (22/40 units significantly increased activity, 11/40 units significantly decreased activity, mean single-unit response across all 40 units to best pup call 48 h before retrieval onset:  $30.4\pm4.6\%$  increase relative to baseline, spiking 24 h before retrieval:  $33.5\pm2.0\%$  increase, spiking on day of first retrieval:  $45.8\pm3.9\%$  increase,  $p=0.009$ , one-way ANOVA,  $n=40$  units). Gray shading on unit waveforms represents SEM. **b**, Immunostained tissue section of PVN showing co-labeling of oxytocin peptide (green) and GCaMP6s expressed together with TdTomato (red). Replicates from  $N=3$  animals.

## **Supplementary information**

### **Supplementary Information**

This file contains a Supplementary Discussion.

### **Reporting Summary**

### **Supplementary Video 1**

Example of spontaneous pup retrieval by mother occurring during co-housing (same as in Fig. 3a).

### **Supplementary Video 2**

Example of virgin spending time in nest with pups after days of co-housing; same damvirgin dyad recorded on days one through four.

### **Supplementary Video 3**

Retrieval onset in virgin co-housed with dam and pups ('D+'). Example pup retrieval testing on day 0, day 1, and day 3.

### **Supplementary Video 4**

Retrieval onset in virgin co-housed with pups ('D-'). Example pup retrieval testing on day 0, day 1, and day 5.

## Supplementary Video 5

Example of dam shepherding virgin back to nest (same event as in Fig. 1g).

## Supplementary Video 6

Video with more examples of dam shepherding from different co-housing pairs.

## Supplementary Video 7

Video of dam shepherding implanted and tethered virgin.

## Supplementary Video 8

Video with more examples of self-generated retrievals from different co-housing pairs.

## Supplementary Video 9

Video of self-generated retrieval by dam with implanted and tethered virgin in nest.

## Supplementary Video 10

Example of dam bringing pup near virgin.

## Supplementary Video 11

Testing social transmission of maternal behaviour with transparent barrier separating virgin from dam and pups; after two days, virgin was retrieving.

## Supplementary Video 12

Testing social transmission of maternal behaviour with opaque barrier separating virgin from dam and pups; virgin did not begin retrieving even after four days.

## **Source data**

### Source Data Fig. 1

### Source Data Fig. 2

[\*\*Source Data Fig. 3\*\*](#)

[\*\*Source Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 3\*\*](#)

[\*\*Source Data Extended Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 6\*\*](#)

[\*\*Source Data Extended Data Fig. 7\*\*](#)

[\*\*Source Data Extended Data Fig. 8\*\*](#)

[\*\*Source Data Extended Data Fig. 9\*\*](#)

[\*\*Source Data Extended Data Fig. 10\*\*](#)

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[\*\*Reprints and Permissions\*\*](#)

## About this article

### Cite this article

Carcea, I., Caraballo, N.L., Marlin, B.J. *et al.* Oxytocin neurons enable social transmission of maternal behaviour. *Nature* **596**, 553–557 (2021).  
<https://doi.org/10.1038/s41586-021-03814-7>

- Received: 13 May 2020
- Accepted: 08 July 2021
- Published: 11 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03814-7>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03814-7>

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

- Article
- [Published: 18 August 2021](#)

# Cotranslational prolyl hydroxylation is essential for flavivirus biogenesis

- [Ranen Aviner](#)<sup>1,2</sup>,
- [Kathy H. Li](#)<sup>3</sup>,
- [Judith Frydman](#) [ORCID: orcid.org/0000-0003-2302-6943](#)<sup>2</sup> &
- [Raul Andino](#) [ORCID: orcid.org/0000-0001-5503-9349](#)<sup>1</sup>

*Nature* volume **596**, pages 558–564 (2021)

- 2820 Accesses
- 35 Altmetric
- [Metrics details](#)

## Subjects

- [Ribosome](#)
- [Virology](#)

## Abstract

Viral pathogens are an ongoing threat to public health worldwide. Analysing their dependence on host biosynthetic pathways could lead to effective antiviral therapies<sup>1</sup>. Here we integrate proteomic analyses of polysomes with functional genomics and pharmacological interventions to define how enteroviruses and flaviviruses remodel host polysomes to synthesize viral proteins and disable host protein production. We find that infection with polio, dengue or Zika virus markedly modifies polysome composition, without major changes to core ribosome stoichiometry. These viruses use different strategies to evict a common set of translation initiation and RNA surveillance factors from polysomes while recruiting host machineries that are specifically required for viral biogenesis. Targeting these specialized viral polysomes could provide a new approach for antiviral interventions.

For example, we find that both Zika and dengue use the collagen proline hydroxylation machinery to mediate cotranslational modification of conserved proline residues in the viral polyprotein. Genetic or pharmacological inhibition of proline hydroxylation impairs nascent viral polyprotein folding and induces its aggregation and degradation. Notably, such interventions prevent viral polysome remodelling and lower virus production. Our findings delineate the modular nature of polysome specialization at the virus–host interface and establish a powerful strategy to identify targets for selective antiviral interventions.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

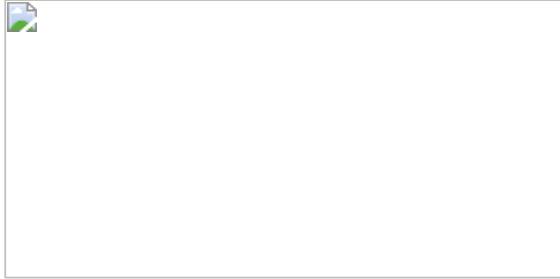
\$19.99/month

Learn more

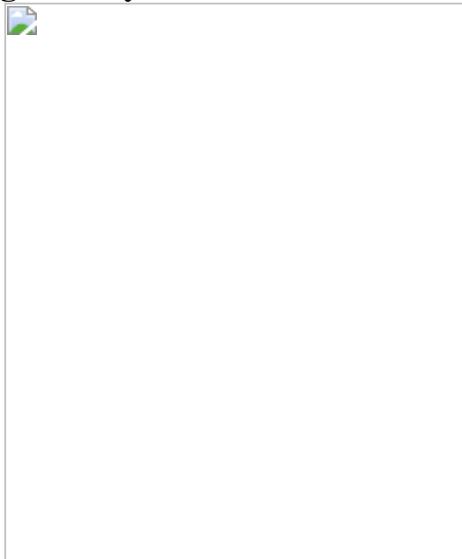
**Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

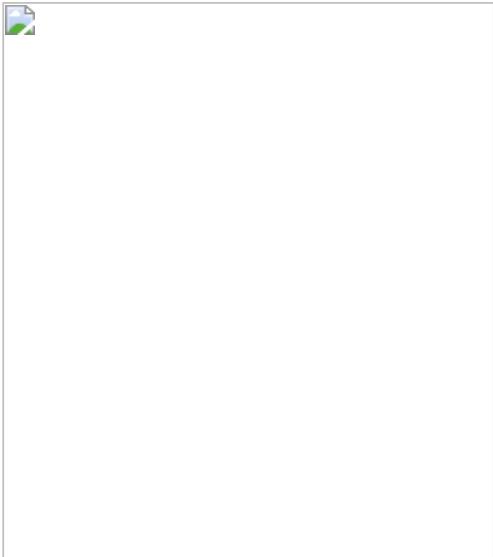
**Fig. 1: Polysome modularity is hijacked by RNA viruses.**



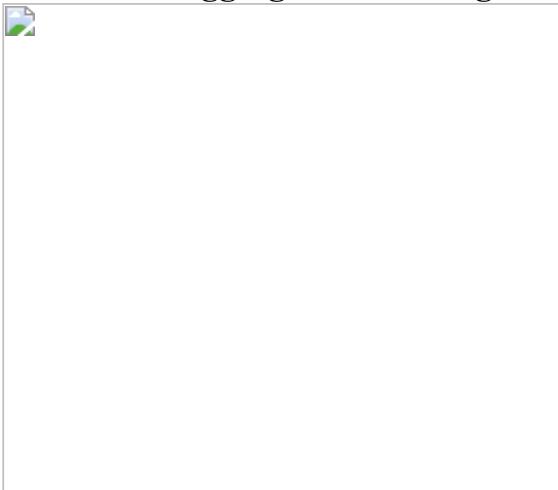
**Fig. 2: Polysome-associated host targets for antiviral intervention.**



**Fig. 3: ZIKV and DENV infection requires collagen prolyl hydroxylases.**



**Fig. 4: Cotranslational proline hydroxylation protects ZIKV and DENV proteins from aggregation and degradation.**



## Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium through the PRIDE<sup>45</sup> partner repository with the dataset identifier PXD024546.

## References

1. 1.

Kumar, N. et al. Host-directed antiviral therapy. *Clin. Microbiol. Rev.* 33, e00168 (2020).

2. 2.

Stern-Ginossar, N., Thompson, S. R., Mathews, M. B. & Mohr, I. Translational control in virus-infected cells. *Cold Spring Harb. Perspect. Biol.* **11**, a033001 (2019).

3. 3.

Nicholson, B. L. & White, K. A. Functional long-range RNA–RNA interactions in positive-strand RNA viruses. *Nat. Rev. Microbiol.* **12**, 493–504 (2014).

4. 4.

Aviner, R. & Frydman, J. Proteostasis in viral infection: unfolding the complex virus–chaperone interplay. *Cold Spring Harb. Perspect. Biol.* **12**, a034090 (2020).

5. 5.

Jaafar, Z. A. & Kieft, J. S. Viral RNA structure-based strategies to manipulate translation. *Nat. Rev. Microbiol.* **17**, 110–123 (2019).

6. 6.

Mizuno, C. M. et al. Numerous cultivated and uncultivated viruses encode ribosomal proteins. *Nat. Commun.* **10**, 752 (2019).

7. 7.

Macejak, D. G. & Sarnow, P. Association of heat shock protein 70 with enterovirus capsid precursor P1 in infected human cells. *J. Virol.* **66**, 1520–1527 (1992).

8. 8.

Yang, W. et al. G3BP1 inhibits RNA virus replication by positively regulating RIG-I-mediated cellular antiviral response. *Cell Death Dis.* **10**, 946 (2019).

9. 9.

Limjindaporn, T. et al. Interaction of dengue virus envelope protein with endoplasmic reticulum-resident chaperones facilitates dengue virus production. *Biochem. Biophys. Res. Commun.* **379**, 196–200 (2009).

10. 10.  
Sweeney, T. R., Abaeva, I. S., Pestova, T. V. & Hellen, C. U. T. The mechanism of translation initiation on type 1 picornavirus IRESs. *EMBO J.* **33**, 76–92 (2014).
11. 11.  
Lee, K.-M., Chen, C.-J. & Shih, S.-R. Regulation mechanisms of viral IRES-driven translation. *Trends Microbiol.* **25**, 546–561 (2017).
12. 12.  
Zhang, Y., Gao, W., Li, J., Wu, W. & Jiu, Y. The role of host cytoskeleton in flavivirus infection. *Virol. Sin.* **34**, 30–41 (2019).
13. 13.  
Dougherty, J., Tsai, W.-C. & Lloyd, R. Multiple poliovirus proteins repress cytoplasmic RNA granules. *Viruses* **7**, 6127–6140 (2015).
14. 14.  
Michalski, D. et al. Zika virus noncoding sfRNAs sequester multiple host-derived RNA-binding proteins and modulate mRNA decay and splicing during infection. *J. Biol. Chem.* **294**, 16282–16296 (2019).
15. 15.  
Moon, S. L. et al. Flavivirus sfRNA suppresses antiviral RNA interference in cultured cells and mosquitoes and directly interacts with the RNAi machinery. *Virology* **485**, 322–329 (2015).
16. 16.  
Ooi, Y. S. et al. An RNA-centric dissection of host complexes controlling flavivirus infection. *Nat. Microbiol.* **4**, 2369–2382 (2019).
17. 17.  
Lenarcic, E. M., Landry, D. M., Greco, T. M., Cristea, I. M. & Thompson, S. R. Thiouracil cross-linking mass spectrometry: a cell-based method to identify host factors involved in viral amplification. *J. Virol.* **87**, 8697–8712 (2013).

18. 18.

Pallarés, H. M. et al. Zika virus subgenomic flavivirus RNA generation requires cooperativity between duplicated RNA structures that are essential for productive infection in human cells. *J. Virol.* **94**, e00343 (2020).

19. 19.

Zeng, J. et al. The Zika virus capsid disrupts corticogenesis by suppressing Dicer activity and miRNA biogenesis. *Cell Stem Cell* **27**, 618–632 (2020).

20. 20.

Rodríguez Pulido, M., Serrano, P., Sáiz, M. & Martínez-Salas, E. Foot-and-mouth disease virus infection induces proteolytic cleavage of PTB, eIF3a,b, and PABP RNA-binding proteins. *Virology* **364**, 466–474 (2007).

21. 21.

Hunt, S. L., Hsuan, J. J., Totty, N. & Jackson, R. J. unr, a cellular cytoplasmic RNA-binding protein with five cold-shock domains, is required for internal initiation of translation of human rhinovirus RNA. *Genes Dev.* **13**, 437–448 (1999).

22. 22.

Andreev, D. E. et al. Glycyl-tRNA synthetase specifically binds to the poliovirus IRES to activate translation initiation. *Nucleic Acids Res.* **40**, 5602–5614 (2012).

23. 23.

Iwasaki, S., Floor, S. N. & Ingolia, N. T. Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature* **534**, 558–561 (2016).

24. 24.

Li-Weber, M. Molecular mechanisms and anti-cancer aspects of the medicinal phytochemicals rocaglamides (=flavaglines). *Int. J. Cancer* **137**, 1791–1799 (2015).

25. 25.

Yang, S. N. Y. et al. RK-33 is a broad-spectrum antiviral agent that targets DEAD-box RNA helicase DDX3X. *Cells* **9**, 170 (2020).

26. 26.

Kramer, G., Shiber, A. & Bukau, B. Mechanisms of cotranslational maturation of newly synthesized proteins. *Annu. Rev. Biochem.* **88**, 337–364 (2019).

27. 27.

Gamerdinger, M., Hanebuth, M. A., Frickey, T. & Deuerling, E. The principle of antagonism ensures protein targeting specificity at the endoplasmic reticulum. *Science* **348**, 201–207 (2015).

28. 28.

Taguwa, S. et al. Zika virus dependence on host Hsp70 provides a protective strategy against infection and disease. *Cell Rep.* **26**, 906–920 (2019).

29. 29.

Taguwa, S. et al. Defining Hsp70 subnetworks in dengue virus replication reveals key vulnerability in flavivirus infection. *Cell* **163**, 1108–1123 (2015).

30. 30.

Kirk, T. Z., Evans, J. S. & Veis, A. Biosynthesis of type I procollagen. Characterization of the distribution of chain sizes and extent of hydroxylation of polysome-associated pro-alpha-chains. *J. Biol. Chem.* **262**, 5540–5545 (1987).

31. 31.

DiChiara, A. S. et al. Mapping and exploring the collagen-I proteostasis network. *ACS Chem. Biol.* **11**, 1408–1421 (2016).

32. 32.

Ishikawa, Y. & Bächinger, H. P. A molecular ensemble in the rER for procollagen maturation. *Biochim. Biophys. Acta* **1833**, 2479–2491 (2013).

33. 33.

Ueno, T. et al. Enhancement of procollagen biosynthesis by p180 through augmented ribosome association on the endoplasmic reticulum in response to

stimulated secretion. *J. Biol. Chem.* **285**, 29941–29950 (2010).

34. 34.

Li, Y., Li, Q., Wong, Y. L., Liew, L. S. Y. & Kang, C. Membrane topology of NS2B of dengue virus revealed by NMR spectroscopy. *Biochim. Biophys. Acta* **1848**, 2244–2252 (2015).

35. 35.

Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L. & Raines, R. T. Conformational stability of collagen relies on a stereoelectronic effect. *J. Am. Chem. Soc.* **123**, 777–778 (2001).

36. 36.

Naduthambi, D. & Zondlo, N. J. Stereoelectronic tuning of the structure and stability of the trp cage miniprotein. *J. Am. Chem. Soc.* **128**, 12430–12431 (2006).

37. 37.

Guo, J. et al. pVHL suppresses kinase activity of Akt in a proline-hydroxylation-dependent manner. *Science* **353**, 929–932 (2016).

38. 38.

Burrill, C. P., Strings, V. R. & Andino, R. Poliovirus: generation, quantification, propagation, purification, and storage. *Curr. Protoc. Microbiol.* **128**, 15H.1.1–15H.1.27 (2013).

39. 39.

Kinney, R. M. et al. Construction of infectious cDNA clones for dengue 2 virus: strain 16681 and its attenuated vaccine derivative, strain PDK-53. *Virology* **230**, 300–308 (1997).

40. 40.

Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

41. 41.

Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).

42. 42.

Shah, P. S. et al. Comparative flavivirus–host protein interaction mapping reveals mechanisms of dengue and Zika virus pathogenesis. *Cell* **175**, 1931–1945 (2018).

43. 43.

Gamarnik, A. V. & Andino, R. Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev.* **12**, 2293–2304 (1998).

44. 44.

Di Veroli, G. Y. et al. An automated fitting procedure and software for dose–response curves with multiphasic features. *Sci. Rep.* **5**, 14701 (2015).

45. 45.

Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

46. 46.

Cox, J. & Mann, M. 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinf.* **13**, S12 (2012).

47. 47.

Imami, K. et al. Phosphorylation of the ribosomal protein RPL12/uL11 affects translation during mitosis. *Mol. Cell* **72**, 84–98 (2018).

48. 48.

Simsek, D. et al. The mammalian ribo-interactome reveals ribosome functional diversity and heterogeneity. *Cell* **169**, 1051–1065 (2017).

49. 49.

Fontaine, K. A. et al. The cellular NMD pathway restricts Zika virus infection and is targeted by the viral capsid protein. *MBio* **9**, e02126-18 (2018).

50. 50.

Göertz, G. P. et al. Subgenomic flavivirus RNA binds the mosquito DEAD/H-box helicase ME31B and determines Zika virus transmission by *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **116**, 19136–19144 (2019).

51. 51.

Kanodia, P. et al. A rapid and simple quantitative method for specific detection of smaller coterminal RNA by PCR (DeSCo-PCR): application to the detection of viral subgenomic RNAs. *RNA* **26**, 888–901 (2020).

52. 52.

Lloyd, R. E. Nuclear proteins hijacked by mammalian cytoplasmic plus strand RNA viruses. *Virology* **479–480**, 457–474 (2015).

53. 53.

Jagdeo, J. M. et al. Heterogeneous nuclear ribonucleoprotein M facilitates enterovirus infection. *J. Virol.* **89**, 7064–7078 (2015).

54. 54.

Gustin, K. E. & Sarnow, P. Effects of poliovirus infection on nucleocytoplasmic trafficking and nuclear pore complex composition. *EMBO J.* **20**, 240–249 (2001).

55. 55.

Silvera, D., Gamarnik, A. V. & Andino, R. The N-terminal K homology domain of the poly(rC)-binding protein is a major determinant for binding to the poliovirus 5'-untranslated region and acts as an inhibitor of viral translation. *J. Biol. Chem.* **274**, 38163–38170 (1999).

56. 56.

Dave, P. et al. Strand-specific affinity of host factor hnRNP C1/C2 guides positive to negative-strand ratio in Coxsackievirus B3 infection. *RNA Biol.* **16**, 1286–1299 (2019).

## Acknowledgements

This research was supported by fellowships from Rothschild-Yad Hanadiv Foundation, the European Molecular Biology Organization (EMBO ALTF 289-2015) and the Human Frontier Science Program (HFSP LT000050-2016) to R. Aviner; National Institutes of Health (NIH) grants AI127447 and GM056433 to J.F.; NIH grants R01 AI36178, AI40085, P01 AI091575; the University of California (California Center for Antiviral Drug Discovery; CCADD); and DoD-DARPA Prophecy to R. Andino.

## Author information

### Affiliations

1. Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, CA, USA

Ranen Aviner & Raul Andino

2. Department of Biology and Genetics, Stanford University, Stanford, CA, USA

Ranen Aviner & Judith Frydman

3. Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA, USA

Kathy H. Li

### Contributions

R. Aviner, R. Andino and J.F. designed the study. R. Aviner carried out experiments, analysed data and performed statistical analyses. K.H.L. performed LC–MS/MS data acquisition. R. Aviner, R. Andino and J.F. wrote the manuscript, and all authors approved the manuscript.

### Corresponding authors

Correspondence to [Judith Frydman](#) or [Raul Andino](#).

### Ethics declarations

## Competing interests

All authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Andreas Pichlmair, Alexander Khromykh and the anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

### [Extended Data Fig. 1 Suppression of global translation is associated with preferential synthesis of viral proteins.](#)

**a, b**, Schematics of vRNA and polyprotein organization for poliovirus (**a**), ZIKV, and DENV (**b**). IRES, internal ribosome entry site; ORF, open reading frame. **c**, Poliovirus, ZIKV and DENV infection is associated with shut-off of global translation and continued synthesis of viral proteins. Top, immunoblots of cells infected with poliovirus, ZIKV or DENV (MOI = 5) and treated with puromycin for 15 min at the indicated time points, to label nascent chains. Results are representative of 4 independent repeats. Bottom, time course of infectious particle production detected by plaque assays (grey) and densitometry of viral proteins (black) or puromycin-labelled nascent chains (pink). Data are means  $\pm$  s.d. **d**, Related to Fig. [1b](#). Polysome profile analysis (top) and immunoblots (bottom) of infected and control cells subjected to sucrose gradient fractionation followed by either qPCR analysis of negative strand viral RNA or puromycin treatment of gradient fractions. Results are representative of 3 independent repeats. **e**, Cells infected as in d were lysed in buffer containing EDTA and no Mg<sup>2+</sup> and subjected to polysome profile analysis and qPCR of positive strand viral RNA. Results are representative of 2 independent repeats.

### [Extended Data Fig. 2 Overview of polysome mass spectrometry in infected and control cells.](#)

**a**, Pairwise comparisons of two biological replicates of polysome MS from mock-infected cells. Each dot represents the level of a single protein as detected by MS.

iBAQ, intensity based absolute quantification.  $r$ , Spearman correlation coefficient. **b**, iBAQ-normalized MS intensities of proteins associated with the indicated cellular complexes and organelles, plotted as percent of total, for either lysate (input) or polysome MS analyses. Data are means  $\pm$  s.d.  $n = 2$  and 3 for Lysate(input) and Polysomes, respectively. **c**, Venn bar diagram of proteins detected by polysome MS across all conditions and time points. **d**, Distribution of ribosomal protein intensities (grey) and median intensities of viral polyproteins (pink), as detected by polysome MS. **e**, To identify co-sedimenting non-polysomal proteins, we performed polysome MS on uninfected cells pretreated with 1  $\mu$ M puromycin for 3 min. Puromycin treated and untreated controls were subjected to polysome profile analysis on sucrose gradients, in triplicates. Dashed lines indicate fractions pooled for MS analysis. Because samples are normalized for total protein content, non-polysomal proteins, whose migration in the gradient is unaffected by puromycin, will appear to be enriched, whereas polysomal proteins, which will shift to lighter fractions in a puromycin dependent manner, will appear depleted or unchanged. To determine fold-changes (FC), we subtracted the means of each  $\log_2$ -transformed iBAQ value in puromycin versus control. FC distribution is shown in green for ribosomal proteins and grey for non-ribosomal proteins (histogram). FC values were then used for 1D annotation enrichment analysis<sup>46</sup>, where the FC distribution in each category is compared to the global distribution using a two-sided Wilcoxon-Mann-Whitney test. FDR values for each category are plotted as a function of mean FC (scatterplot). **f**, Most proteins detected by polysome MS are either previously identified ribosome-interacting partners<sup>47,48</sup> or known to be involved in protein biogenesis.

### **Extended Data Fig. 3 Polysome composition is gradually and extensively remodelled during infection.**

**a**, Infection is not associated with changes to the core ribosome composition. Volcano plots of median-adjusted iBAQ intensities of ribosomal proteins measured by polysome MS in infected and uninfected cells. Two-sided Student's  $t$ -test, FDR < 0.01, S0 = 1. **b**, Differential detection of peptides mapping to ribosomal proteins may reflect changes in post-translational modifications. Same analysis as in **a**, only at the peptide level. Red, ribosomal peptides with statistically significant increases in polysome MS of DENV at 60 hpi versus mock. Bottom bar graph, number of changing peptides mapping to each indicated protein, plotted as percent of all detected peptides that map to that protein. All changing peptides were found to harbor at least one—and most more than two—potential PTM sites based on Phosphosite data. **c**, Pairwise comparisons of non-ribosomal proteins detected by polysome MS (left) and total proteome analysis of matching lysate samples (right) in infected cells ( $y$  axis) versus uninfected control ( $x$  axis).  $r$ , Spearman correlation coefficient. **d**, Euclidean distance hierarchical clustering of Spearman correlation

coefficients across all polysome MS samples. Highest correlations are in white, lowest in red. **e**, Protein rank plots for total proteome of infected cells. Viral polyproteins are indicated in colour. **f**, Related to Fig. [1d](#), line plots show medians±interquartile range of fold change from mock in  $\log_2$ -transformed iBAQ intensities, based on clustering shown in Fig. [1d](#).

**Extended Data Fig. 4 Differentially associated polysome interactors in mock versus the terminal time point for each virus.**

**a–c**, Volcano plots comparing mock polysome MS with that of poliovirus (**a**), ZIKV (**b**) and DENV (**c**) at 4, 36 and 60 h, respectively. Two-sided Student's *t*-test, FDR < 0.01, S0 = 1.

**Extended Data Fig. 5 Antiviral host factors are evicted from polysomes by distinct yet convergent mechanisms.**

**a**, Major RNA QC factors are strongly evicted from polysomes by all three viruses. STRING protein-protein interaction network for host factors involved in RNA surveillance and quality control depleted from (blue) or recruited to (pink) polysomes during infection. Differentially associated proteins (ANOVA FDR < 0.01 across all polysome MS samples) were grouped based on annotations in Fig. [1g](#). Each node is divided in three to reflect individual changes for each virus and coloured based on changes in polysome association from Mock. Node size represents relative abundance, which is the maximum iBAQ intensity measured for each indicated protein across all polysome MS samples. **b**, Line plots of the change over time in polysome association of G3BP complexes (top) and UPF1/MOV10 (bottom) during infection. Each line represents a single replicate of polysome MS time course. Dotted grey line, translation shut-off. **c**, G3BP1 is cleaved during poliovirus but not ZIKV or DENV infection. Immunoblot analysis of whole-cell lysates from cells infected with poliovirus, ZIKV or DENV (MOI = 5). Time courses are 1–6 h for poliovirus and 12–72 h for ZIKV and DENV. Arrowhead marks the position of G3BP1 cleavage fragment. Results are representative of 2 independent repeats. **d**, Volcano plot comparing polysome interactors in cells infected with poliovirus for 4h versus mock. Host proteins known to be cleaved by poliovirus protease are indicated. **e**, Polysome MS captures most direct interactors of vRNA. Venn diagram of proteins detected by polysome MS and those detected by pull-down of poliovirus, ZIKV and DENV vRNA from the indicated studies<sup>[17,18](#)</sup>. **f**, High affinity flavivirus RNA interactors are depleted from polysomes during infection. Violin plots of proteins detected by both vRNA pulldowns and polysome MS. Proteins were divided into two groups based on their polysome association in this study: depleted or recruited by more than 2 folds after infection with ZIKV or DENV at 36 or 60 h, compared to

mock. y-axis, protein abundance in vRNA pulldown-MS. *P*, Mann Whitney U test p-value. For each violin plot, white circles indicate the median of each distribution, black rectangles indicate interquartile range and vertical lines indicate the 95% confidence intervals. **g**, Polysome interactors depleted during ZIKV infection consist of high affinity ZIKV 3'UTR or capsid interactors. Pairwise comparison of the fold change in polysomes interactors (*x* axis) versus their protein abundance in ZIKV vRNA pulldown-MS (*y* axis). Known ZIKV sfRNA and capsid interactors from [14,15,49,50](#) are coloured blue and purple, respectively. **h**, Schematic of the production of subgenomic flaviviral RNA (sfRNA) by limited 5'-to-3' XRN1 exonucleolytic cleavage of full-length genomes. **i, j**, Most ZIKV sfRNA does not associate with polysomes. Cells were infected with ZIKV (MOI = 5) for 24 h and lysates were fractionated on sucrose gradients followed by: qPCR analysis using primers that amplify either full-length genome (pink) or 3'UTR sequence present in both full-length genome and sfRNA (blue). Values for each fraction are plotted as percent of total across the entire gradient. Data are means  $\pm$  s.d (*n* = 3) (i); or PCR analysis for specific detection of either gRNA or sfRNA using blocking oligonucleotides (j) [51](#). **k**, Most ZIKV capsid does not associate with polysomes. Levels of ZIKV capsid detected by MS analysis of either polysomes or lysate input samples from cells infected with ZIKV for 36 h. Values are sum of peptides mapping to the capsid region of the viral polyprotein divided by the sum of peptides mapping to ribosomal proteins, with average ratios in the lysate set as 1. Data are means  $\pm$  s.d (*n* = 3). *P*, two-tailed Student's *t*-test p-value. **l**, A similar set of RNA QC factors is evicted from polysomes by either viral protease cleavage (poliovirus) or sfRNA/capsid sequestration (ZIKV) to prevent translational silencing of gRNA.

## Extended Data Fig. 6 Rearrangement of translation initiation complexes during viral infection.

**a**, Line plots of the change over time in polysome association of initiation factors 3 (EIF3), 4G1 (EIF4G1), 4A1 (EIF4A1) and poly-A-binding protein (PABPC1). Included in the quantification of EIF3 are the 8 subunits of the core octamer complex. Data are means  $\pm$  s.d. (*n* = 3) Asterisk denotes statistically significant differences in EIF4A1 polysome association between 24 and 36 hpi for ZIKV and 48 to 60 hpi for DENV ( $p=0.0037$  and 0.0071, respectively, two-tailed Student's *t*-test). Dotted grey line, translation shut-off. **b**, EIF4G1 c-terminal fragment, generated by poliovirus protease cleavage to inhibit cap-dependent translation while promoting polio-virus-specific translation initiation<sup>[21](#)</sup>, is retained on polysomes during poliovirus infection. Mapping of polysome-associated EIF4G1 peptides, as detected by polysome MS of cells infected with poliovirus, ZIKV or DENV for 4, 36 and 60 h, respectively, plotted as cumulative fraction from N- to C-termini. Protein–protein interaction domains are shown below. **c**, EIF3, a multisubunit complex, forms two

distinct subcomplexes in virus-infected cells. Line plots of the change over time in polysome association of EIF3 subunits, divided into the octamer (top) and yeast-like core (bottom) subcomplexes, coloured blue and cyan on the structure of a preinitiating 48S ribosome (right; PDB 6FEC). Each line represents a single replicate of polysome MS time course. Dotted grey line, translation shut-off. **d**, Toxicity of RocA at 36 h. Data are means  $\pm$  s.d. ( $n = 3$ ). **e**, Heat map of the median fold change from mock in polysome association of RNA helicases during infection with each of the three viruses.

### Extended Data Fig. 7 RNA-binding proteins control the fate of polysome-associated poliovirus vRNA.

Many RNA-binding proteins have been implicated in regulating poliovirus translation and replication by binding and restructuring its 5' UTR<sup>52</sup>. **a**, Multiple RNA-binding proteins are uniquely recruited to polysomes by poliovirus. STRING protein–protein interaction network for host factors involved in alternative splicing and mRNA processing depleted from (blue) or recruited to (pink) polysomes during infection. Differentially associated proteins (ANOVA FDR $<0.01$  across all polysome MS samples) were grouped based on annotations in Fig. 1g. Each node is divided into three to reflect individual changes for each virus and coloured based on change in polysome association from Mock. Node size represents relative abundance, which is the maximum iBAQ intensity measured for each indicated protein across all polysome MS samples. **b**, To identify novel RNA interactors regulating poliovirus biogenesis, we used the temporal pattern of change in CSDE1/Unr, a bona fide poliovirus ITAF<sup>21</sup>, across all polysome MS samples. We then generated a list of the top 15 polysome-associated RNA-binding proteins based on similarity in temporal patterns using Euclidian distances. Shown are medians $\pm$ interquartile range of the top 15 interactors identified. Specific protein names are noted on the right. Dotted grey line, translation shut-off. Seven of the top 15 hits are known interactors, including bona fide ITAFs e.g. PCBP2, STRAP/UnrIP<sup>21</sup> and glycyl-tRNA synthetase (GARS)<sup>22</sup>. Recruitment of additional factors that mirror the same temporal pattern, for example, HNRNPR, SRSF10 and TRA2B, is likely to reflect previously unknown involvement in poliovirus biogenesis. **c**, Line plots of the change over time in polysome association of known poliovirus 5' UTR binders, plotted as fold change from mock. Each line is interquartile range. Dotted grey line, translation shut-off. Green, factors that stimulate poliovirus translation; pink, factors that inhibit poliovirus translation and stimulate replication; grey, alternative splicing factor cleaved by poliovirus protease to facilitate infection<sup>53</sup>. Values for GARS were recalculated manually based on unique peptides only, as many peptides also map to IARS. Cytosolic ITAFs are recruited earlier than nuclear factors (3 versus 4 hpi), consistent with a later disruption of nuclear pores by poliovirus protease<sup>54</sup>. This

coincides with cleavage of PCBP2, which converts it from activator to inhibitor of translation<sup>55</sup>, as well as increased association of nuclear HNRNPC, an essential component of viral replication that competes with and displaces ITAFs from the 5' UTR of picornaviruses<sup>54,56</sup>. **d**, Poliovirus translation peaks by 3 hpi while replication lags by about an hour. (+)vRNA and 3CD polymerase were quantified by qPCR and immunoblot analysis, respectively, in cells infected with poliovirus (MOI = 5) for the indicated times. Shown are representative blots from 2 independent repeats, and qPCR values of (+)vRNA normalized to GAPDH (pink) and 3CD densitometry normalized to actin (black) ( $n = 1$ ). **e**, Schematic of the change in poliovirus 5' UTR-binding proteins upon transition from translation to replication. Early during translation, viral genome binds cytosolic ITAFs that stimulate its translation. Later during infection, poliovirus 3C protease cleaves nuclear pore proteins, and alternative splicing factors are released into the cytoplasm to inhibit new rounds of translation initiation on polysome-associated vRNA. This releases vRNA from polysomes for use as a template in replication. Thus, the observed rearrangement of 5' UTR-binding proteins between 3 and 4 hpi acts as a timer that shifts viral genomes from translation to replication.

### **Extended Data Fig. 8 Cytosolic and ER proteostasis of viral nascent chains.**

**a**, Line plots of the change over time in polysome association of cytosolic HSP70 chaperones and co-chaperones. Each line represents a single replicate of polysome MS time course. Dotted grey line, translation shut-off. **b**, Efficiency of KD in Huh7 cells based on qPCR using primers specific to the indicated transcripts, normalized to GAPDH ( $n = 2$ ). **c**, Toxicity of JG40 and JG345 at 36 h. Data are means  $\pm$  s.d. ( $n = 3$ ). **d**, JG345 inhibits poliovirus, ZIKV and DENV, whereas JG40 only inhibits ZIKV and DENV. Virus production by plaque assays of cells infected with poliovirus, ZIKV or DENV (MOI = 0.1) for 5, 24 and 36 h, respectively, in the presence or absence of HSP70 inhibitor, plotted as % of DMSO control. Data are means  $\pm$  s.d. ( $n = 3$ ). **e**, Virus production by plaque assays of cells infected with poliovirus (MOI = 0.1) in the presence or absence of HSP70 inhibitors. Data are means  $\pm$  s.d. ( $n = 3$ ). **f**, Eviction of nascent chain-associated complex (NAC) from polysomes during flavivirus infection is associated with increased binding of ER-resident factors. Line plots of the change over time in polysome association of NAC subunits (NACA and BTF3, blue) and ER-resident factors with GOCC ER lumen annotation (pink). Data are means  $\pm$  s.d. for NAC or median  $\pm$  interquartile range for GOCC ER Lumen. Dotted grey line, translation shut-off. **g**, Cumulative fraction of polysome interactors with GOCC ER Part annotation.  $P$ , two-sided Mann Whitney U test p-value. **h**, Efficiency of KD in Huh7 cells based on qPCR using primers

specific to the indicated transcripts, normalized to GAPDH ( $n = 2$ ). **i**, Toxicity of indicated compounds at 36 h. Data are means  $\pm$  s.d. ( $n = 3$ ).

### Extended Data Fig. 9 Collagen prolyl hydroxylases facilitate infection of ZIKV and DENV.

**a**, Collagen-specific chaperone, but not collagens or an RNA-binding protein required for their translation, is recruited to polysomes by ZIKV and DENV. Line plots of the change over time in polysome association of collagen-specific chaperone (SERPINH1/HSP47), all collagens identified by polysome MS, and collagen translation regulator RRPB1. Each line represents a single replicate of polysome MS time course. Dotted grey line, translation shut-off. **b**, Related to Fig. 3c. Polysome profiles (top) and immunoblot analysis of gradient fractions treated with puromycin to label nascent chains (bottom). Results are representative of 2 independent repeats. **c**, P4HA1 co-sedimentation with ZIKV polysomes is disrupted by ribosome runoff. Wild-type cells were infected with ZIKV (MOI = 5) for 24 h then treated with 20  $\mu$ M harringtonine for 10 min followed by addition of 1  $\mu$ M puromycin for another 10 min. Lysates were subjected to sucrose gradient fractionation and puromycin labelling. Shown are rRNA absorbance profiles (top) and immunoblots (bottom). Results are representative of 2 independent repeats. **d**, Efficiency of KD in Huh7 cells based on qPCR using primers specific to the indicated transcripts, normalized to GAPDH. ( $n = 2$ ). **e**, **f**, Time course of virus production by plaque assays in wild-type and cP4H-KD cells for ZIKV (**e**) and DENV (**f**) at MOI = 0.1 ( $n = 3$ ). **g**, Wild-type or cP4H-KD cells were infected with ZIKV at the indicated MOIs. 200  $\mu$ M ascorbic acid (VitC) was added at the start of infection and again 12 h later, and virus production at 24 h was monitored using plaque assays. Data are means  $\pm$  s.d. ( $n = 3$ ).

### Extended Data Fig. 10 Proline hydroxylation affects viral protein stability.

**a**, Multiple DENV proline residues are hydroxylated cotranslationally by cP4H. Bubble chart shows individual hydroxyproline sites, as detected by polysome MS in wild-type cells infected with DENV for 60 h. Bubble size is proportional to posterior error probability (PEP) of hydroxyproline sites detected by MS. **b**, Loss of proline hydroxylation prevents full induction of translation shut-off by ZIKV infection. Wild-type and cP4H-KD cells were infected with ZIKV (MOI = 10) for 24 h and metabolically labelled for 1h followed by autoradiography (left) and quantification (right). Data are means  $\pm$  s.d. ( $n = 4$ ).  $P$ , two-tailed Student's *t*-test p-value. **c**, Proline hydroxylation stabilizes multiple viral proteins during infection. Wild-type or cP4H-KD cells were infected with ZIKV (MOI = 10) and MG132, CHX or MG132+CHX were added at 24 h for additional 12 h prior to immunoblot analysis. Results are

representative of 3 independent repeats. **d**, Lack of proline hydroxylation does not affect NS3 protease activity or NS1 protein stability/oligomerization. Strep-tagged ZIKV proteins were transfected into wild-type or cP4H-KD cells and analyzed by nonreducing immunoblot at 24 h. Representative of 3 independent repeats. **e**, Proline hydroxylation reduces aggregation of ZIKV and DENV NS2B. Strep-tagged NS2B3 was transfected into wild-type and cP4H-KD cells and MG132 was added at 24 h for additional 12 h, prior to immunoblot analysis. Results are representative of 2 independent repeats. **f**, Wild-type or cP4H-KD cells were infected with DENV (MOI = 10) and MG132 was added at 36 for additional 12 h prior to immunoblot analysis. Results are representative of 3 independent repeats.

## Supplementary information

### Supplementary Figure

This file contains the uncropped blots.

### Reporting Summary

### Supplementary Table 1

Mass spectrometry data, polysome mass spectrometry and global proteome profiling.

### Supplementary Table 2

Polysome mass spectrometry after puromycin pretreatment (related to Extended Data Fig. 2e).

### Supplementary Table 3

Ribosomal proteins in polysome mass spectrometry (related to Extended Data Fig. 3a, b).

### Supplementary Table 4

Differentially associated polysome interactors (related to Fig. 1d).

### Supplementary Table 5

Hydroxyproline sites in mass spectrometry data.

## **Supplementary Table 6**

Oligonucleotides and shRNA.

## **Peer Review File**

## **Rights and permissions**

### **Reprints and Permissions**

## **About this article**

### **Cite this article**

Aviner, R., Li, K.H., Frydman, J. *et al.* Cotranslational prolyl hydroxylation is essential for flavivirus biogenesis. *Nature* **596**, 558–564 (2021).  
<https://doi.org/10.1038/s41586-021-03851-2>

- Received: 16 September 2020
- Accepted: 23 July 2021
- Published: 18 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03851-2>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03851-2>

- Article
- [Published: 07 July 2021](#)

# Antibody epitopes in vaccine-induced immune thrombotic thrombocytopenia

- [Angela Huynh](#)<sup>1</sup>,
- [John G. Kelton](#) ORCID: [orcid.org/0000-0001-9418-2383](https://orcid.org/0000-0001-9418-2383)<sup>1,2</sup>,
- [Donald M. Arnold](#) ORCID: [orcid.org/0000-0003-0943-8853](https://orcid.org/0000-0003-0943-8853)<sup>1,2</sup>,
- [Mercy Daka](#) ORCID: [orcid.org/0000-0002-7149-2781](https://orcid.org/0000-0002-7149-2781)<sup>3</sup> &
- [Ishac Nazy](#) ORCID: [orcid.org/0000-0002-5695-4343](https://orcid.org/0000-0002-5695-4343)<sup>1,2</sup>

[Nature](#) volume 596, pages 565–569 (2021)

- 57k Accesses
- 1020 Altmetric
- [Metrics details](#)

## Subjects

- [Immunological disorders](#)
- [Translational research](#)

## Abstract

Vaccine-induced immune thrombotic thrombocytopenia (VITT) is a rare adverse effect of COVID-19 adenoviral vector vaccines<sup>1,2,3</sup>. VITT resembles heparin-induced thrombocytopenia (HIT) in that it is associated with platelet-activating antibodies against platelet factor 4 (PF4)<sup>4</sup>; however, patients with VITT develop thrombocytopenia and thrombosis without exposure to heparin. Here we sought to determine the binding site on PF4 of antibodies from patients with VITT. Using alanine-scanning mutagenesis<sup>5</sup>, we found that the binding of anti-PF4 antibodies from patients with VITT ( $n = 5$ ) was restricted to eight surface amino acids on PF4, all of which were located within the heparin-binding site, and that the binding was inhibited by heparin. By contrast, antibodies from patients with HIT ( $n = 10$ ) bound to amino

acids that corresponded to two different sites on PF4. Biolayer interferometry experiments also revealed that VITT anti-PF4 antibodies had a stronger binding response to PF4 and PF4-heparin complexes than did HIT anti-PF4 antibodies, albeit with similar dissociation rates. Our data indicate that VITT antibodies can mimic the effect of heparin by binding to a similar site on PF4; this allows PF4 tetramers to cluster and form immune complexes, which in turn causes Fc $\gamma$  receptor IIa (Fc $\gamma$ RIIa; also known as CD32a)-dependent platelet activation. These results provide an explanation for VITT-antibody-induced platelet activation that could contribute to thrombosis.

[Download PDF](#)

## Main

VITT is a rare but serious adverse effect of adenoviral vector vaccines against SARS-CoV-2. The clinical picture of VITT is moderate to severe thrombocytopenia together with arterial and/or venous thrombi, often occurring in unusual locations<sup>1,2,3</sup>. These findings resemble the immunological drug reaction HIT, which presents clinically as thrombocytopenia and thrombosis in patients who have previously been exposed to heparin<sup>4</sup>. VITT most closely resembles the exceptionally rare spontaneous HIT, which occurs in the absence of heparin<sup>6,7</sup>.

HIT is caused by immunoglobulin G (IgG) antibodies that bind to neoepitopes on PF4 (also known as CXCL4), a 70-amino-acid cationic protein that is contained within platelets<sup>8,9</sup>. The neoepitopes become exposed after heparin, a large anionic polysaccharide, binds to a specific site on PF4, which causes tetramers to cluster together. The IgG-specific antibodies bind to PF4-heparin to form immune complexes, which activate platelets through Fc $\gamma$ RIIa receptors, leading to an intense activation of platelets and the release of procoagulant-rich microparticles<sup>10</sup>. Other cells, including monocytes, are also activated by these immune complexes, which amplifies the hypercoagulable state in patients with HIT<sup>11</sup>. It has been postulated that VITT has a similar pathophysiology to HIT, and several studies have shown that high levels of anti-PF4 antibodies are present in samples from patients with VITT<sup>1,2,3,12</sup>. However, VITT is a unique syndrome as it occurs without heparin exposure, and the pattern of platelet reactivity in vitro does not exhibit the typical heparin dependence that is seen with HIT.

In this report, we describe the binding site and characteristics of anti-PF4 antibodies that developed in patients with VITT in response to vaccination against COVID-19 with an adenoviral vector. We found that patients with VITT had anti-PF4 antibodies that bound to a highly restricted site on PF4 that corresponds to the heparin-binding

site. These antibodies can form platelet-activating immune complexes without heparin, potentially causing the thrombocytopenia and clotting that are observed in VITT.

## Patient demographics

Samples from patients with VITT ( $n = 5$ ) were referred to the McMaster Platelet Immunology Laboratory for diagnosis. All patients with VITT had received a single dose of the ChAdOx1 nCoV-19 vaccine (AstraZeneca COVID-19 vaccine, AstraZeneca; COVISHIELD, Verity Pharmaceuticals and Serum Institute of India) and subsequently developed thrombocytopenia and thrombosis; the mean age of patients was 44 years (range, 35–72 years) and 2 out of 5 (40%) were female. The time from first dose of the ChAdOx1 nCoV-19 vaccine to sample collection was 14–40 days (mean, 28 days). All samples from patients with VITT (hereafter, VITT samples) had antibodies against PF4 (mean optical density (OD), 2.71; range, 0.763–3.347).

The samples from patients with VITT were compared to samples from patients with HIT ( $n = 10$ ) who had thrombocytopenia after receiving heparin and had a high clinical probability score (4Ts score of at least 4), with detectable anti-PF4-heparin antibodies and evidence of platelet activation in vitro. The mean age of patients with HIT was 69 years (range, 52–81 years) and 5 out of 10 (50%) were female. Nine of 10 (90%) patients with HIT experienced thrombosis. The time from heparin initiation to sample collection was 6–27 days (mean, 14.3 days). All samples from patients with HIT (hereafter, HIT samples) had detectable anti-PF4-heparin antibodies (mean OD, 3.10; range, 2.329–3.897).

## Platelet activation profiles of VITT antibodies

Using a functional platelet activation assay (serotonin release assay; SRA), we showed that VITT and HIT samples had unique patterns of platelet reactivity in vitro. VITT samples did not exhibit any dependence on heparin for platelet activation in vitro (Extended Data Fig. [1a](#)). All VITT samples showed strong, dose-dependent PF4-mediated platelet activation with  $50 \mu\text{g ml}^{-1}$  PF4 (Extended Data Fig. [1b](#)). Complete inhibition of platelet activation by VITT and HIT samples was achieved with the addition of IV.3, an Fc $\gamma$ RIIa-blocking monoclonal antibody<sup>4</sup>. These results indicate that VITT antibodies require PF4 for platelet activation, and that this is mediated by the engagement of Fc $\gamma$ RIIa receptors, consistent with the effects of an immune complex. By contrast, HIT samples activated platelets with two distinct profiles. Patients with HIT with heparin-dependent antibodies were patients who were exposed to heparin and whose antibodies activated platelets in the SRA in the presence of heparin (Extended Data Fig. [1c](#)). Patients with HIT with heparin-independent antibodies were patients who were exposed to heparin and whose antibodies activated

platelets in the SRA both in the presence and in the absence of heparin (Extended Data Fig. 1c). All HIT samples also activated platelets with the addition of PF4 (Extended Data Fig. 1d).

## Binding site of VITT antibodies on PF4

Previous studies have shown that the binding epitopes of HIT antibodies are non-contiguous and conformation-dependent<sup>13</sup>. Therefore, to identify the specific amino acid targets of VITT antibodies on PF4, we used alanine-scanning mutagenesis and produced 70 unique recombinant PF4 mutants, each differing by a single amino acid<sup>5</sup>. We defined a critical binding amino acid as one that when mutated caused a reduction in binding of more than 50% in the corresponding PF4 mutant compared to wild-type PF4. We identified eight surface amino acids that were necessary for binding of VITT samples (Arg22, His23, Glu28, Lys46, Asn47, Lys50, Lys62 and Lys66; Table 1, Fig. 1a). The PF4 mutants R22A and E28A affected the binding for all five VITT samples. This restricted epitope is consistent with limited B cell clonality and suggests that the binding site of VITT antibodies is to a specific site on PF4. We observed that four of the eight amino acids that comprised the VITT epitope (Arg22, His23, Lys46 and Lys66) corresponded to amino acids on PF4 to which heparin binds<sup>14</sup> (Fig. 1b). We postulate that binding of VITT anti-PF4 antibodies to the heparin-binding site on the PF4 tetramer explains why the addition of heparin in vitro does not augment platelet activation, as seen in HIT; rather, it inhibits platelet activation, presumably by displacing VITT anti-PF4 antibodies. Furthermore, in a PF4 enzyme immunoassay (EIA), we showed that the binding of VITT antibodies to PF4 was inhibited by therapeutic concentrations of heparin in four of the available VITT samples (Extended Data Fig. 2). Our results explain why some VITT samples tested in previous studies<sup>1,2</sup> were inhibited by therapeutic doses of heparin. These findings suggest that VITT antibodies cause platelet activation through a heparin-like mechanism by stabilizing complexes of PF4, aligning the Fc regions of the VITT antibodies to be in close proximity to one another and cross-link FcγRIIa receptors on platelets—similar to the mechanism of the monoclonal antibody 1E12, which can activate platelets independent of heparin<sup>15</sup>.

**Table 1 Key amino acids in PF4 for the binding of VITT and HIT antibodies**  
**Fig. 1: Identification of amino acids in PF4 that are critical for the binding of antibodies using alanine-scanning mutagenesis.**

---

 **figure1**

**a, b**, For VITT samples ( $n = 5$ ), the binding site (coloured in red; **a**) aligns within the heparin-binding site on PF4 (coloured in brown; **b**). **c**, For HIT samples ( $n = 10$ ), one main binding site was identified for all samples (coloured in blue). **d**, An additional binding site was found in 6 out of 10 (60%) of the HIT samples (coloured in red),

which aligns within the heparin-binding site on PF4. **e**, The amino acids that are predicted to be part of the binding sites for VITT and heparin-dependent (HD) and heparin-independent (HI) antibodies (abs) in HIT are highlighted on the primary sequence of PF4. Images are modified from the Protein Data Bank (PDB) entry 1RHP.

The binding site of VITT antibodies on PF4 was then compared with that of HIT antibodies. When all of the critical amino acids that were identified through screening of the 10 patients with HIT were combined, there was a total of 10 amino acids (Leu8, Cys10, Cys12, Thr16, Arg22, Gln40, Asn47, Cys52, Leu53, Asp54, Lys61, Lys66 and Leu67; Table 1) that were part of the HIT epitopes, in different combinations. No single common amino acid was critical for the binding of all 10 HIT samples, probably owing to the polyclonal nature of the antibodies<sup>5</sup>. The loss of binding in PF4 mutants L8A, C10A, C12A, T16A, C52A, D54A and K61A corresponded to the most common amino acids that affect the binding of HIT antibodies and were common among six of the HIT samples. When displayed on the PF4 tetramer, we identified one site on PF4 that all HIT samples (10 out of 10) targeted (Fig. 1c, d). In addition, 6 of the 10 HIT samples targeted an additional site (Fig. 1d) similar to the VITT site that was within the heparin-binding site. Unlike the VITT samples, none of the HIT samples were restricted to the heparin-binding site. This is consistent with previous observations that some HIT samples contain two types of platelet-activating anti-PF4-heparin antibodies<sup>16</sup>. The HIT antibodies bound to a similar site as KKO, a monoclonal antibody against PF4-heparin complexes, thus explaining why HIT antibodies, but not VITT antibodies, require heparin to cross-link PF4 tetramers<sup>5</sup>.

In addition to clarifying the binding site of VITT antibodies, these results provide an explanation for why some rapid HIT immunoassays may yield false-negative results for VITT<sup>17</sup>. One rapid HIT immunoassay, the latex immunoturbidimetric assay (HemosIL HIT-Ab<sub>(PF4-H)</sub>), uses KKO to aggregate complexes of PF4-heparin. As all HIT samples have antibodies that bind to the same site as KKO, the HIT samples compete with KKO in binding the PF4-heparin complexes. By contrast, VITT antibodies bind to a different site on PF4 than KKO, and thus do not compete for binding.

## Binding kinetics of VITT antibodies

The binding responses of VITT samples ( $n = 5$ ), HIT samples ( $n = 10$ ) and samples from healthy control individuals ( $n = 10$ ) to PF4 and PF4-heparin were tested using bio-layer interferometry (BLI; Fig. 2). Binding responses are a measure of the antigen-specific antibodies present in a given sample. When healthy control samples ( $n = 10$ ) were tested with immobilized PF4 and PF4-heparin, the mean binding response (measured as the average wavelength shift (in nm)  $\pm$  2 standard deviations (2 s.d.))

was  $0.0059 \pm 0.11$  nm and  $0.031 \pm 0.11$  nm, respectively. All VITT ( $n = 5$ ) and HIT ( $n = 10$ ) samples were above the mean + 2 s.d. cut-off of 0.12 nm with immobilized PF4 and 0.14 nm with immobilized PF4–heparin, indicative of a positive binding result. The mean binding response (nm shift  $\pm$  s.d.) was  $1.82 \pm 0.88$  nm for VITT samples and  $0.82 \pm 0.72$  nm for HIT samples ( $P < 0.05$ ) with immobilized PF4. Similarly, the mean binding response was  $1.24 \pm 0.70$  nm for VITT samples and  $0.62 \pm 0.45$  nm for HIT samples ( $P < 0.05$ ) with PF4–heparin complexes. Therefore, the binding response in VITT samples was significantly higher than that in HIT samples and healthy control samples for both PF4 and PF4–heparin, indicating a stronger antibody response in patients with VITT.

We further compared the binding response for the VITT samples to the two groups of HIT samples, with heparin-dependent ( $n = 4$ ) and heparin-independent ( $n = 6$ ) antibodies. The mean binding response for patients with HIT with heparin-dependent antibodies was  $0.29 \pm 0.18$  nm, and for patients with HIT with heparin-independent antibodies the mean binding response was  $1.18 \pm 0.73$  nm against immobilized PF4 (Fig. [2c](#)). When samples were tested against immobilized PF4–heparin complexes, the mean binding response was  $0.30 \pm 0.09$  nm for patients with HIT with heparin-dependent antibodies and  $0.83 \pm 0.48$  nm for patients with HIT with heparin-independent antibodies (Fig. [2d](#)). The binding response of VITT samples was significantly higher than that of samples from patients with HIT with heparin-dependent antibodies and healthy control samples for both PF4 and PF4–heparin (VITT versus HIT with heparin-dependent antibodies:  $P < 0.01$ ; VITT versus healthy controls:  $P < 0.001$ ). By contrast, the binding response of VITT samples was similar to that of samples from patients with HIT with heparin-independent antibodies. VITT samples also had a significantly higher binding response to PF4 when compared to PF4–heparin ( $P < 0.05$ ), consistent with an inhibitory effect of heparin on the binding of VITT antibodies to PF4 (Extended Data Fig. [2](#)).

**Fig. 2: Association and dissociation of antibodies against PF4 and PF4–heparin using BLI.**

---

 **figure2**

---

**a, b**, Representative binding responses of antibodies from a patient with VITT, a patient with HIT and a healthy control participant to biotinylated PF4 (**a**) and biotinylated PF4–heparin (**b**) immobilized on streptavidin biosensors. **c–f**, Binding responses with PF4 (**c**) and PF4–heparin (**d**) and dissociation rates ( $k_{\text{off}}$ ) with PF4 (**e**) and PF4–heparin (**f**) of samples from patients with VITT ( $n = 5$ ), patients with HIT with heparin-dependent antibodies ( $n = 4$ ), patients with HIT with heparin-independent antibodies ( $n = 6$ ) and healthy control individuals ( $n = 10$ ). Patients with HIT with heparin-dependent antibodies ( $n = 4$ ) are patients with typical HIT and patients with HIT with heparin-independent antibodies ( $n = 6$ ) are patients with HIT who exhibited heparin-independent platelet activation in the SRA. The results show that VITT samples have higher binding responses to both PF4 and PF4–heparin than do HIT samples with heparin-dependent antibodies (PF4,  $P = 0.0017$ ; PF4–heparin,  $P = 0.0090$ ) and healthy control samples. Values are shown as the mean binding response (nm)  $\pm$  s.d. and mean dissociation rate  $k_{\text{off}}$  ( $s^{-1}$ )  $\pm$  s.d. HD antibodies, heparin-dependent antibodies; HI antibodies, heparin-independent antibodies. Statistical analysis by two-tailed Mann–Whitney test; \*\* $P < 0.01$ , \*\*\* $P < 0.0005$ .

To confirm that the binding responses in samples were due to anti-PF4 antibodies and not other serum factors, total IgG antibodies were purified from two VITT samples and then tested for their ability to bind PF4 and PF4–heparin. However, owing to sample constraints, further purification of specific VITT anti-PF4 antibodies was not performed, which represents a potential limitation for this study. The binding responses of total IgG antibodies isolated from both VITT samples against PF4 and PF4–heparin were similar to the binding responses observed with their respective sera, indicating that the binding responses are due to the anti-PF4 antibodies (Extended

Data Fig. 3). In addition, two VITT, one HIT and two healthy control samples were re-tested in separate experiments and showed reproducibility (Extended Data Fig. 4).

When polyclonal antibodies interact with an antigen, the dissociation rate, which is concentration-independent, can be measured. The mean dissociation rate ( $k_{\text{off}}$   $\text{s}^{-1} \pm \text{s.d.}$ ) was  $6.32 \times 10^{-3} \pm 0.0077 \text{ s}^{-1}$  for VITT samples,  $2.34 \times 10^{-3} \pm 0.0041 \text{ s}^{-1}$  for HIT samples with heparin-dependent antibodies and  $9.65 \times 10^{-4} \pm 0.0016 \text{ s}^{-1}$  for HIT samples with heparin-independent antibodies, for immobilized PF4 (Fig. 2e). Similarly, the mean dissociation rate ( $k_{\text{off}} \text{ s}^{-1} \pm \text{s.d.}$ ) was  $3.47 \times 10^{-3} \pm 0.0029 \text{ s}^{-1}$  for VITT samples,  $4.56 \times 10^{-4} \pm 0.0002 \text{ s}^{-1}$  for HIT samples with heparin-dependent antibodies and  $1.07 \times 10^{-3} \pm 0.0018 \text{ s}^{-1}$  for HIT samples with heparin-independent antibodies, for immobilized PF4-heparin (Fig. 2f). There was no statistically significant difference in the dissociation rates of the VITT group and the two HIT groups with both PF4 and PF4-heparin (PF4: VITT versus HIT with heparin-dependent antibodies,  $P = 0.482$ ; VITT versus HIT with heparin-independent antibodies,  $P = 0.220$ ; PF4-heparin: VITT versus HIT with heparin-dependent antibodies,  $P = 0.104$ ; VITT versus HIT with heparin-dependent antibodies,  $P = 0.160$ ). The low dissociation rates probably allow for sufficient binding and formation of immune complexes to induce platelet activation by cross-linking Fc $\gamma$ RIIa receptors on platelets. As human antibodies in serum are polyclonal, affinity could not be determined; however, binding responses and dissociation rates demonstrate the strength of the immune response and the avidity of the polyclonal samples, respectively.

Monoclonal antibodies against PF4, such as KKO (ref. 18) and 1E12 (ref. 15) facilitate the formation of ultra-large complexes of PF4 on the platelet surface. Previous studies using antibodies from patients with HIT<sup>19</sup> and monoclonal antibodies that resemble HIT antibodies<sup>15</sup> have shown that the higher affinity of heparin-independent antibodies in some patients with HIT can cluster PF4 tetramers and create platelet-activating immune complexes in the absence of heparin. Similarly, we found that anti-PF4 antibodies from patients with VITT have a comparable binding response and avidity to the antibodies from patients with HIT—especially patients with HIT with heparin-independent antibodies—implying that they can cluster PF4 tetramers and create the same ultra-large platelet-activating immune complexes (Extended Data Fig. 5).

It has been observed that 50% of HIT samples tested in the SRA cause platelet activation in vitro without the addition of heparin<sup>20,21,22</sup>. As such, the HIT samples were separated into two groups: those with heparin-dependent antibodies and those with both heparin-dependent and heparin-independent antibodies. The distinguishing feature that defines anti-PF4 antibodies from patients with VITT is that these patients

—unlike patients with HIT—have not been exposed to heparin, and their anti-PF4 antibodies are restricted to the heparin-binding site.

In this report, we show that anti-PF4 antibodies in patients with VITT can induce platelet activation through Fc $\gamma$ RIIa receptors in the presence of PF4, without heparin. However, other serum factors could also contribute to platelet activation. Previous studies found that antibodies from patients with VITT were able to activate platelets and cause platelet aggregation in the presence of adenoviral particles in a dose-dependent manner<sup>1,23,24</sup>. Thus, it is possible that platelet activation caused by anti-PF4 antibodies in patients with VITT is not the only factor that leads to the development of thrombotic events. HIT is also propagated by various pro-thrombotic mechanisms that could also be important in VITT, including Fc-receptor polymorphisms<sup>25</sup>, monocyte activation and tissue factor production<sup>26</sup>, and the generation of procoagulant microparticles<sup>10</sup>.

This study offers an explanation for VITT-mediated platelet activation. The patients with VITT in our study exhibited similar antibody characteristics to one another and their antibodies bound PF4 at the same site as heparin. VITT antibodies form immune complexes without the addition of heparin or other co-factors, and activate platelets and potentially other cells through Fc $\gamma$ RIIa receptors, which, in turn, could initiate coagulation at multiple points to cause thrombocytopaenia and thrombosis.

## 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.

### Study participants

Participants included patients diagnosed with VITT ( $n = 5$ ), patients diagnosed with HIT ( $n = 10$ ) and healthy volunteers ( $n = 10$ ). VITT diagnosis was based on four criteria: recent AstraZeneca vaccination; positive for anti-PF4 IgG antibodies; positive in the PF4-enhanced SRA; and no previous exposure to heparin. HIT diagnosis was confirmed based on three criteria: the 4Ts score, in which all patients with HIT had a clinical score of at least 4; a positive commercially available PF4-enhanced heparin-dependent IgG/A/M-specific EIA (Immucor, OD  $\geq 0.45$ ) and a positive SRA (SRA  $\geq 20\%$   $^{14}\text{C}$ -serotonin release)<sup>27</sup>. This study was approved by the Hamilton

Integrated Research Ethics Board (HiREB) and informed written consent was obtained from all participants.

## Platelet activation assays

Platelet activation assays were performed in the presence of heparin using the SRA, and including a modification in which increasing doses of exogenous PF4 were added, rather than heparin (PF4-SRA)<sup>27,28</sup>. Some assays were performed with high concentrations of unfractionated heparin (100 IU ml<sup>-1</sup>), or with an Fc-receptor-blocking monoclonal antibody (IV.3).

## Epitope mapping of antibody binding to PF4 from VITT and HIT samples using alanine-scanning mutagenesis

The full-length DNA coding sequence of human PF4 was cloned into the pET22b expression vector using restriction sites NdeI and HindIII (GenScript). The PF4 mutants were expressed and purified as previously described<sup>5,29</sup>. In brief, PF4 mutants were designed whereby non-alanine amino acids in wild-type PF4 were mutated to alanine and the alanine amino acids in wild-type PF4 were mutated to valine. PF4 mutants were introduced into *Escherichia coli* ArcticExpress (DE3) cells (Agilent Technologies). For the overexpression of PF4 mutants, cultures were grown at 37 °C to mid-exponential phase before induction with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown at 37 °C for 3 h. *E. coli* cells for each wild-type PF4 or PF4 mutant were lysed by sonication in 20 mM sodium phosphate, pH 7.2, 400 mM sodium chloride, 1.4 mM β-mercaptoethanol, 5% (v/v) glycerol, 1% (v/v) Triton X-100 (Thermo Fisher Scientific), and 0.5% (w/v) sodium deoxycholate (MilliporeSigma) with 2 mM MgCl<sub>2</sub>, 10 µg ml<sup>-1</sup> DNaseI (MilliporeSigma) and EDTA-free protease inhibitor cocktail (Roche). The supernatant was then cleared by centrifugation at 40,000g for 40 min at 4 °C and applied onto a HiTrap Q HP column (Cytiva Life Sciences) equilibrated with 20 mM sodium phosphate, pH 7.2, 400 mM sodium chloride, 1.4 mM β-mercaptoethanol and 5% (v/v) glycerol. The flow-through of the Q HP column was then stored at 4 °C, overnight. The following day, the serum was diluted twofold to yield a sodium chloride concentration of 200 mM with 20 mM sodium phosphate, pH 7.2, 1.4 mM β-mercaptoethanol and 5% (v/v) glycerol, syringe-filtered with a 0.2-µm filter (Acrodisc) and loaded onto a HiTrap Heparin HP column (Cytiva Life Sciences). Contaminants were eluted with 0.5 M sodium chloride and PF4 was eluted with a linear gradient from 0.5 to 2 M sodium chloride. Fractions containing pure wild-type or PF4 mutants were pooled, concentrated and buffer-exchanged to phosphate-buffered saline (PBS) and 1.5 M sodium chloride. The concentration of PF4 was determined using a bicinchoninic acid (BCA) assay (Thermo

Fisher Scientific). Protein expression and purity was assessed for each PF4 mutant using 4–18% denaturing SDS-PAGE.

The effect of the 70 amino acids on the binding of anti-PF4-heparin antibodies in patient samples was measured and analysed in a similar manner to that described previously<sup>5</sup>. The binding of anti-PF4-heparin antibodies to wild-type PF4 and PF4 mutants was measured using a modified PF4-heparin IgG-specific EIA<sup>5,13</sup>. Microtitre well plates (384 wells; Thermo Scientific Nunc) were coated with 10 µg ml<sup>-1</sup> streptavidin and 1 IU ml<sup>-1</sup> biotinylated-heparin and blocked with PBS supplemented with 3% (v/v) bovine serum albumin (BSA) for 2 h at ambient temperature. Wild-type PF4 or PF4 mutants at 5 µg ml<sup>-1</sup> were then added and incubated for 1 h at ambient temperature. Diluted patient samples (1:50 prepared in 1% BSA in PBS) in technical duplicates were added to the plates and incubated for 1 h at room temperature. After washing, alkaline-phosphatase-conjugated goat anti-human IgG ( $\gamma$ -chain-specific, Jackson ImmunoResearch Laboratories) was added at a 1:3,000 dilution and incubated for 1 h at ambient temperature. The addition of 1 mg ml<sup>-1</sup> para-nitrophenylphosphate (MilliporeSigma) substrate dissolved in 1 M diethanolamine buffer (pH 9.6) was used for detection. The OD was measured at 405 nm and 490 nm (as a reference) using a BioTek 800TS microplate reader (BioTek) to assess the binding of antibodies to wild-type PF4 and PF4 mutants. Results were reported as a percentage of the loss of binding relative to wild-type PF4 binding.

## Heparin inhibition of VITT anti-PF4 antibodies

Microtitre well plates (96 wells, Nunc Maxisorp) were coated overnight at 4 °C with 100 µl per well of PF4 (60 µg ml<sup>-1</sup>) diluted in 50 mM carbonate-bicarbonate buffer (pH 9.6). The plates were then blocked with 200 µl per well of 3% (v/v) BSA prepared in PBS at room temperature for 2 h. VITT samples ( $n = 4$ ) that were available were diluted 1:50 with 1% BSA in PBS and pre-incubated with increasing concentrations of unfractionated heparin (final concentrations of 0.1, 0.5, 1, 5 and 100 IU ml<sup>-1</sup>; Pfizer) for 1 h at room temperature. The blocking solution was removed from the microtitre well plates and the VITT samples and heparin mixtures (100 µl per well) in technical duplicates were added to the plates, which were then incubated for 1 h at room temperature. The plates were washed twice with PBS–0.05% Tween 20 and three times with PBS. Bound human IgG antibodies were detected with 100 µl per well of alkaline-phosphatase-conjugated goat anti-human IgG ( $\gamma$ -chain-specific, 1:3,000, Jackson Immuno-Research Laboratories) antibody prepared in 1% (v/v) BSA in PBS. Plates were washed as before and followed with the addition of 100 µl substrate (para-nitrophenylphosphate dissolved in diethanolamine buffer (MilliporeSigma)). The OD was measured at 405 nm and 490 nm (as a reference) using a BioTek 800TS microplate reader (BioTek).

## Purification of total IgG antibodies

Total IgG antibodies from two VITT samples were purified for further analysis of binding kinetics. A volume of 2 ml protein G-coated sepharose beads (Thermo Fisher Scientific) was washed three times with PBS at room temperature, 300g for 5 min. Patient samples were heat-inactivated at 56 °C for 30 min and diluted three times in PBS. The samples were then transferred to protein G sepharose beads and incubated at room temperature for 1 h before extensive rinsing with 30 ml PBS. Total IgG was eluted from the protein G sepharose beads with 0.1 M glycine, pH 2.7 and neutralized by Tris buffer, pH 8.0.

## Binding kinetics of VITT and HIT antibodies using BLI

Wild-type PF4 was labelled with biotin as previously described<sup>30</sup>. In brief, wild-type PF4 and PF4 mutants were incubated with 5× the volume of heparin sepharose 6 fast flow affinity chromatography medium (Cytiva Life Sciences) for 1 h with shaking at ambient temperature. EZ-Link Sulfo-NHS-LC-Biotin (Thermo Fisher Scientific) was added to the PF4 and heparin sepharose mixture in 20 molar excess and allowed to react for 1 h with shaking at ambient temperature. The biotinylated wild-type PF4 or PF4 mutants were eluted from the heparin sepharose using PBS and 2 M sodium chloride. Absorbance at 280 nm was measured using a spectrophotometer (Eppendorf AG) and used to calculate the concentration. Biotinylation of PF4 was checked using a streptavidin-coated anti-PF4-heparin EIA.

BLI experiments were performed using the Octet Red 96 instrument (FortéBio). Samples or buffer were dispensed into 96-well black flat-bottom microtitre plates (Greiner Bio-One) diluted in PBS supplemented with 1% (v/v) BSA at a volume of 200 µl per well with an operating temperature maintained at 23 °C. Streptavidin-coated biosensor tips (FortéBio) were hydrated with 1% BSA in PBS (MilliporeSigma) to establish a baseline before antigen immobilization for 60 s. Biotinylated recombinant PF4 (final concentration 7.5 µg ml<sup>-1</sup>), alone or complexed with 0.125 IU ml<sup>-1</sup> unfractionated heparin (LEO Pharma), was then immobilized on the biosensor tips for 1,200 s at 1,000 rpm followed by a baseline re-establishment for 1,800 s at 1,000 rpm. Antigen-coated sensors were then reacted with heat-inactivated patient samples or purified total IgG at a 1:32 dilution in duplicate for 780 s at 1,000 rpm followed by a dissociation step for 3,000 s at 1,000 rpm. Data were analysed using Octet User Software v.3.1 using the 2:1 heterogenous ligand-binding model. Reference values from control wells with buffer alone were subtracted and all results were aligned to the measured baseline. The binding profile response of each sample was expressed as the average wavelength shift in nm.

## Data acquisition, statistical analysis and reproducibility

Differences between data were tested for statistical significance using the paired or unpaired *t*-test and the Mann–Whitney test. *P*-values are reported as two-tailed and *P* < 0.05 was considered to be statistically significant. All statistical analyses were conducted using GraphPad Prism v.9.1.0 (GraphPad Software). Experiments were repeated with technical duplicates independently twice with similar results unless otherwise stated.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

The datasets generated during and/or analysed during the current study are not publicly available to allow for commercialization of research findings but are available from the corresponding author on reasonable request.

## References

1. 1. Greinacher, A. et al. Thrombotic thrombocytopenia after ChAdOx1 nCov-19 vaccination. *N. Engl. J. Med.* **384**, 2092–2101 (2021).
2. 2. Schultz, N. H. et al. Thrombosis and thrombocytopenia after ChAdOx1 nCoV-19 vaccination. *N. Engl. J. Med.* **384**, 2124–2130 (2021).
3. 3. Scully, M. et al. Pathologic antibodies to platelet factor 4 after ChAdOx1 nCoV-19 vaccination. *N. Engl. J. Med.* **384**, 2202–2211 (2021).
4. 4. Kelton, J. G. et al. Heparin-induced thrombocytopenia: laboratory studies. *Blood* **72**, 925–930 (1988).
5. 5.

Huynh, A. et al. Characterization of platelet factor 4 amino acids that bind pathogenic antibodies in heparin-induced thrombocytopenia. *J. Thromb. Haemost.* **17**, 389–399 (2019).

6. 6.

Warkentin, T. E., Makris, M., Jay, R. M. & Kelton, J. G. A spontaneous prothrombotic disorder resembling heparin-induced thrombocytopenia. *Am. J. Med.* **121**, 632–636 (2008).

7. 7.

Hwang, S. R. et al. Cerebral venous sinus thrombosis associated with spontaneous heparin-induced thrombocytopenia syndrome after total knee arthroplasty. *Platelets* <https://doi.org/10.1080/09537104.2020.1828574> (2020).

8. 8.

Amiral, J. et al. Antibodies to macromolecular platelet factor 4-heparin complexes in heparin-induced thrombocytopenia: a study of 44 cases. *Thromb. Haemost.* **73**, 21–28 (1995).

9. 9.

Cai, Z. et al. Atomic description of the immune complex involved in heparin-induced thrombocytopenia. *Nat. Commun.* **6**, 8277 (2015).

10. 10.

Warkentin, T. E. et al. Sera from patients with heparin-induced thrombocytopenia generate platelet-derived microparticles with procoagulant activity: an explanation for the thrombotic complications of heparin-induced thrombocytopenia. *Blood* **84**, 3691–3699 (1994).

11. 11.

Rauova, L. et al. Monocyte-bound PF4 in the pathogenesis of heparin-induced thrombocytopenia. *Blood* **116**, 5021–5031 (2010).

12. 12.

Thiele, T. et al. Frequency of positive anti-PF4/polyanion antibody tests after COVID-19 vaccination with ChAdOx1 nCoV-19 and BNT162b2. *Blood* <https://doi.org/10.1182/blood.2021012217> (2021).

13. 13.

Horsewood, P., Warkentin, T. E., Hayward, C. P. & Kelton, J. G. The epitope specificity of heparin-induced thrombocytopenia. *Br. J. Haematol.* **95**, 161–167 (1996).

14. 14.

Mayo, K. H. et al. Heparin binding to platelet factor-4. An NMR and site-directed mutagenesis study: arginine residues are crucial for binding. *Biochem. J.* **312**, 357–365 (1995).

15. 15.

Vayne, C. et al. Characterization of new monoclonal PF4-specific antibodies as useful tools for studies on typical and autoimmune heparin-induced thrombocytopenia. *Thromb. Haemost.* **121**, 322–331 (2021).

16. 16.

Nguyen, T. H., Medvedev, N., Delcea, M. & Greinacher, A. Anti-platelet factor 4/polyanion antibodies mediate a new mechanism of autoimmunity. *Nat. Commun.* **8**, 14945 (2017).

17. 17.

Vayne, C. et al. PF4 immunoassays in vaccine-induced thrombotic thrombocytopenia. *N. Engl. J. Med.* (2021).

18. 18.

Sachais, B. S. et al. Dynamic antibody-binding properties in the pathogenesis of HIT. *Blood* **120**, 1137–1142 (2012).

19. 19.

Bui, V. C. & Nguyen, T. H. The role of single-molecule force spectroscopy in unraveling typical and autoimmune heparin-induced thrombocytopenia. *Int. J. Mol. Sci.* **19**, 1054 (2018).

20. 20.

Socher, I., Kroll, H., Jorks, S., Santoso, S. & Sachs, U. J. Heparin-independent activation of platelets by heparin-induced thrombocytopenia antibodies: a

common occurrence. *J. Thromb. Haemost.* **6**, 197–200 (2008).

21. 21.

Warkentin, T. E., Arnold, D. M., Nazi, I. & Kelton, J. G. The platelet serotonin-release assay. *Am. J. Hematol.* **90**, 564–572 (2015).

22. 22.

Padmanabhan, A. et al. Heparin-independent, PF4-dependent binding of HIT antibodies to platelets: implications for HIT pathogenesis. *Blood* **125**, 155–161 (2015).

23. 23.

Tiede, A. et al. Prothrombotic immune thrombocytopenia after COVID-19 vaccine. *Blood* <https://doi.org/10.1182/blood.2021011958> (2021).

24. 24.

Althaus, K. et al. Antibody-mediated procoagulant platelets in SARS-CoV-2-vaccination associated immune thrombotic thrombocytopenia. *Haematologica* <https://doi.org/10.3324/haematol.2021.279000> (2021).

25. 25.

Arepally, G., McKenzie, S. E., Jiang, X. M., Poncz, M. & Cines, D. B. Fc $\gamma$ RIIA H/R<sup>131</sup> polymorphism, subclass-specific IgG anti-heparin/platelet factor 4 antibodies and clinical course in patients with heparin-induced thrombocytopenia and thrombosis. *Blood* **89**, 370–375 (1997).

26. 26.

Kasthuri, R. S. et al. PF4/heparin-antibody complex induces monocyte tissue factor expression and release of tissue factor positive microparticles by activation of Fc $\gamma$ RI. *Blood* **119**, 5285–5293 (2012).

27. 27.

Sheridan, D., Carter, C. & Kelton, J. G. A diagnostic test for heparin-induced thrombocytopenia. *Blood* **67**, 27–30 (1986).

28. 28.

Nazi, I. et al. Distinguishing between anti-platelet factor 4/heparin antibodies that can and cannot cause heparin-induced thrombocytopenia. *J. Thromb. Haemost.* **13**, 1900–1907 (2015).

29. 29.

Huynh, A. et al. Development of a high-yield expression and purification system for platelet factor 4. *Platelets* **29**, 249–256 (2018).

30. 30.

Huynh, A. et al. The role of fluid-phase immune complexes in the pathogenesis of heparin-induced thrombocytopenia. *Thromb. Res.* **194**, 135–141 (2020).

31. 31.

Rubino, J. G. et al. A comparative study of platelet factor 4-enhanced platelet activation assays for the diagnosis of heparin-induced thrombocytopenia. *J. Thromb. Haemost.* **19**, 1096–1102 (2021).

## Acknowledgements

We thank R. Clare, N. Ivetic, A. Hucik, J. C. Moore and J. W. Smith for their technical assistance and A. Bissola and C. Groves for the input on the figures. Funding support for this work was provided by a grant from the Canadian Institutes of Health Research (CIHR 452655) awarded to I.N and a grant from the Public Health Agency of Canada awarded to D.M.A.

## Author information

### Affiliations

1. Department of Medicine, Michael G. DeGroote School of Medicine, McMaster University, Hamilton, Ontario, Canada

Angela Huynh, John G. Kelton, Donald M. Arnold & Ishac Nazy

2. McMaster Centre for Transfusion Research, McMaster University, Hamilton, Ontario, Canada

John G. Kelton, Donald M. Arnold & Ishac Nazy

3. Department of Biochemistry and Biomedical Sciences, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

Mercy Daka

## Contributions

A.H. performed epitope mapping and BLI, analysed data and wrote the manuscript. J.G.K. and D.M.A. designed the research, interpreted data and wrote the manuscript. M.D. carried out the purification of PF4 mutants, performed epitope mapping and wrote the manuscript. I.N. designed the research, analysed and interpreted data and wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

## Corresponding author

Correspondence to [Ishac Nazy](#).

## Ethics declarations

## Competing interests

A US provisional patent was submitted by I.N., D.M.A., J.G.K. and A.H. covering the products and methods for the diagnosis of VITT and the differentiation of VITT from HIT using the research in this study. The authors declare no other competing interests.

## Additional information

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

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

## Extended data figures and tables

[Extended Data Fig. 1 Activation of platelets in samples from patients with VITT and patients with HIT in different SRAs.](#)

Platelet activation is shown as the percentage release of  $^{14}\text{C}$ -serotonin. **a, c**, Samples from patients with VITT ( $n = 5$ ; red lines) (**a**) and patients with HIT ( $n = 10$ ; blue lines) (**c**) in the standard SRA<sup>27</sup>. **b, d**, The same VITT (**b**) and HIT (**d**) samples in the PF4-enhanced SRA<sup>28,31</sup> with added PF4. The results showed that in the standard SRA, the HIT samples activated platelets with ( $n = 4$ ) or without ( $n = 6$ ) heparin or with added PF4, whereas the VITT samples only activated platelets with the addition of PF4. Experiments were repeated with technical duplicates independently twice with similar results. All platelet activation was inhibited by the addition of the IV.3 monoclonal antibody that binds the Fc $\gamma$ RIIa receptors on platelets.

#### **Extended Data Fig. 2 Inhibition of VITT antibody binding to PF4 with increasing heparin concentrations.**

VITT samples ( $n = 4$ ) were tested with increasing concentrations of heparin to investigate its effect on the binding of VITT antibodies to PF4. The experiment was repeated with technical duplicates independently twice with similar results. VITT antibody binding was disrupted by therapeutic concentrations and higher than therapeutic concentrations of heparin in all four VITT samples that were available for testing. Results are shown as the mean of the technical replicates. OD  $\geq 0.45$  is considered positive for anti-PF4 antibodies.

#### **Extended Data Fig. 3 BLI curves comparing binding kinetics in serum and purified total IgG of two VITT samples.**

Binding kinetics of VITT serum ( $n = 2$ ) and its corresponding purified total IgG ( $n = 2$ ) were compared to demonstrate anti-PF4 antibody specific responses. The experiment was repeated with technical duplicates independently twice with similar results. **a, b**, Association and dissociation steps are shown for biotinylated PF4 (**a**) and PF4-heparin (**b**) immobilized on streptavidin biosensor tips. Black lines represent the serum of VITT sample 1, red lines represent the purified total IgG of VITT sample 1, grey lines represent the serum of VITT sample 2 and purple lines represent the purified total IgG of VITT sample 2.

#### **Extended Data Fig. 4 Binding responses of replicate samples from patients with VITT, a patient with HIT and healthy control individuals.**

VITT ( $n = 2$ ), HIT ( $n = 1$ ) and healthy control ( $n = 2$ ) samples were tested on different days to determine the reproducibility of the biolayer interferometry experiments. Experiments were repeated with technical duplicates independently three times with similar results. **a, b**, Binding responses against immobilized PF4 (**a**) and PF4-heparin (**b**) were measured as the average wavelength shift in nm. The binding response for

each antigen and sample was reproducible on separate assays. Red points represent VITT samples, blue points represent the HIT sample and black points represent healthy control samples.

## **Extended Data Fig. 5 Proposed mechanism for VITT antibodies binding to and clustering PF4 tetramers, independent of heparin, and forming platelet-activating immune complexes.**

**a, b**, We postulate that VITT antibodies bind the antigen, PF4 (**a**), which in turn can cluster PF4 tetramers and create platelet-activating immune complexes in the absence of heparin (**b**). **c, d**, The immune complexes can be found on the platelet surface and in solution, resulting in the aligning and close proximity of the Fc part of these antibodies, which then are able to cross-link Fc $\gamma$ RIIa receptors (**c**) and lead to platelet activation (**d**).

## **Supplementary information**

### **Reporting Summary**

### **Rights and permissions**

### **Reprints and Permissions**

### **About this article**

#### **Cite this article**

Huynh, A., Kelton, J.G., Arnold, D.M. *et al.* Antibody epitopes in vaccine-induced immune thrombotic thrombocytopenia. *Nature* **596**, 565–569 (2021).  
<https://doi.org/10.1038/s41586-021-03744-4>

- Received: 04 May 2021
- Accepted: 21 June 2021
- Published: 07 July 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03744-4>

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03744-4>

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

- Article
- [Published: 21 July 2021](#)

# Tonic prime-boost of STING signalling mediates Niemann–Pick disease type C

- [Ting-Ting Chu<sup>1</sup>](#),
- [Xintao Tu](#) ORCID: [orcid.org/0000-0001-8789-5297<sup>1</sup>](#),
- [Kun Yang](#) ORCID: [orcid.org/0000-0002-2011-7084<sup>1</sup>](#),
- [Jianjun Wu<sup>1</sup>](#),
- [Joyce J. Repa<sup>2,3</sup>](#) &
- [Nan Yan](#) ORCID: [orcid.org/0000-0002-0637-3989<sup>1,4</sup>](#)

[Nature](#) volume **596**, pages 570–575 (2021)

- 10k Accesses
- 136 Altmetric
- [Metrics details](#)

## Subjects

- [Innate immunity](#)
- [Neuroimmunology](#)

This article has been [updated](#)

## Abstract

The classic mode of STING activation is through binding the cyclic dinucleotide 2'3'-cyclic GMP–AMP (cGAMP), produced by the DNA sensor cyclic GMP–AMP synthase (cGAS), which is important for the innate immune response to microbial infection and autoimmune disease. Modes of STING activation that are independent of cGAS are much less well understood. Here, through a spatiotemporally resolved

proximity labelling screen followed by quantitative proteomics, we identify the lysosomal membrane protein Niemann–Pick type C1 (NPC1) as a cofactor in the trafficking of STING. NPC1 interacts with STING and recruits it to the lysosome for degradation in both human and mouse cells. Notably, we find that knockout of *Npc1* ‘primes’ STING signalling by physically linking or ‘tethering’ STING to SREBP2 trafficking. Loss of NPC1 protein also ‘boosts’ STING signalling by blocking lysosomal degradation. Both priming and boosting of STING signalling are required for severe neurological disease in the *Npc1*<sup>−/−</sup> mouse. Genetic deletion of *Sting1* (the gene that encodes STING) or *Irf3*, but not that of *Cgas*, significantly reduced the activation of microglia and relieved the loss of Purkinje neurons in the cerebellum of *Npc1*<sup>−/−</sup> mice, leading to improved motor function. Our study identifies a cGAS- and cGAMP-independent mode of STING activation that affects neuropathology and provides a therapeutic target for the treatment of Niemann–Pick disease type C.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

# Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

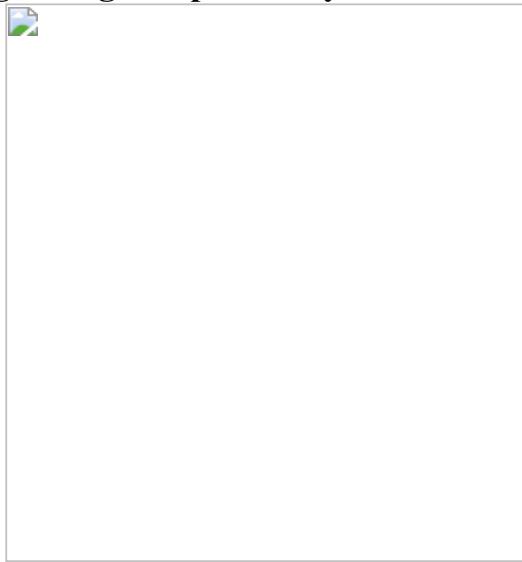
\$19.99/month

[Learn more](#)

## Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: NPC1 deficiency chronically activates STING-mediated type I IFN signalling independently of cGAS.**



**Fig. 2: SREBP2 trafficking primes STING trafficking and activation in *Npc1*-deficient cells.**



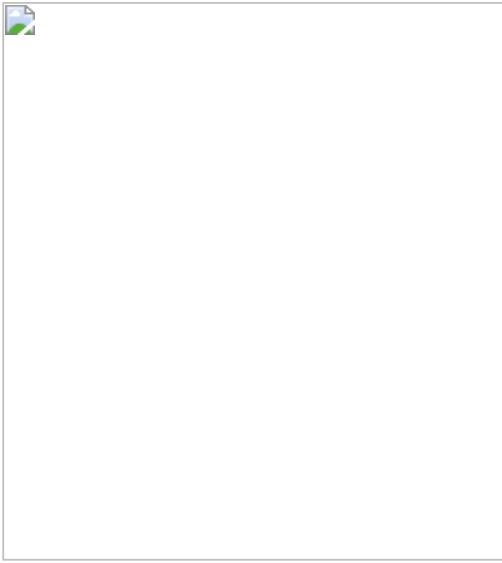
**Fig. 3: NPC1 is a lysosomal adaptor that mediates STING degradation.**



**Fig. 4: STING drives neuroinflammation and neuropathology in *Npc1*<sup>-/-</sup> mice.**



**Fig. 5: IRF3, but not cGAS, is required for disease pathogenesis in *Npc1*<sup>-/-</sup> mice.**



## Data availability

There are no restrictions on data availability in this manuscript. All of the information is included in the manuscript. Original western blots and TMT-MS raw data are included as Supplementary Fig. 1 and Supplementary Data, respectively.

## Change history

- **30 July 2021**

This Article was amended to correct the linking to the Supplementary Data and Peer Review File

## References

1. 1.  
Chin, A. C. Neuroinflammation and the cGAS–STING pathway. *J. Neurophysiol.* **121**, 1087–1091 (2019).
2. 2.  
Yan, N. Immune diseases associated with TREX1 and STING dysfunction. *J. Interferon Cytokine Res.* **37**, 198–206 (2017).
3. 3.

Sliter, D. A. et al. Parkin and PINK1 mitigate STING-induced inflammation. *Nature* **561**, 258–262 (2018).

4. 4.

Yang, K., Huang, R., Fujihira, H., Suzuki, T. & Yan, N. N-glycanase NGLY1 regulates mitochondrial homeostasis and inflammation through NRF1. *J. Exp. Med.* **215**, 2600–2616 (2018).

5. 5.

Yu, C.-H. et al. TDP-43 triggers mitochondrial DNA release via mPTP to activate cGAS/STING in ALS. *Cell* **183**, 636–649 (2020).

6. 6.

Wu, J. et al. STING-mediated disruption of calcium homeostasis chronically activates ER stress and primes T cell death. *J. Exp. Med.* **216**, 867–883 (2019).

7. 7.

Warner, J. D. et al. STING-associated vasculopathy develops independently of IRF3 in mice. *J. Exp. Med.* **214**, 3279–3292 (2017).

8. 8.

Luksch, H. et al. STING-associated lung disease in mice relies on T cells but not type I interferon. *J. Allergy Clin. Immunol.* **144**, 254–266.e8 (2019).

9. 9.

Motwani, M. et al. Hierarchy of clinical manifestations in SAVI N153S and V154M mouse models. *Proc. Natl Acad. Sci. USA* **116**, 7941–7950 (2019).

10. 10.

Dobbs, N. et al. STING activation by translocation from the ER is associated with infection and autoinflammatory disease. *Cell Host Microbe* **18**, 157–168 (2015).

11. 11.

Barber, G. N. STING-dependent signaling. *Nat. Immunol.* **12**, 929–930 (2011).

12. 12.

Wu, J., Dobbs, N., Yang, K. & Yan, N. Interferon-independent activities of mammalian STING mediate antiviral response and tumor immune evasion. *Immunity* **53**, 115–126 (2020).

13. 13.

Gonugunta, V. K. et al. Trafficking-mediated STING degradation requires sorting to acidified endolysosomes and can be targeted to enhance anti-tumor response. *Cell Rep.* **21**, 3234–3242 (2017).

14. 14.

Pokatayev, V. et al. Homeostatic regulation of STING protein at the resting state by stabilizer TOLLIP. *Nat. Immunol.* **21**, 158–167 (2020).

15. 15.

Hung, V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456–475 (2016).

16. 16.

Gui, X. et al. Autophagy induction via STING trafficking is a primordial function of the cGAS pathway. *Nature* **567**, 262–266 (2019).

17. 17.

Srikanth, S. et al. The  $\text{Ca}^{2+}$  sensor STIM1 regulates the type I interferon response by retaining the signaling adaptor STING at the endoplasmic reticulum. *Nat. Immunol.* **20**, 152–162 (2019).

18. 18.

Brown, M. S., Radhakrishnan, A. & Goldstein, J. L. Retrospective on cholesterol homeostasis: the central role of Scap. *Annu. Rev. Biochem.* **87**, 783–807 (2018).

19. 19.

Chen, W. et al. ER adaptor SCAP translocates and recruits IRF3 to perinuclear microsome induced by cytosolic microbial DNAs. *PLoS Pathog.* **12**, e1005462

(2016).

20. 20.

Taylor, A. M., Liu, B., Mari, Y., Liu, B. & Repa, J. J. Cyclodextrin mediates rapid changes in lipid balance in *Npc1*<sup>-/-</sup> mice without carrying cholesterol through the bloodstream. *J. Lipid Res.* **53**, 2331–2342 (2012).

21. 21.

Duncan, E. A., Brown, M. S., Goldstein, J. L. & Sakai, J. Cleavage site for sterol-regulated protease localized to a Leu–Ser bond in the luminal loop of sterol regulatory element-binding protein-2. *J. Biol. Chem.* **272**, 12778–12785 (1997).

22. 22.

Hua, X. et al. SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. *Proc. Natl Acad. Sci. USA* **90**, 11603–11607 (1993).

23. 23.

Li, X. et al. Structure of human Niemann–Pick C1 protein. *Proc. Natl Acad. Sci. USA* **113**, 8212–8217 (2016).

24. 24.

Chu, B.-B. et al. Cholesterol transport through lysosome–peroxisome membrane contacts. *Cell* **161**, 291–306 (2015).

25. 25.

Fog, C. K. & Kirkegaard, T. Animal models for Niemann–Pick type C: implications for drug discovery & development. *Expert Opin. Drug Discov.* **14**, 499–509 (2019).

26. 26.

Elrick, M. J. et al. Conditional Niemann–Pick C mice demonstrate cell autonomous Purkinje cell neurodegeneration. *Hum. Mol. Genet.* **19**, 837–847 (2010).

27. 27.

Guyenet, S. J. et al. A simple composite phenotype scoring system for evaluating mouse models of cerebellar ataxia. *J. Vis. Exp.* **39**, e1787 (2010).

28. 28.

Shin, S. D. K. et al. Interferon downstream signaling is activated early in pre-symptomatic Niemann–Pick disease type C. *Neurosci. Lett.* **706**, 43–50 (2019).

29. 29.

Parra, J. et al. Npc1 deficiency in the C57BL/6J genetic background enhances Niemann–Pick disease type C spleen pathology. *Biochem. Biophys. Res. Commun.* **413**, 400–406 (2011).

30. 30.

Cougnoux, A. et al. Microglia activation in Niemann–Pick disease, type C1 is amendable to therapeutic intervention. *Hum. Mol. Genet.* **27**, 2076–2089 (2018).

31. 31.

Sleat, D. E. et al. Genetic evidence for nonredundant functional cooperativity between NPC1 and NPC2 in lipid transport. *Proc. Natl Acad. Sci. USA* **101**, 5886–5891 (2004).

32. 32.

Dixit, S. S. et al. Loss of Niemann–Pick C1 or C2 protein results in similar biochemical changes suggesting that these proteins function in a common lysosomal pathway. *PLoS ONE* **6**, e23677 (2011).

33. 33.

Decout, A., Katz, J. D., Venkatraman, S. & Ablasser, A. The cGAS–STING pathway as a therapeutic target in inflammatory diseases. *Nat. Rev. Immunol.* <https://doi.org/10.1038/s41577-021-00524-z> (2021).

## Acknowledgements

We thank G. Barber for *Sting*<sup>−/−</sup> mice, K. Fitzgerald and T. Taniguchi for *Irf3*<sup>−/−</sup> mice, Z. Chen for *Cgas*<sup>−/−</sup> and *Mavs*<sup>−/−</sup> mice and J. Dietschy for *Npc2*<sup>+−</sup> mice; D. Ramirez, W. Li and A. Nawaby at the University of Texas Southwestern (UTSW) Whole Brain Microscopy Core Facility (RRID:SCR\_017949) for assistance with tissue sectioning, histological staining, immunostaining and slide scanning; A. Lemoff at the UTSW Proteomics Core Facility for assistance with TMT-MS experiments; R. Hammer and M. Nguyen at the UTSW Transgenic Technology Center for creating *Npc1*<sup>−/−</sup>*Sting*<sup>−/−</sup> mice; K. Phelps and A. Bugde at the UTSW Live Cell Imaging Facility, a Shared Resource of the Harold C. Simmons Cancer Center, supported in part by NCI Cancer Center Support Grant 1P30 CA142543-01 and National Institutes of Health (NIH) Shared Instrumentation Award 1S10 OD021684-01; I. Raman at the UTSW Medical Center Genomics and Microarray Core; and P. Tsai and members of the Yan laboratory for discussions. This work was supported by the NIH (AI151708 and NS122825 to N.Y.), the Cancer Prevention and Research Institute of Texas (CPRIT; RP180288 to N.Y.), the Burroughs Wellcome Fund (N.Y.) and the Ara Parseghian Medical Research Foundation (J.J.R.).

## Author information

### Affiliations

1. Department of Immunology, University of Texas Southwestern Medical Center, Dallas, TX, USA

Ting-Ting Chu, Xintao Tu, Kun Yang, Jianjun Wu & Nan Yan

2. Department of Physiology, University of Texas Southwestern Medical Center, Dallas, TX, USA

Joyce J. Repa

3. Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, USA

Joyce J. Repa

4. Department of Microbiology, University of Texas Southwestern Medical Center, Dallas, TX, USA

Nan Yan

## Contributions

T.-T.C. and N.Y. conceived and designed the study. T.-T.C. performed most of the experiments. X.T. helped with siRNA validation of the primary proteomic screen and in vitro experiments. K.Y. and J.W. helped with in vitro experiments. J.J.R. provided BALB/c wild-type, *Npc1*<sup>-/-</sup> and *Npc2*<sup>-/-</sup> mice and provided reagents. T.-T.C. and N.Y. wrote the paper with inputs from all co-authors.

## Corresponding author

Correspondence to [Nan Yan](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Veit Hornung and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

## Extended data figures and tables

### [Extended Data Fig. 1 Screen of STING trafficking cofactors using spatiotemporally resolved proximity labelling and quantitative proteomics.](#)

**a**, Diagram for proximity labelling and proteomic discovery of trafficking cofactors using *Sting1*<sup>-/-</sup> MEFs stably expressing STING–APEX2. **b**, Representative microscopy images of STING–APEX2 trafficking at various time points after stimulation (HT-DNA, 1 µg ml<sup>-1</sup>). BFA blocks the exit of STING from the ER; BafA1 blocks the degradation of STING by lysosomes. Scale bars, 10 µm. **c**, Immunoblot analysis of cell lysates. STING–APEX2 MEFs were mock-treated or stimulated with HT-DNA (1 µg ml<sup>-1</sup>) for the indicated times with or without treatment with BFA or BafA1 (top). Then, proximity labelling was performed and biotinylated proteins were detected by streptavidin–HRP. Immunoblot is

representative of at least three independent experiments. **d**, TMT-MS quantitative proteomics data filtering from two independent experiments and candidate discovery. **e**, Diagram of the STING trafficking route after stimulation, including time points (top) and organelles. **f**, Heat map showing selected STING trafficking cofactor candidates at each organelle (top). MS value was normalized to the 0-h time point. Data are representative of three independent experiments.

### Extended Data Fig. 2 NPC1 deficiency primes STING trafficking and activation independently of cGAMP.

**a**, qRT-PCR analysis of the baseline expression of ISGs (*Oas3*, *Oas1a* and *Usp18*) in *Npc1*<sup>WT</sup>, *Npc1*<sup>KO</sup> and *Npc1*<sup>KO</sup> MEFs stably expressing wild-type *Npc1* ( $n = 4$ ). **b**, qRT-PCR analysis of *Ifnb1* mRNA expression in wild-type and *Npc1*-knockdown MEFs after stimulation with an increasing amount of Poly I:C (0, 1, 2, 4  $\mu$ g ml<sup>-1</sup>) for 3 h ( $n = 3$ ). **c**, Immunoblot analysis of the indicated proteins (left) in *Npc1*<sup>WT</sup>, *Npc1*<sup>KO</sup> and *Npc1*<sup>KO</sup>*Sting1*<sup>KO</sup> MEFs. **d**, Heat map showing the baseline expression of ISGs in *Npc1*<sup>WT</sup>, *Npc1*<sup>KO</sup> and *Npc1*<sup>KO</sup>*Sting1*<sup>KO</sup> MEFs reconstituted with vector, STING wild type (FL), R232A or S366A mutants ( $n = 2$ ). The mRNA expression of each ISG was measured by qRT-PCR. **e**, Immunoblot analysis of the indicated proteins (left) in different reconstituted MEFs (shown on top) as in **d**. **f**, Fluorescent microscopy analysis of endogenous STING localization at the resting state in *Npc1*<sup>WT</sup> and two independent clones of *Npc1*<sup>KO</sup> MEFs. Endogenous STING is shown in green, an ERGIC marker (ERGIC53) is shown in red and DAPI is shown in blue. Scale bars, 10  $\mu$ m. **g**, Quantification of STING co-localization with the ERGIC in **f** (*Npc1*<sup>WT</sup>,  $n = 13$ ; *Npc1*<sup>KO-1</sup>,  $n = 12$ ; *Npc1*<sup>KO-2</sup>,  $n = 13$ ). **h**, qRT-PCR analysis of the baseline expression of ISGs in fibroblasts from healthy control individuals ( $n = 2$ ) and unrelated patients with Niemann–Pick disease type C (*NPC1*<sup>H061T</sup>) ( $n = 4$ ). **a, b, g**, Data are mean  $\pm$  s.d. **a, g**, Unpaired two-tailed Student's *t*-test. Data are representative of at least two independent experiments.

### Extended Data Fig. 3 SREBP2 trafficking primes STING signalling in *Npc1*<sup>KO</sup> cells independently of its transcriptional activity.

**a**, Diagram showing mechanisms of action for HP- $\beta$ -CD and triparanol on cholesterol synthesis and SREBP2 activation. HP- $\beta$ -CD promotes the egress of lysosomal cholesterol to the ER, thus limiting SREBP2 trafficking and activation in *Npc1*<sup>KO</sup> cells. Triparanol inhibits the biochemical conversion of desmosterol into cholesterol, thus promoting SREBP2 trafficking and activation in wild-type cells. **b**, **c**, qRT-PCR analysis of cholesterol-synthesis genes (**b**) and ISGs (**c**) in *Npc1*<sup>WT</sup> and *Npc1*<sup>KO</sup> MEFs mock-treated or treated with HP- $\beta$ -CD (0.3 mM) overnight ( $n = 3$ ). **d**,

**e**, qRT–PCR analysis of mock-, DMXAA- ( $50 \mu\text{g ml}^{-1}$ , 2 h) or poly(I:C)- ( $1 \mu\text{g ml}^{-1}$ , 2 h) induced expression of cholesterol-synthesis genes (**d**) and ISGs (**e**) with mock- or HP- $\beta$ -CD treatment (0.3 mM, 8 h) in  $Npc1^{WT}$  MEFs ( $n = 3$ ). **f**, qRT–PCR analysis of the baseline expression of ISGs in  $Npc1^{WT}$  and  $Npc1^{KO}$  MEFs transfected with siCtrl or siSrebf2 for 48 h ( $n = 3$ ). **g**, Immunoblot analysis of the indicated proteins (left) in  $Npc1^{WT}$ ,  $Npc1^{KO}$ ,  $Npc1^{KO}Srebf2^{KD}$  and  $Npc1^{KO}Srebf2^{KD}$  MEFs reconstituted with SREBP2 wild type (FL) or transcription-inactive mutants (L511A/S512A,  $\Delta$ bHLH). **h**, **i**, qRT–PCR analysis of the expression of cholesterol-synthesis genes (**h**) and ISGs (**i**) in  $Npc1^{WT}$ ,  $Npc1^{KO}$ ,  $Npc1^{KO}Srebf2^{KD}$  and  $Npc1^{KO}Srebf2^{KD}$  MEFs reconstituted with SREBP2 wild type (FL) or transcription-inactive mutants (L511A/S512A,  $\Delta$ bHLH) ( $n = 3$ ). **b–f**, **h**, **i**, Data are mean  $\pm$  s.d. Unpaired two-tailed Student’s *t*-test. NS, not significant. Data are representative of at least two independent experiments.

#### Extended Data Fig. 4 SREBP2 primes STING signalling independently of cGAS or MAVS.

**a**, Immunoblot analysis of the SREBP2 and STING activation in  $Npc1^{WT}$  MEFs that were treated with mock or triparanol ( $14 \mu\text{M}$ ) for the indicated times (shown on the top). SREBP2 cleavage indicates activation. STING activation induces pSTING and pTBK1. **b**, Fluorescent microscopy analysis of SREBP2 and STING activation. Cleaved SREBP2 translocates to the nucleus (red). The STING activation markers pSTING, pTBK1 and pIRF3 are shown in green. The nucleus dye DAPI is shown in blue. Scale bars,  $10 \mu\text{m}$ . **c**, qRT–PCR analysis of cholesterol-synthesis genes in BMDMs that were either mock-treated or treated with triparanol ( $14 \mu\text{M}$ ) for 12 h ( $n = 3$ ). **d**, qRT–PCR analysis of the baseline expression of ISGs in BMDMs that were either mock-treated or treated with triparanol ( $14 \mu\text{M}$ ) alone or in combination with the STING inhibitor C-176 ( $0.5 \mu\text{M}$ ) for 12 h ( $n = 3$ ). **e**, **f**, qRT–PCR analysis of cholesterol-synthesis genes (**e**) and ISGs (**f**) in wild-type,  $Sting1^{-/-}$  and  $Cgas^{-/-}$  BMDMs that were either mock-treated or treated with triparanol ( $14 \mu\text{M}$ ) for 12 h ( $n = 3$ ). **g**, **h**, qRT–PCR analysis of the expression of cholesterol-synthesis genes (**g**) and ISGs (**h**) in  $Mavs^{-/-}$  MEFs that were either mock-treated or treated with triparanol ( $14 \mu\text{M}$ ) for 12 h ( $n = 3$ ). **i**, **j**, qRT–PCR analysis of the cholesterol-synthesis gene  $Hmgcs1$ , knockdown efficiency of  $Insig1$  (**i**) and the expression of ISGs (**j**) in wild-type,  $Cgas^{-/-}$  and  $Sting1^{-/-}$  MEFs ( $n = 3$ ). **c–j**, Data are mean  $\pm$  s.d. Unpaired two-tailed Student’s *t*-test. Data are representative of at least two independent experiments.

#### Extended Data Fig. 5 NPC1 interacts with STING and mediates the lysosomal degradation of STING.

**a**, Quantification of immunoblots in Fig. 3a. **b**, Immunoblot analysis of STING degradation in *Npc1*<sup>WT</sup>, *Npc1*<sup>KO</sup> or *Npc1*<sup>KO</sup> MEFs stably expressing wild-type NPC1. Cells were mock-treated or stimulated with HT-DNA (1 µg ml<sup>-1</sup>) for 8 h. **c**, **d**, Live-cell microscopy images of STING–EGFP and lysosomes. *Npc1*<sup>WT</sup> and *Npc1*<sup>KO</sup> MEFs stably expressing STING–EGFP were mock-treated (left) or stimulated with HT-DNA (right, 4 µg ml<sup>-1</sup>). STING–EGFP is shown in green and LysoTracker-Red is shown in red. Scale bar, 10 µm. Representative images in **c** and quantifications in **d**. *Npc1*<sup>WT</sup> 0 h (*n* = 13), *Npc1*<sup>KO</sup> 0 h (*n* = 17), *Npc1*<sup>WT</sup> 15 h (*n* = 15), and *Npc1*<sup>KO</sup> 15 h (*n* = 16) for quantification. Data are mean ± s.d. Unpaired two-tailed Student's *t*-test. **e**, Immunoblot analysis of NPC1 and STING co-immunoprecipitation in HEK293T cells. HEK293T cells were transfected with the indicated plasmids (top), and 24 h later, anti-IgG (mouse) or anti-Flag was used for the pull-down. HA–STING co-immunoprecipitation was analysed by anti-STING immunoblot. Whole-cell lysates were blotted by anti-Flag (NPC1), anti-STING and anti-tubulin as input. **f**, Immunoblot analysis of endogenous STING and NPC1 co-immunoprecipitation in wild-type MEFs. Anti-IgG (rabbit) or anti-NPC1 (rabbit) was used for the pull-down. Both immunoprecipitation and lysate were blotted for endogenous STING (rabbit), NPC1 (mouse) and tubulin. **g**, Immunoblot analysis of endogenous STING and NPC1 co-immunoprecipitation in wild-type MEFs with mock-treated, DMXAA (30 µg ml<sup>-1</sup>, 8 h)-treated or DMXAA combined with BafA1-treated (to prevent STING degradation). Anti-IgG (rabbit) or Anti-NPC1 (rabbit) was used for the pull-down. Both immunoprecipitation and lysate input were blotted for endogenous STING (rabbit) and NPC1 (mouse). **h**, Immunoblot analysis of endogenous STING interaction with NPC1 or LAMP1 in wild-type MEFs. Anti-IgG (rabbit) or anti-STING (rabbit) was used for the pull-down. Both immunoprecipitation and lysate input were blotted for endogenous STING, NPC1 or LAMP1. Red arrow, co-immunoprecipitated NPC1 band. **i**, Live-cell fluorescent microscopy analysis of STING–EGFP and NPC1–mCherry localization. *Npc1*<sup>KO</sup>*Sting1*<sup>KO</sup> MEFs stably expressing STING–EGFP (green) and NPC1–mCherry (red) were stimulated with HT-DNA (4 µg ml<sup>-1</sup>) and imaged at the indicated times. Scale bars, 10 µm. **j**, **k**, Diagrams showing STING (**j**) and NPC1 (**k**) full length and truncations. **l**, Immunoblot analysis of STING degradation in *Npc1*<sup>WT</sup>, *Npc1*<sup>KO</sup> or *Npc1*<sup>KO</sup> MEFs stably expressing indicated NPC1 truncations. Cells were mock-treated or stimulated with HT-DNA (1 µg ml<sup>-1</sup>) for 8 h. Cell lysates were analysed for the proteins indicated on the left. **m**, Immunoblot analysis of STING interaction with NPC1 transmembrane bundles in HEK293T cells. HEK293T cells were transfected with the indicated plasmids (top), and 24 h later, anti-IgG (mouse) or anti-Flag (mouse) was used for the pull-down. Both immunoprecipitation and whole-cell lysates were analysed by anti-Flag (NPC1), anti-STING and anti-tubulin. Data are representative of at least two independent experiments.

## Extended Data Fig. 6 STING drives neuroinflammation in *Npc1*<sup>-/-</sup> mice.

**a**, Immunoblot analysis of the indicated proteins (left) in whole-brain lysates from BALB/c *Npc1*<sup>+/+</sup>, *Npc1*<sup>-/-</sup> and *Npc1*<sup>-/-</sup>*Sting1*<sup>-/-</sup> mice ( $n = 2$ ). **b**, qRT–PCR analysis of the expression of ISGs in the cerebellum of BALB/c *Npc1*<sup>+/+</sup>, *Npc1*<sup>-/-</sup> and *Npc1*<sup>-/-</sup>*Sting1*<sup>-/-</sup> mice ( $n = 4$ ). **c**, Heat map showing the expression of ISGs in BMDMs from BALB/c *Npc1*<sup>+/+</sup>, *Npc1*<sup>-/-</sup> and *Npc1*<sup>-/-</sup>*Sting1*<sup>-/-</sup> mice ( $n = 3$ ). The mRNA expression level of each ISG was measured by qRT–PCR. **b**, Data are mean ± s.d. Unpaired two-tailed Student’s *t*-test. Data are representative of at least two independent experiments.

## Extended Data Fig. 7 STING and IRF3, but not cGAS, are required for *Npc1*<sup>-/-</sup> disease pathogenesis.

**a**, Representative image for the body size of eight-week-old mice. **b**, Immunoblot analysis of proteins (left) in whole-brain lysates of mice of the indicated genotypes (C57BL/6J;  $n = 3$ ). **c**, Heat map showing the expression of ISGs in BMDMs from mice of the indicated genotypes (C57BL/6J;  $n = 3$ ). The mRNA expression level of each ISG was measured by qRT–PCR. **d**, Serum cytokine levels in four-week-old mice of the indicated genotypes measured by multiplex ELISA ( $n = 4$ ). **e**, Heat map showing the expression of ISGs in *Npc1*<sup>+/+</sup> and *Npc1*<sup>-/-</sup> BMDMs that were mock-treated or treated with the STING inhibitor C-176 (0.5 μM) or C-178 (0.5 μM) overnight ( $n = 2$ ). The mRNA expression of each ISG was measured by qRT–PCR. **d**, Data are mean ± s.d. Unpaired two-tailed Student’s *t*-test. Data are representative of at least two independent experiments.

## Extended Data Fig. 8 STING function in Purkinje cells and microglia.

**a**, Progression of *Npc1*<sup>-/-</sup> neuropathology. Fluorescent IHC staining of mouse cerebellum of *Npc1*<sup>+/+</sup> and *Npc1*<sup>-/-</sup> mice at P7, four weeks old and eight weeks old. Calbindin is shown in red, CD68 is shown in green and DAPI is shown in blue. Representative images are shown ( $n = 3$ ). Bottom panels are enlarged views. Scale bars, 200 μm (top); 30 μm (bottom). **b**, STING antibody validation for IHC. Fluorescent IHC staining with IgG (rabbit), anti-STING (rabbit) in wild-type C57BL/6J cerebellum or staining with anti-STING (rabbit) in *Sting1*<sup>-/-</sup> C57BL/6J cerebellum. Scale bars, 30 μm. Representative images are shown ( $n = 3$ ). **c**, Heat map showing the baseline expression of ISGs in ex-vivo-cultured microglia isolated from P10–P12 mouse brains of the indicated genotypes ( $n = 3$ ). The mRNA

expression of each ISG was measured by qRT–PCR. Representative microscopy images of isolated microglia from indicated genotype of mice are shown at the bottom. Scale bars, 30 µm.

### **Extended Data Fig. 9 STING signalling activation in *Npc1*<sup>-/-</sup> and *Npc2*<sup>-/-</sup> cells and mice.**

**a**, Immunoblot analysis of proteins (left) in whole-brain lysates of BALB/c mice of the indicated genotypes ( $n = 3$ ). **b**, qRT–PCR analysis of cholesterol-synthesis gene (*Hmgcr*) expression in the cerebellums of mice of the indicated genotypes (BALB/c) ( $n = 3$ ). **c**, Heat map showing the expression of ISGs in BALB/c *Npc1*<sup>+/+</sup>, *Npc1*<sup>-/-</sup> and *Npc2*<sup>-/-</sup> mouse cerebellum ( $n = 3$ ). The mRNA expression of each ISG was measured by qRT–PCR. **d**, qRT–PCR analysis of the expression of some ISGs in the cerebellum of mice of the indicated genotypes (BALB/c) ( $n = 3$ ). **e**, Immunoblot analysis of the STING signalling cascade. BALB/c *Npc1*<sup>+/+</sup>, *Npc1*<sup>-/-</sup> and *Npc2*<sup>-/-</sup> mouse BMDMs were stimulated with DMXAA (50 µg ml<sup>-1</sup>) for 0, 2 or 4 h. Phosphorylated and total proteins blotted are shown on the left. **f**, qRT–PCR analysis of the baseline expression of ISGs in *Npc2*<sup>WT</sup>, *Npc2*<sup>KD</sup> and *Npc2*<sup>KD</sup>*Sting1*<sup>KD</sup> MEFs ( $n = 3$ ). **g**, qRT–PCR analysis of the baseline expression of ISGs in BALB/c *Npc2*<sup>+/+</sup> and *Npc2*<sup>-/-</sup> mouse BMDMs that were mock-treated or treated with STING inhibitor C-176 (0.5 µM) overnight ( $n = 3$ ). **b**, **d**, **f**, **g**, Data are mean ± s.d. Unpaired two-tailed Student's *t*-test. Data are representative of at least two independent experiments.

## **Supplementary information**

### **Supplementary Figure 1**

This file contains Western blot full scans.

### **Reporting Summary**

### **Supplementary Data**

This file contains the STING-APEX2 TMT-MS raw data.

### **Peer Review File**

# Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Chu, TT., Tu, X., Yang, K. *et al.* Tonic prime-boost of STING signalling mediates Niemann–Pick disease type C. *Nature* **596**, 570–575 (2021).  
<https://doi.org/10.1038/s41586-021-03762-2>

- Received: 12 October 2020
- Accepted: 23 June 2021
- Published: 21 July 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03762-2>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03762-2>

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

- Article
- [Published: 11 August 2021](#)

# Cycling cancer persister cells arise from lineages with distinct programs

- [Yaara Oren<sup>1,2</sup>](#),
- [Michael Tsabar](#) ORCID: [orcid.org/0000-0001-9470-6167](#)<sup>1,3,4 na1</sup>,
- [Michael S. Cuoco](#) ORCID: [orcid.org/0000-0003-2163-5120](#)<sup>1 na1</sup>,
- [Liat Amir-Zilberstein<sup>1</sup>](#),
- [Heidie F. Cabanos<sup>5,6</sup>](#),
- [Jan-Christian Hütter](#) ORCID: [orcid.org/0000-0002-1219-4821](#)<sup>1</sup>,
- [Bomiao Hu<sup>7</sup>](#),
- [Pratiksha I. Thakore<sup>1 nAff18</sup>](#),
- [Marcin Tabaka<sup>1</sup>](#),
- [Charles P. Fulco<sup>8 nAff19</sup>](#),
- [William Colgan](#) ORCID: [orcid.org/0000-0002-8328-4672](#)<sup>8</sup>,
- [Brandon M. Cuevas](#) ORCID: [orcid.org/0000-0002-8275-8258](#)<sup>1</sup>,
- [Sara A. Hurvitz<sup>9,10</sup>](#),
- [Dennis J. Slamon<sup>9,10</sup>](#),
- [Amy Deik<sup>11</sup>](#),
- [Kerry A. Pierce<sup>11</sup>](#),
- [Clary Clish](#) ORCID: [orcid.org/0000-0001-8259-9245](#)<sup>11</sup>,
- [Aaron N. Hata](#) ORCID: [orcid.org/0000-0002-6127-318X](#)<sup>5,6</sup>,
- [Elma Zaganjor<sup>12</sup>](#),
- [Galit Lahav<sup>3</sup>](#),
- [Katerina Politi<sup>13,14</sup>](#),
- [Joan S. Brugge](#) ORCID: [orcid.org/0000-0002-2547-4814](#)<sup>2,15 na2</sup> &
- [Aviv Regev](#) ORCID: [orcid.org/0000-0003-3293-3158](#)<sup>1,16,17 na2 nAff18</sup>

*Nature* volume **596**, pages 576–582 (2021)

- 17k Accesses
- 1 Citations

- 219 Altmetric
- [Metrics details](#)

## Subjects

- [Cell-cycle exit](#)
- [Targeted therapies](#)

## Abstract

Non-genetic mechanisms have recently emerged as important drivers of cancer therapy failure<sup>1</sup>, where some cancer cells can enter a reversible drug-tolerant persister state in response to treatment<sup>2</sup>. Although most cancer persisters remain arrested in the presence of the drug, a rare subset can re-enter the cell cycle under constitutive drug treatment. Little is known about the non-genetic mechanisms that enable cancer persisters to maintain proliferative capacity in the presence of drugs. To study this rare, transiently resistant, proliferative persister population, we developed Watermelon, a high-complexity expressed barcode lentiviral library for simultaneous tracing of each cell's clonal origin and proliferative and transcriptional states. Here we show that cycling and non-cycling persisters arise from different cell lineages with distinct transcriptional and metabolic programs. Upregulation of antioxidant gene programs and a metabolic shift to fatty acid oxidation are associated with persister proliferative capacity across multiple cancer types. Impeding oxidative stress or metabolic reprogramming alters the fraction of cycling persisters. In human tumours, programs associated with cycling persisters are induced in minimal residual disease in response to multiple targeted therapies. The Watermelon system enabled the identification of rare persister lineages that are preferentially poised to proliferate under drug pressure, thus exposing new vulnerabilities that can be targeted to delay or even prevent disease recurrence.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

## Rent or Buy

All prices are NET prices.

## Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

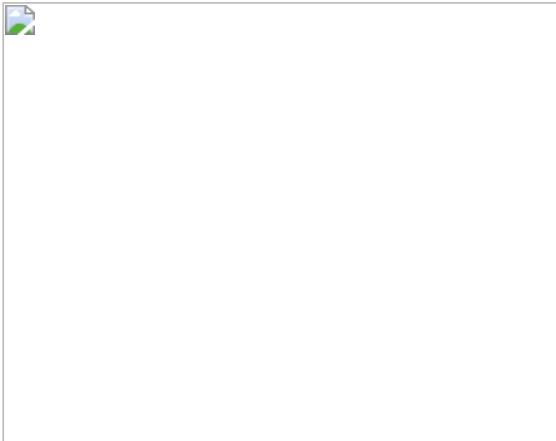
\$19.99/month

Learn more

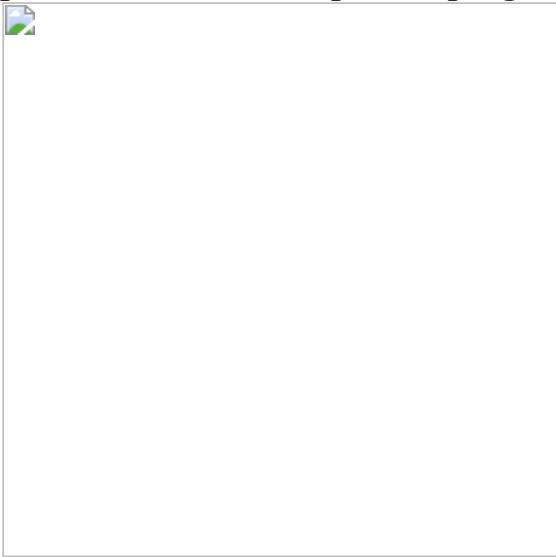
## Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

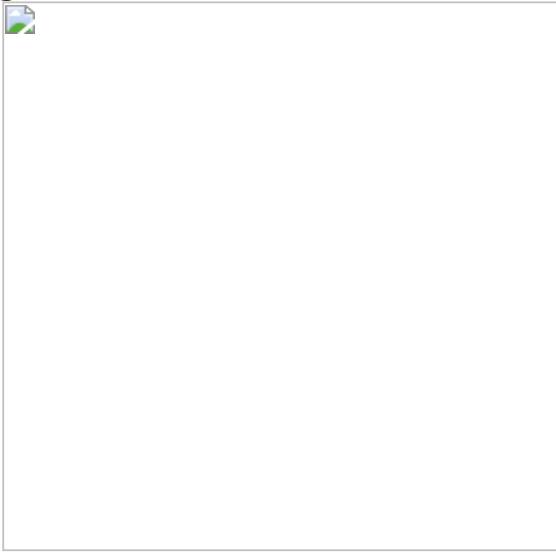
**Fig. 1: Persister cells contain a rare proliferative subpopulation.**



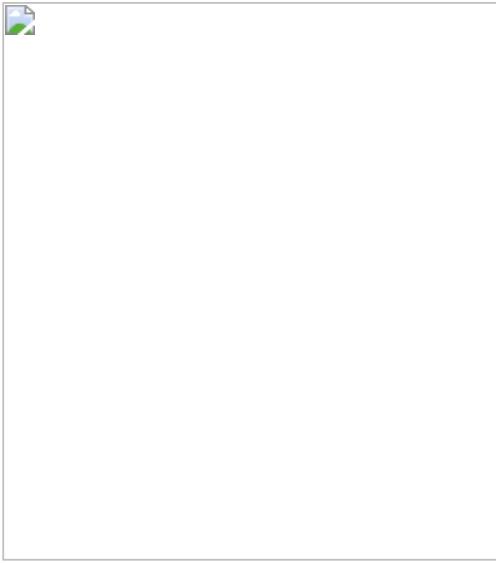
**Fig. 2: Cycling and non-cycling persisters arise from different cell lineages that express distinct transcriptional programs.**



**Fig. 3: Persister cells shift their metabolism to FAO.**



**Fig. 4: Metabolic shift in tumours treated with oncogene-targeted therapy.**



## Data availability

RNA-seq data have been deposited in the NCBI Genome Expression Omnibus (GEO) under the accession code [GSE150949](#). The Watermelon library and plasmid are available on Addgene (Addgene IDs 155257 and 155258). [Source data](#) are provided with this paper.

## Code availability

Code used in this study is available from <https://github.com/yaaraore/watermelon> (additional code is available upon request).

## References

1. 1.

Salgia, R. & Kulkarni, P. The genetic/non-genetic duality of drug ‘resistance’ in cancer. *Trends Cancer* **4**, 110–118 (2018).

2. 2.

Valette, F. M. et al. Dormant, quiescent, tolerant and persister cells: four synonyms for the same target in cancer. *Biochem. Pharmacol.* **162**, 169–176 (2019).

3. 3.

Sharma, S. V. et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141**, 69–80 (2010).

4. 4.

Hata, A. N. et al. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.* **22**, 262–269 (2016).

5. 5.

Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

6. 6.

Taniguchi, H. et al. AXL confers intrinsic resistance to osimertinib and advances the emergence of tolerant cells. *Nat. Commun.* **10**, 259 (2019).

7. 7.

Dhimolea, E. et al. An embryonic diapause-like adaptation with suppressed Myc activity enables tumor treatment persistence. *Cancer Cell* **39**, 240–256 (2021).

8. 8.

Rehman, S. K. et al. Colorectal cancer cells enter a diapause-like DTP state to survive chemotherapy. *Cell* **184**, 226–242 (2021).

9. 9.

Raoof, S. et al. Targeting FGFR overcomes EMT-mediated resistance in EGFR mutant non-small cell lung cancer. *Oncogene* **38**, 6399–6413 (2019).

10. 10.

Zhu, X., Chen, L., Liu, L. & Niu, X. EMT-Mediated Acquired EGFR-TKI Resistance in NSCLC: Mechanisms and Strategies. *Front. Oncol.* **9**, 1044 (2019).

11. 11.

Garon, E. B. et al. Antiestrogen fulvestrant enhances the antiproliferative effects of epidermal growth factor receptor inhibitors in human non-small-cell lung cancer. *J. Thorac. Oncol.* **8**, 270–278 (2013).

12. 12.

Hangauer, M. J. et al. Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. *Nature* **551**, 247–250 (2017).

13. 13.

Friedmann Angeli, J. P., Krysko, D. V. & Conrad, M. Ferroptosis at the crossroads of cancer-acquired drug resistance and immune evasion. *Nat. Rev. Cancer* **19**, 405–414 (2019).

14. 14.

Fox, D. B. et al. NRF2 activation promotes the recurrence of dormant tumour cells through regulation of redox and nucleotide metabolism. *Nat. Metab.* **2**, 318–334 (2020).

15. 15.

Foggetti, G. et al. Genetic determinants of EGFR-driven lung cancer growth and therapeutic response *in vivo*. *Cancer Discov.* **11**, 1736–1753 (2021).

16. 16.

Reczek, C. R. & Chandel, N. S. The two faces of reactive oxygen species in cancer. *Ann. Rev. Cancer Biol.* **1**, 79–98 (2017).

17. 17.

Zaugg, K. et al. Carnitine palmitoyltransferase 1C promotes cell survival and tumor growth under conditions of metabolic stress. *Genes Dev.* **25**, 1041–1051 (2011).

18. 18.

Kinker, G. S. et al. Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.* **52**, 1208–1218 (2020).

19. 19.

McFarland, J. M. et al. Multiplexed single-cell transcriptional response profiling to define cancer vulnerabilities and therapeutic mechanism of action. *Nat. Commun.* **11**, 4296 (2020).

20. 20.

Maynard, A. et al. Therapy-induced evolution of human lung cancer revealed by single-cell RNA sequencing. *Cell* **182**, 1232–1251 (2020).

21. 21.

Patel, A. P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).

22. 22.

Kwong, L. N. et al. Co-clinical assessment identifies patterns of BRAF inhibitor resistance in melanoma. *J. Clin. Invest.* **125**, 1459–1470 (2015).

23. 23.

Viswanathan, V. S. et al. Dependency of a therapy-resistant state of cancer cells on a lipid peroxidase pathway. *Nature* **547**, 453–457 (2017).

24. 24.

Shaffer, S. M. et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017).

25. 25.

Shah, K. N. et al. Aurora kinase A drives the evolution of resistance to third-generation EGFR inhibitors in lung cancer. *Nat. Med.* **25**, 111–118 (2019).

26. 26.

Havas, K. M. et al. Metabolic shifts in residual breast cancer drive tumor recurrence. *J. Clin. Invest.* **127**, 2091–2105 (2017).

27. 27.

Liau, B. B. et al. Adaptive chromatin remodeling drives glioblastoma stem cell plasticity and drug tolerance. *Cell Stem Cell* **20**, 233–246 (2017).

28. 28.

Guler, G. D. et al. Repression of stress-induced LINE-1 expression protects cancer cell subpopulations from lethal drug exposure. *Cancer Cell* **32**, 221–237 (2017).

29. 29.

Yang, C., Tian, C., Hoffman, T. E., Jacobsen, N. K. & Spencer, S. L. Melanoma subpopulations that rapidly escape MAPK pathway inhibition incur DNA damage and rely on stress signalling. *Nat. Commun.* **12**, 1747 (2021).

30. 30.

Reyes, J. et al. Fluctuations in p53 signaling allow escape from cell-cycle arrest. *Mol. Cell* **71**, 581–591 (2018).

31. 31.

Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).

32. 32.

Jerby-Arnon, L. et al. A cancer cell program promotes t cell exclusion and resistance to checkpoint blockade. *Cell* **175**, 984–997 (2018).

33. 33.

Chong, J., Yamamoto, M. & Xia, J. MetaboAnalystR 2.0: from raw spectra to biological insights. *Metabolites* **9**, E57 (2019).

34. 34.

Kang, H. M. et al. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).

35. 35.

Erik Garrison, G. M. Haplotype-based variant detection from short-read sequencing. Preprint at <https://arxiv.org/abs/1207.3907> (2012).

## Acknowledgements

We thank the Broad Cytometry Facility (P. Rogers, S. Saldi, C. Otis and N. Pirete); L. Gaffney and A. Hupalowska for help with figure preparation and artwork; A. Martinez Gakidis for scientific editing; the Yale Center for Genome Analysis (F. Lopez-Giraldez and D. Zhao); and the Yale Center for Research Computing for data processing, guidance, RNA-seq and use of the research computing infrastructure, specifically the Ruddle Cluster. Y.O. is supported by the Hope Fund for Cancer Research, Grillo-Marxuach Postdoctoral Fellowship and the Rivkin Scientific Scholar Award. M. Tsabar was supported by the American Cancer Society–New England Pay-if Group Postdoctoral Fellowship, PF-18-126-01-DMC. H.F.C. was supported by a V Foundation Scholar Award (D2015-027). P.I.T. is supported by an NIH F32 Postdoctoral Fellowship from National Institute of Allergy and Infectious Disease (1F32AI138458-01). Funding for breast cancer patient study came from grants from Sanofi and GlaxoSmithKline. A.N.H. was supported by NIH/NCI K08 CA197389. S.A.H. was supported in part by NCI/NIH CA016042 as well as the Marni Levine Memorial Research Award. B.H. received support from a Lo Graduate Fellowship for Excellence in Stem Cell Research. Additional support came from R01CA121210, R01CA120247 and P50CA196530. Yale Cancer Center Shared Resources used in this article were in part supported by NIH/NCI Cancer Center Support Grant P30 CA016359. The work was supported by the Klarman Cell Observatory, the NHGRI Center for Cell Circuits, and Howard Hughes Medical Institute (A.R.) as well as by the Breast Cancer Research Foundation- BCRF-16-020 and the Sheldon and Miriam Adelson Medical Research Foundation (J.S.B.).

## Author information

### Author notes

1. Pratiksha I. Thakore & Aviv Regev

Present address: Genentech, South San Francisco, CA, USA

2. Charles P. Fulco

Present address: Bristol Myers Squibb, Cambridge, MA, USA

3. These authors contributed equally: Michael Tsabar, Michael S. Cuoco

4. These authors jointly supervised this work: Joan S. Brugge, Aviv Regev

## Affiliations

1. Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Yaara Oren, Michael Tsabar, Michael S. Cuoco, Liat Amir-Zilberstein, Jan-Christian Hütter, Pratiksha I. Thakore, Marcin Tabaka, Brandon M. Cuevas & Aviv Regev

2. Department of Cell Biology, Harvard Medical School, Boston, MA, USA

Yaara Oren & Joan S. Brugge

3. Department of Systems Biology, Harvard Medical School, Boston, MA, USA

Michael Tsabar & Galit Lahav

4. Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA, USA

Michael Tsabar

5. Department of Medicine, Massachusetts General Hospital, Boston, MA, USA

Heidie F. Cabanos & Aaron N. Hata

6. Departments of Medicine, Harvard Medical School, Boston, MA, USA

Heidie F. Cabanos & Aaron N. Hata

7. Department of Pathology, Yale School of Medicine, New Haven, CT, USA

Bomiao Hu

8. Broad Institute of MIT and Harvard, Cambridge, MA, USA

Charles P. Fulco & William Colgan

9. David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

Sara A. Hurvitz & Dennis J. Slamon

10. Jonsson Comprehensive Cancer Center, Los Angeles, CA, USA

Sara A. Hurvitz & Dennis J. Slamon

11. Metabolomics Platform, Broad Institute, Cambridge, MA, USA

Amy Deik, Kerry A. Pierce & Clary Clish

12. Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA

Elma Zaganjor

13. Section of Medical Oncology, Department of Pathology, Yale School of Medicine, New Haven, CT, USA

Katerina Politi

14. Yale Cancer Center, New Haven, CT, USA

Katerina Politi

15. Ludwig Center at Harvard, Boston, MA, USA

Joan S. Brugge

16. Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

Aviv Regev

17. Howard Hughes Medical Institute, Chevy Chase, MD, USA

Aviv Regev

## Contributions

Study design, data interpretation and preparation of the manuscript: Y.O., J.S.B. and A.R. Execution of experiments: Y.O., M. Tsabar, H.F.C., M.S.C., L.A.-Z., E.Z., P.I.T., B.H., A.D., K.A.P. and B.M.C. Computational and statistical analysis: Y.O., J.-C.H., M.S.C., W.C., M. Tabaka and M. Tsabar . Resources: M. Tabaka, C.P.F., S.A.H., D.J.S., G.L., C.C., A.N.H. and K.P.

## Corresponding authors

Correspondence to [Joan S. Brugge](#) or [Aviv Regev](#).

## Ethics declarations

### Competing interests

A.R. is a co-founder and equity holder of Celsius Therapeutics, an equity holder in Immunitas, and was a scientific advisory board member of ThermoFisher Scientific, Syros Pharmaceuticals, Neogene Therapeutics and Asimov until 31 July 2020. From 1 August 2020, A.R. has been an employee of Genentech. J.S.B. is a consultant for Agios Pharmaceuticals, eFFECTOR Therapeutics, and Frontier Medicines. Y.O., A.R. and J.S.B. are inventors on US patent application 16/563,450 filed by the Broad Institute to expressed barcode libraries as described in this manuscript. C.P.F. is now an employee of Bristol Myers Squibb. K.P. is co-inventor on a patent licensed to Molecular MD for EGFR(T790M) mutation testing (through MSKCC). K.P. has received Honoraria/Consulting fees from Takeda, NCCN, Novartis, Merck, AstraZeneca, Tocagen, Maverick Therapeutics, Dynamo Therapeutics, Halda and research support from AstraZeneca, Kolltan, Roche Boehringer Ingelheim and Symphogen. A.N.H. a consultant for Nuvalent and is supported by Novartis, Pfizer, Amgen, Blueprint Medicines, Lilly, Roche/Genetech, Nuvalent, Relay Therapeutics. S.A.H. has contracted research with Ambrx, Amgen, Astra Zeneca, Arvinas, Bayer, Daiichi-Sankyo, Genentech/Roche, Gilead, GSK, Immunomedics, Lilly, Macrogenics, Novartis, Pfizer, OBI Pharma, Pieris, PUMA, Radius, Samumed, Sanofi, Seattle Genetics, Dignitana, Zymeworks, Phoenix Molecular Designs and Lilly, and stock options in NK Max.

### Additional information

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

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

### Extended data figures and tables

#### [Extended Data Fig. 1 Lineage detection efficacy and fluorescent dilution capacity of the Watermelon library.](#)

**a**, Viability of PC9 cells treated osimertinib. % of viable PC9 cells (y axis) after 72 h of treatment with osimertinib at different concentrations (x axis). **b**, Schematic of tracking. (1). Tracking of non-persister cells. Red arrows: cell that was tracked in the

lineage. (2). Tracking of persister cells. Colonies detected at day 14, and the time lapse is viewed and tracked from day 14 backward to day 0 to detect the common initiator cell. After the initiator cell is detected, this initiator is tracked forward as in **a** (red line indicates the tracked cell). **c**, Watermelon library complexity. Distribution of number of unique lineage barcodes (*y* axis, red bars) in a Watermelon plasmid library sequenced at a depth of  $\sim 68 \times 10^6$  reads. Blue curve: cumulative wealth distribution of unique barcodes. **d**, Watermelon library sequence diversity. Sequence logo of nucleotide composition at each position (*x* axis) relative to the beginning of the barcode sequence of 5,472,944 unique lineage barcodes detected in the Watermelon library. **e**, PC9-Watermelon cell line grown in dox containing media. A merge of the green, red and bright field channels is shown. Scale bar 20 $\mu$ m. **f**, Fluorescence dilution of H2B-mCherry over time reports proliferative history. Distributions of mCherry fluorescence level (*x* axis) for  $n = 3000$  cells analysed by flow cytometry at each time point (colour legend) from cells transduced with the Watermelon vector, exposed to dox for 48 h, sorted for red positive cells and seeded in separate wells at  $t = 0$  (Methods). Data are representative of two independent experiments

[Source data.](#)

### **Extended Data Fig. 2 scRNA-seq along a time course of osimertinib treated PC9-Watermelon cells.**

**a**, Sorting strategy. Distribution of mCherry fluorescence level (*x* axis) in Watermelon-PC9 cells gates at day 14 of osimertinib treatment, marked by representative sorting gates used to sort three persister subpopulations: cycling, moderate cyclers and non-cycling. **b**, Number of high-quality cells profiled in each sample. **c**, Changes in expression profiles following treatment. *t*-stochastic neighbourhood embedding (tSNE) of 56,419 PC9-Watermelon cell profiles (dots), coloured (red) by the labelled time point. **d**, Assignment of cells to lineages by lineage barcode. Percent of cells (*y* axis) at each time point/subpopulation (*x* axis) that have a detected lineage barcode. **e**, Identification of cycling cells. Percent of cells (*y* axis) at each time point/subpopulation (*x* axis) that express either the G2/M or S phase signature. **f**, Majority fate. Clone size on day 14 (*y* axis) of each persister lineage barcode inferred from scRNA-seq ordered by ascending rank order (*x* axis) and coloured by majority fate based on flow sample provenance of the captured cells

[Source data.](#)

### **Extended Data Fig. 3 Estimates of lineage diversity.**

**a**, Difference in number of cells profiled per time point. Number of cells (*y* axis) with captured lineage barcode at each day (*x* axis). Day 14 cells are partitioned by the three mCherry populations (legend). **b–e**, Species diversity estimators can be biased by coverage. Estimated sample coverage (cumulative proportion of all lineages in the total population that were observed, top, *y* axis, Methods), estimated number of lineages in the population (middle, *y* axis, Methods), and estimated inverse Simpson Index, also known as Hill number of order 2 (bottom, *y* axis, Methods) at each time point (*x* axis), computed from all cells with barcodes (left) or subsampled without replacement to match the smallest number of cells per time point, 4,656 cells on day 7 (right). Confidence bands (shaded area) indicate the empirical pointwise 95% coverage confidence interval over 1,000 subsampling repetitions. Since standard species richness estimators are not suited for the analysis of estimated proportions from stratified sampling, we randomly subsampled 8,320, 1,949, and 1,276 day 14 cells without replacement from the cycling and moderate cyclers and non-cycling population, respectively (left panel, for unsorted population proportions see Extended Data Fig. 2a). P-values obtained by (asymptotic) two-sided Welch's *t*-test with bootstrap estimated standard errors, Holm-corrected with level 5% (Methods,  $n = 5,087$ , day 0,  $n = 11,348$ , day 3,  $n = 4,656$ , day 7,  $n = 11,545$ , day 14 subsampled). **c**, Alternative estimates of number of lineages with rarefaction. Rarefaction curves for the expected observed number of different lineages (*y* axis) at varying hypothetical sample sizes (*x* axis) for each time point (coloured lines). Actual number of observed lineages: marker; Interpolated results: solid lines; Extrapolation beyond the observed number: dashed lines. Day 14 cells were subsampled as done for the estimation of the number of lineages in the right-hand side panels of **b**. Shaded areas: confidence bands at 95% confidence level. **d, e**, Estimated cumulative proportion (eCDF) of lineages in the total population (*y* axis) sorted in decreasing order of estimated lineage proportion (*x* axis) for each time point (coloured lines) when estimating the proportion from all sequenced cells with barcodes (**d**) or subsampled to 4,656 cells (**e**) as in **b**. Subsampling (**b, e**) and rarefaction (**c**) facilitate comparison between different time points since estimators of population diversity are strongly biased by sample size. Confidence bands indicate the empirical pointwise 95% coverage confidence interval over 1000 repetitions of the subsampling.

#### Extended Data Fig. 4 Lineage fate analysis.

**a**, Single cell-derived clone size by sample. In each sample, detected barcodes were sorted in descending order by the sum of their counts. Each unique lineage barcode was accounted as a separate clone. **b**, The number of observed multi-fate lineages is significantly smaller ( $P = 1 \times 10^{-5}$ ) than expected by chance. Distribution of the number of multi-fate lineages (*x* axis) in simulated data. Red line: observed number

of multi-fate lineages. **c, d**, Clone size reproducibility is significantly higher than expected by chance. **c**. Clone size on day 7 of each persister lineage barcode inferred from scRNA-seq ( $x, y$  axes) in each of two independent experiments seeded from the same barcoded founding cell population. Top: linear correlation coefficient. **d**, Distribution of  $r^2$  values ( $x$  axis) in simulated day 14 data. Red line: observed  $r^2$  between the two replicates at day 14.

### Extended Data Fig. 5 Differences in transcriptional programs and drug response in cycling and non-cycling persisters.

**a**, EMT signature expression is similar in cycling and non-cycling persisters. Distribution of expression levels of EMT ( $y$  axis,  $\log_2(\text{TPM}+1)$ ) across time points and subpopulations. Effect size (ES): difference between the mean signature score of cycling and non-cycling persisters. **b, c**, Higher expression of glutathione metabolism and NRF2 signatures in cycling vs. non-cycling persisters. Signature score ( $y$  axis) of glutathione metabolism (**b**) and NRF2 pathway (**c**) signatures in cells profiled at each time point ( $x$  axis) stratified by their lineage majority fate at day 14 (colour legend). Effect size (ES) indicates difference between the mean signature score of cycling and non-cycling persisters. **d–f**, Persister populations show differential sensitivity to fulvestrant. Effect of fulvestrant co-treatment (300nM fulvestrant during days 14-20 of 300nM osimertinib) on overall survival (**d**) non-cycling (**e**) and cycling cells (**f**). **g–i**, Persister populations show differential sensitivity to RSL3. Effect of different doses of RSL3 co-treatment (during days 3-14 of 300nM osimertinib) on overall survival (**g**) non-cycling (**h**) and cycling cells (**i**). **j**, Co-treatment with 2  $\mu\text{M}$  of RSL3 shifts surviving persister cells to cycling. Distributions of mCherry fluorescence level ( $x$  axis) for cells analysed by flow cytometry with and without RSL3 co-treatment (panel legend). Error bars are mean  $\pm/-$  s.d. of two (**e, f**) or three biologically independent experiments (**d, h, i**). \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ ; two-tailed  $t$ -tests compared to osimertinib only condition (d-f, h-i)

[Source data.](#)

### Extended Data Fig. 6 Correlation between clone size at day 14 and gene expression.

**a**, Gene expression is progressively more predictive of persister lineages over time. For each time point ( $x$  axis), maximum (blue) and minimum (orange) over the correlation coefficients ( $y$  axis) of each gene and the lineage size at day 14 (also see Fig. 2g). **b**, Choosing expanded and non-expanded lineages for gene expression comparisons. Cut-offs (vertical lines) for highly expanded and non-expanded

lineages on day 14 based on the estimated proportion of each lineage in the population (y axis), sorted by decreasing proportion (x axis, log scale) for each time point (coloured lines). **c**, For each time point (x axis), maximum (blue line) and minimum (orange line) over the correlation coefficients (y axis) of each gene and the lineage size at day 14 as in **a**, but restricted to cells of highly expanded lineages as in **b**. **d**, Genes with top correlation to lineage expansion. Top five rows: distribution of gene expression of top correlated genes (log normalized counts, y axis) at each time point (x axis), comparing cells from non-expanded (red) and expanded (pink) lineages, as defined in **b**. Bottom row: numbers of cells (y axis) per time point in non-expanded, (dark grey) and expanded (light grey) lineages. Distributions are visualized as enhanced box plots indicating median (grey bar) and geometric progression of quantiles (progressively decreasing box widths for 75th, 87.5th, 93.75th, 96.875th, etc. percentiles, and analogously for 25th, 12.5th, 6.25th, 3.125th, etc. percentiles, labelling up to 1.5625% of the data as outliers). Bonferroni-Holm adjusted P values, determined by a two-sided Mann–Whitney U-test with continuity correction, or no significance (NS,  $P > 5\%$ ). **e**, Increase in correlation of top correlated genes as early as day 3. For each time point (x axis), rank of selected genes' (coloured solid lines) correlation with the lineage size at Day 14 among all genes (y axis), normalized to lie between 0 and 1, and average relative correlation rank of genes with similar mean expression as determined by grouping genes by their mean log-normalized expression over all time points combined into 20 bins (coloured dashed lines) (Methods).

### **Extended Data Fig. 7 Metabolite profiles of cycling persisters, non-cycling persisters and untreated parental cells and FAO measurements.**

**a, b**, PCA loadings (x axis) for the top 46 metabolites (y axis) associated with PC1 (**a**) and PC2 (**b**). **c**, UMAP representation of metabolomics data. **d**, Mean FAO level (y axis, relative to mean of the untreated controls) measured by  ${}^3\text{H}$ -palmitic acid oxidation in PC9-Watermelon cells either untreated, treated only with 100  $\mu\text{M}$  etomoxir for 3 days, or treated with 300nM osimertinib for 14 days. **e**, Mean FAO levels (y axis, relative to cells seeded at 300,000 per well, as used for the osimertinib time course) in PC9-Watermelon cells seeded at different densities (x axis) 24 h before measurement. two tailed  $t$ -tests; \*\* $P < 0.01$ ; NS – not significant (compared to 300,000 cells per well). **f**, Mean confluence (y axis) of PC9-Watermelon cells treated with 100 $\mu\text{M}$  Etomoxir for 14 days (Methods). **g**, Mean fraction of cycling persisters for control and sgCPT1A PC9 cells. Error bars are mean  $\pm$  s.d. of two (**f**) or three biologically independent experiments (**d, e, g**). \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ ; NS, not significant ( $P > 0.05$ ); two-tailed  $t$ -tests (**e–g**)

[Source data](#).

## Extended Data Fig. 8 Single cell analysis of multiple Watermelon persister models.

**a, b**, UMAP representation of cells coloured by cell line (**a**) and cluster identity (**b**). **c–e**, Fraction of cells (y axis) in each cluster (x axis) coloured by cell line (**c**), treatment (**d**) and experimental replicate (**e**). **f**, Proportion of cells (y axis) in each cell cycle phase (coloured stacked bars) based on cell cycle scores inferred from scRNA-seq data across cell line models.

## Extended Data Fig. 9 Modelling minimal residual disease using an engineered transgenic mouse model.

**a**, Transgenic composition of the EGFR<sup>L858R</sup> genetically engineered mouse model used in this study. **b**, Tumour burden in a representative osimertinib-treated mouse as measured by MRI. **c**, IHC staining for EGFR L858R mutant and mKate in treatment-naive mouse lung tumours. **d**, Experimental schema for isolation and molecular profiling of mKate<sup>+</sup> cells. **e**, Flow gating scheme for mKate<sup>+</sup> cells. R4 cells from untreated and mice bearing minimal residual disease were sorted and subjected to sequencing. mKate negative cells (R6). Images in **c** are representative of three independent experiments.

## Extended Data Fig. 10 Changes in expression of metabolic programs in patient tumours.

**a, b**, Increase in FAM and ROS pathway signatures in drug-treated human lung adenocarcinoma. Distribution of expression scores of FAM (**a**) and ROS (**b**) signatures in cells from individual EGFR-driven lung adenocarcinoma tumours (with more than 10 cells) across different treatment time points (x axis). Box plots are represented by centre line, median; box limits, upper and lower quartiles; whiskers extend at most  $1.5 \times$  interquartile range past upper and lower quartiles. \\ (\bar{(\rm{mean}}\{x\}))\\: Mean signature level for time point. For number of cells per patient see Supplementary Table 3. **c–e**, Correlation between ROS (y axis) and FAM (x axis) signature scores in (**c**) treatment-naive (TN), residual disease (RD) and progressive disease (PD) human lung adenocarcinoma, (**d**) treatment-naive melanoma, and (**e**) treatment-naive breast cancer. Significance based on bootstrap test (**c**, Methods) and t distribution (**d, e**). 95% confidence interval (**d, e**, shaded area).

## Supplementary information

## **Supplementary Information**

This file contains Supplementary Methods and Supplementary References.

### **Reporting Summary**

#### **Supplementary Table 1**

Mean signatures expression of the three different persister subpopulations.

#### **Supplementary Table 2**

Pooled experiment cell number by model information.

#### **Supplementary Table 3**

Patient information

#### **Supplementary Table 4**

A list of primers used in this study.

#### **Supplementary Table 5**

Gene signature list

## **Source data**

### **Source Data Fig. 1**

### **Source Data Fig. 2**

### **Source Data Fig. 3**

### **Source Data Fig. 4**

### **Source Data Extended Data Fig. 1**

### **Source Data Extended Data Fig. 2**

## Source Data Extended Data Fig. 5

## Source Data Extended Data Fig. 7

## Rights and permissions

### Reprints and Permissions

## About this article

### Cite this article

Oren, Y., Tsabar, M., Cuoco, M.S. *et al.* Cycling cancer persister cells arise from lineages with distinct programs. *Nature* **596**, 576–582 (2021).  
<https://doi.org/10.1038/s41586-021-03796-6>

- Received: 01 June 2020
- Accepted: 02 July 2021
- Published: 11 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03796-6>

## Further reading

- [\*\*A persistent look at how tumours evade therapy\*\*](#)
  - Karen Gomez
  - & Raul Rabadan

*Nature* (2021)

## [\*\*A persistent look at how tumours evade therapy\*\*](#)

- Karen Gomez
- Raul Rabadan

## News & Views 11 Aug 2021

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03796-6>

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

- Article
- Open Access
- [Published: 15 July 2021](#)

# Highly accurate protein structure prediction with AlphaFold

- [John Jumper](#) [ORCID: orcid.org/0000-0001-6169-6580](#)<sup>1</sup> na1,
- [Richard Evans](#)<sup>1</sup> na1,
- [Alexander Pritzel](#)<sup>1</sup> na1,
- [Tim Green](#) [ORCID: orcid.org/0000-0002-3227-1505](#)<sup>1</sup> na1,
- [Michael Figurnov](#)<sup>1</sup> na1,
- [Olaf Ronneberger](#)<sup>1</sup> na1,
- [Kathryn Tunyasuvunakool](#)<sup>1</sup> na1,
- [Russ Bates](#)<sup>1</sup> na1,
- [Augustin Žídek](#)<sup>1</sup> na1,
- [Anna Potapenko](#)<sup>1</sup> na1,
- [Alex Bridgland](#)<sup>1</sup> na1,
- [Clemens Meyer](#)<sup>1</sup> na1,
- [Simon A. A. Kohl](#) [ORCID: orcid.org/0000-0003-4271-4418](#)<sup>1</sup> na1,
- [Andrew J. Ballard](#)<sup>1</sup> na1,
- [Andrew Cowie](#)<sup>1</sup> na1,
- [Bernardino Romera-Paredes](#)<sup>1</sup> na1,
- [Stanislav Nikolov](#)<sup>1</sup> na1,
- [Rishabh Jain](#)<sup>1</sup> na1,
- [Jonas Adler](#) [ORCID: orcid.org/0000-0001-9928-3407](#)<sup>1</sup>,
- [Trevor Back](#)<sup>1</sup>,
- [Stig Petersen](#)<sup>1</sup>,
- [David Reiman](#)<sup>1</sup>,
- [Ellen Clancy](#)<sup>1</sup>,
- [Michał Zielinski](#)<sup>1</sup>,
- [Martin Steinegger](#) [ORCID: orcid.org/0000-0001-8781-9753](#)<sup>2,3</sup>,
- [Michałina Pacholska](#) [ORCID: orcid.org/0000-0002-2160-6226](#)<sup>1</sup>,
- [Tamas Berghammer](#)<sup>1</sup>,
- [Sebastian Bodenstein](#)<sup>1</sup>,

- [David Silver](#) [ORCID: orcid.org/0000-0002-5197-2892<sup>1</sup>](#),
- [Oriol Vinyals<sup>1</sup>](#),
- [Andrew W. Senior](#) [ORCID: orcid.org/0000-0002-2401-5691<sup>1</sup>](#),
- [Koray Kavukcuoglu<sup>1</sup>](#),
- [Pushmeet Kohli<sup>1</sup>](#) &
- [Demis Hassabis](#) [ORCID: orcid.org/0000-0003-2812-9917<sup>1 na1</sup>](#)

[\*Nature\*](#) volume **596**, pages 583–589 (2021)

- 313k Accesses
- 2737 Altmetric
- [Metrics details](#)

## Subjects

- [Computational biophysics](#)
- [Machine learning](#)
- [Protein structure predictions](#)
- [Structural biology](#)

## Abstract

Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort<sup>1,2,3,4</sup>, the structures of around 100,000 unique proteins have been determined<sup>5</sup>, but this represents a small fraction of the billions of known protein sequences<sup>6,7</sup>. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to enable large-scale structural bioinformatics. Predicting the three-dimensional structure that a protein will adopt based solely on its amino acid sequence—the structure prediction component of the ‘protein folding problem’<sup>8</sup>—has been an important open research problem for more than 50 years<sup>9</sup>. Despite recent progress<sup>10,11,12,13,14</sup>, existing methods fall far short of atomic accuracy, especially when no homologous structure is available. Here we provide the first computational method that can regularly predict protein structures with atomic accuracy even in cases in which no similar structure is known. We validated an entirely redesigned version of our neural network-based model, AlphaFold, in the challenging 14th Critical Assessment of protein Structure Prediction (CASP14)<sup>15</sup>, demonstrating accuracy competitive with experimental structures in a majority of

cases and greatly outperforming other methods. Underpinning the latest version of AlphaFold is a novel machine learning approach that incorporates physical and biological knowledge about protein structure, leveraging multi-sequence alignments, into the design of the deep learning algorithm.

[Download PDF](#)

## Main

The development of computational methods to predict three-dimensional (3D) protein structures from the protein sequence has proceeded along two complementary paths that focus on either the physical interactions or the evolutionary history. The physical interaction programme heavily integrates our understanding of molecular driving forces into either thermodynamic or kinetic simulation of protein physics<sup>16</sup> or statistical approximations thereof<sup>17</sup>. Although theoretically very appealing, this approach has proved highly challenging for even moderate-sized proteins due to the computational intractability of molecular simulation, the context dependence of protein stability and the difficulty of producing sufficiently accurate models of protein physics. The evolutionary programme has provided an alternative in recent years, in which the constraints on protein structure are derived from bioinformatics analysis of the evolutionary history of proteins, homology to solved structures<sup>18,19</sup> and pairwise evolutionary correlations<sup>20,21,22,23,24</sup>. This bioinformatics approach has benefited greatly from the steady growth of experimental protein structures deposited in the Protein Data Bank (PDB)<sup>5</sup>, the explosion of genomic sequencing and the rapid development of deep learning techniques to interpret these correlations. Despite these advances, contemporary physical and evolutionary-history-based approaches produce predictions that are far short of experimental accuracy in the majority of cases in which a close homologue has not been solved experimentally and this has limited their utility for many biological applications.

In this study, we develop the first, to our knowledge, computational approach capable of predicting protein structures to near experimental accuracy in a majority of cases. The neural network AlphaFold that we developed was entered into the CASP14 assessment (May–July 2020; entered under the team name ‘AlphaFold2’ and a completely different model from our CASP13 AlphaFold system<sup>10</sup>). The CASP assessment is carried out biennially using recently solved structures that have not been deposited in the PDB or publicly disclosed so that it is a blind test for the participating methods, and has long served as the gold-standard assessment for the accuracy of structure prediction<sup>25,26</sup>.

In CASP14, AlphaFold structures were vastly more accurate than competing methods. AlphaFold structures had a median backbone accuracy of  $0.96 \text{ \AA r.m.s.d.}_{.95}$  ( $\text{Ca}$  root-mean-square deviation at 95% residue coverage) (95% confidence interval = 0.85–1.16 Å) whereas the next best performing method had a median backbone accuracy of  $2.8 \text{ \AA r.m.s.d.}_{.95}$  (95% confidence interval = 2.7–4.0 Å) (measured on CASP domains; see Fig. 1a for backbone accuracy and Supplementary Fig. 14 for all-atom accuracy). As a comparison point for this accuracy, the width of a carbon atom is approximately 1.4 Å. In addition to very accurate domain structures (Fig. 1b), AlphaFold is able to produce highly accurate side chains (Fig. 1c) when the backbone is highly accurate and considerably improves over template-based methods even when strong templates are available. The all-atom accuracy of AlphaFold was  $1.5 \text{ \AA r.m.s.d.}_{.95}$  (95% confidence interval = 1.2–1.6 Å) compared with the  $3.5 \text{ \AA r.m.s.d.}_{.95}$  (95% confidence interval = 3.1–4.2 Å) of the best alternative method. Our methods are scalable to very long proteins with accurate domains and domain-packing (see Fig. 1d for the prediction of a 2,180-residue protein with no structural homologues). Finally, the model is able to provide precise, per-residue estimates of its reliability that should enable the confident use of these predictions.

**Fig. 1: AlphaFold produces highly accurate structures.**

 figure1

- a, The performance of AlphaFold on the CASP14 dataset ( $n = 87$  protein domains) relative to the top-15 entries (out of 146 entries), group numbers correspond to the numbers assigned to entrants by CASP. Data are median and the 95% confidence

interval of the median, estimated from 10,000 bootstrap samples. **b**, Our prediction of CASP14 target T1049 (PDB 6Y4F, blue) compared with the true (experimental) structure (green). Four residues in the C terminus of the crystal structure are *B*-factor outliers and are not depicted. **c**, CASP14 target T1056 (PDB 6YJ1). An example of a well-predicted zinc-binding site (AlphaFold has accurate side chains even though it does not explicitly predict the zinc ion). **d**, CASP target T1044 (PDB 6VR4)—a 2,180-residue single chain—was predicted with correct domain packing (the prediction was made after CASP using AlphaFold without intervention). **e**, Model architecture. Arrows show the information flow among the various components described in this paper. Array shapes are shown in parentheses with *s*, number of sequences ( $N_{\text{seq}}$  in the main text); *r*, number of residues ( $N_{\text{res}}$  in the main text); *c*, number of channels.

We demonstrate in Fig. 2a that the high accuracy that AlphaFold demonstrated in CASP14 extends to a large sample of recently released PDB structures; in this dataset, all structures were deposited in the PDB after our training data cut-off and are analysed as full chains (see [Methods](#), Supplementary Fig. 15 and Supplementary Table 6 for more details). Furthermore, we observe high side-chain accuracy when the backbone prediction is accurate (Fig. 2b) and we show that our confidence measure, the predicted local-distance difference test (pLDDT), reliably predicts the  $\text{C}\alpha$  local-distance difference test (lDDT- $\text{C}\alpha$ ) accuracy of the corresponding prediction (Fig. 2c). We also find that the global superposition metric template modelling score (TM-score)<sup>27</sup> can be accurately estimated (Fig. 2d). Overall, these analyses validate that the high accuracy and reliability of AlphaFold on CASP14 proteins also transfers to an uncurated collection of recent PDB submissions, as would be expected (see [Supplementary Methods 1.15](#) and Supplementary Fig. 11 for confirmation that this high accuracy extends to new folds).

**Fig. 2: Accuracy of AlphaFold on recent PDB structures.**

---

 figure2

The analysed structures are newer than any structure in the training set. Further filtering is applied to reduce redundancy (see [Methods](#)). **a**, Histogram of backbone r.m.s.d. for full chains ( $C\alpha$  r.m.s.d. at 95% coverage). Error bars are 95% confidence intervals (Poisson). This dataset excludes proteins with a template (identified by

hmmsearch) from the training set with more than 40% sequence identity covering more than 1% of the chain ( $n = 3,144$  protein chains). The overall median is 1.46 Å (95% confidence interval = 1.40–1.56 Å). Note that this measure will be highly sensitive to domain packing and domain accuracy; a high r.m.s.d. is expected for some chains with uncertain packing or packing errors. **b**, Correlation between backbone accuracy and side-chain accuracy. Filtered to structures with any observed side chains and resolution better than 2.5 Å ( $n = 5,317$  protein chains); side chains were further filtered to  $B$ -factor <30 Å<sup>2</sup>. A rotamer is classified as correct if the predicted torsion angle is within 40°. Each point aggregates a range of IDDT-Cα, with a bin size of 2 units above 70 IDDT-Cα and 5 units otherwise. Points correspond to the mean accuracy; error bars are 95% confidence intervals (Student *t*-test) of the mean on a per-residue basis. **c**, Confidence score compared to the true accuracy on chains. Least-squares linear fit IDDT-Cα = 0.997 × pLDDT – 1.17 (Pearson's *r* = 0.76).  $n = 10,795$  protein chains. The shaded region of the linear fit represents a 95% confidence interval estimated from 10,000 bootstrap samples. In the companion paper<sup>39</sup>, additional quantification of the reliability of pLDDT as a confidence measure is provided. **d**, Correlation between pTM and full chain TM-score. Least-squares linear fit TM-score = 0.98 × pTM + 0.07 (Pearson's *r* = 0.85).  $n = 10,795$  protein chains. The shaded region of the linear fit represents a 95% confidence interval estimated from 10,000 bootstrap samples.

## The AlphaFold network

AlphaFold greatly improves the accuracy of structure prediction by incorporating novel neural network architectures and training procedures based on the evolutionary, physical and geometric constraints of protein structures. In particular, we demonstrate a new architecture to jointly embed multiple sequence alignments (MSAs) and pairwise features, a new output representation and associated loss that enable accurate end-to-end structure prediction, a new equivariant attention architecture, use of intermediate losses to achieve iterative refinement of predictions, masked MSA loss to jointly train with the structure, learning from unlabelled protein sequences using self-distillation and self-estimates of accuracy.

The AlphaFold network directly predicts the 3D coordinates of all heavy atoms for a given protein using the primary amino acid sequence and aligned sequences of homologues as inputs (Fig. 1e; see [Methods](#) for details of inputs including databases, MSA construction and use of templates). A description of the most important ideas and components is provided below. The full network architecture and training procedure are provided in the [Supplementary Methods](#).

The network comprises two main stages. First, the trunk of the network processes the inputs through repeated layers of a novel neural network block that we term

Evoformer to produce an  $N_{\text{seq}} \times N_{\text{res}}$  array ( $N_{\text{seq}}$ , number of sequences;  $N_{\text{res}}$ , number of residues) that represents a processed MSA and an  $N_{\text{res}} \times N_{\text{res}}$  array that represents residue pairs. The MSA representation is initialized with the raw MSA (although see [Supplementary Methods 1.2.7](#) for details of handling very deep MSAs). The Evoformer blocks contain a number of attention-based and non-attention-based components. We show evidence in ‘Interpreting the neural network’ that a concrete structural hypothesis arises early within the Evoformer blocks and is continuously refined. The key innovations in the Evoformer block are new mechanisms to exchange information within the MSA and pair representations that enable direct reasoning about the spatial and evolutionary relationships.

The trunk of the network is followed by the structure module that introduces an explicit 3D structure in the form of a rotation and translation for each residue of the protein (global rigid body frames). These representations are initialized in a trivial state with all rotations set to the identity and all positions set to the origin, but rapidly develop and refine a highly accurate protein structure with precise atomic details. Key innovations in this section of the network include breaking the chain structure to allow simultaneous local refinement of all parts of the structure, a novel equivariant transformer to allow the network to implicitly reason about the unrepresented side-chain atoms and a loss term that places substantial weight on the orientational correctness of the residues. Both within the structure module and throughout the whole network, we reinforce the notion of iterative refinement by repeatedly applying the final loss to outputs and then feeding the outputs recursively into the same modules. The iterative refinement using the whole network (which we term ‘recycling’ and is related to approaches in computer vision<sup>28,29</sup>) contributes markedly to accuracy with minor extra training time (see [Supplementary Methods 1.8](#) for details).

## Evoformer

The key principle of the building block of the network—named Evoformer (Figs. 1e, [3a](#))—is to view the prediction of protein structures as a graph inference problem in 3D space in which the edges of the graph are defined by residues in proximity. The elements of the pair representation encode information about the relation between the residues (Fig. [3b](#)). The columns of the MSA representation encode the individual residues of the input sequence while the rows represent the sequences in which those residues appear. Within this framework, we define a number of update operations that are applied in each block in which the different update operations are applied in series.

**Fig. 3: Architectural details.**

---

 **figure3**

---

**a**, Evoformer block. Arrows show the information flow. The shape of the arrays is shown in parentheses. **b**, The pair representation interpreted as directed edges in a graph. **c**, Triangle multiplicative update and triangle self-attention. The circles represent residues. Entries in the pair representation are illustrated as directed edges and in each diagram, the edge being updated is  $ij$ . **d**, Structure module including Invariant point attention (IPA) module. The single representation is a copy of the first row of the MSA representation. **e**, Residue gas: a representation of each residue as one free-floating rigid body for the backbone (blue triangles) and  $\chi$  angles for the side chains (green circles). The corresponding atomic structure is shown below. **f**, Frame aligned point error (FAPE). Green, predicted structure; grey, true structure;  $(R_k, t_k)$ , frames;  $x_i$ , atom positions.

The MSA representation updates the pair representation through an element-wise outer product that is summed over the MSA sequence dimension. In contrast to previous work<sup>30</sup>, this operation is applied within every block rather than once in the network, which enables the continuous communication from the evolving MSA representation to the pair representation.

Within the pair representation, there are two different update patterns. Both are inspired by the necessity of consistency of the pair representation—for a pairwise description of amino acids to be representable as a single 3D structure, many constraints must be satisfied including the triangle inequality on distances. On the basis of this intuition, we arrange the update operations on the pair representation in terms of triangles of edges involving three different nodes (Fig. 3c). In particular, we add an extra logit bias to axial attention<sup>31</sup> to include the ‘missing edge’ of the triangle and we define a non-attention update operation ‘triangle multiplicative update’ that uses two edges to update the missing third edge (see [Supplementary Methods 1.6.5](#) for details). The triangle multiplicative update was developed originally as a more symmetric and cheaper replacement for the attention, and networks that use only the attention or multiplicative update are both able to produce high-accuracy structures. However, the combination of the two updates is more accurate.

We also use a variant of axial attention within the MSA representation. During the per-sequence attention in the MSA, we project additional logits from the pair stack to bias the MSA attention. This closes the loop by providing information flow from the pair representation back into the MSA representation, ensuring that the overall Evoformer block is able to fully mix information between the pair and MSA representations and prepare for structure generation within the structure module.

## End-to-end structure prediction

The structure module (Fig. 3d) operates on a concrete 3D backbone structure using the pair representation and the original sequence row (single representation) of the MSA representation from the trunk. The 3D backbone structure is represented as  $N_{\text{res}}$  independent rotations and translations, each with respect to the global frame (residue gas) (Fig. 3e). These rotations and translations—representing the geometry of the N-Cα-C atoms—prioritize the orientation of the protein backbone so that the location of the side chain of each residue is highly constrained within that frame. Conversely, the peptide bond geometry is completely unconstrained and the network is observed to frequently violate the chain constraint during the application of the structure module as breaking this constraint enables the local refinement of all parts of the chain without solving complex loop closure problems. Satisfaction of the peptide bond geometry is encouraged during fine-tuning by a violation loss term. Exact enforcement of peptide bond geometry is only achieved in the post-prediction relaxation of the structure by gradient descent in the Amber<sup>32</sup> force field. Empirically, this final relaxation does not improve the accuracy of the model as measured by the global distance test (GDT)<sup>33</sup> or IDDT-Cα<sup>34</sup> but does remove distracting stereochemical violations without the loss of accuracy.

The residue gas representation is updated iteratively in two stages (Fig. 3d). First, a geometry-aware attention operation that we term ‘invariant point attention’ (IPA) is used to update an  $N_{\text{res}}$  set of neural activations (single representation) without changing the 3D positions, then an equivariant update operation is performed on the residue gas using the updated activations. The IPA augments each of the usual attention queries, keys and values with 3D points that are produced in the local frame of each residue such that the final value is invariant to global rotations and translations (see [Methods](#) ‘IPA’ for details). The 3D queries and keys also impose a strong spatial/locality bias on the attention, which is well-suited to the iterative refinement of the protein structure. After each attention operation and element-wise transition block, the module computes an update to the rotation and translation of each backbone frame. The application of these updates within the local frame of each residue makes the overall attention and update block an equivariant operation on the residue gas.

Predictions of side-chain  $\chi$  angles as well as the final, per-residue accuracy of the structure (pLDDT) are computed with small per-residue networks on the final activations at the end of the network. The estimate of the TM-score (pTM) is obtained from a pairwise error prediction that is computed as a linear projection from the final pair representation. The final loss (which we term the frame-aligned point error (FAPE) (Fig. 3f)) compares the predicted atom positions to the true positions under many different alignments. For each alignment, defined by aligning the predicted frame ( $R_k, \mathbf{t}_k$ ) to the corresponding true frame, we compute the distance of all predicted atom positions  $\mathbf{x}_i$  from the true atom positions. The resulting  $N_{\text{frames}} \times N_{\text{atoms}}$  distances are penalized with a clamped  $L^1$  loss. This creates a strong bias for atoms to be correct relative to the local frame of each residue and hence correct with respect to its side-chain interactions, as well as providing the main source of chirality for AlphaFold ([Supplementary Methods 1.9.3](#) and [Supplementary Fig. 9](#)).

## Training with labelled and unlabelled data

The AlphaFold architecture is able to train to high accuracy using only supervised learning on PDB data, but we are able to enhance accuracy (Fig. 4a) using an approach similar to noisy student self-distillation<sup>35</sup>. In this procedure, we use a trained network to predict the structure of around 350,000 diverse sequences from Uniclust30<sup>36</sup> and make a new dataset of predicted structures filtered to a high-confidence subset. We then train the same architecture again from scratch using a mixture of PDB data and this new dataset of predicted structures as the training data, in which the various training data augmentations such as cropping and MSA subsampling make it challenging for the network to recapitulate the previously predicted structures. This self-distillation procedure makes effective use of the unlabelled sequence data and considerably improves the accuracy of the resulting network.

**Fig. 4: Interpreting the neural network.**

 figure4

**a**, Ablation results on two target sets: the CASP14 set of domains ( $n = 87$  protein domains) and the PDB test set of chains with template coverage of  $\leq 30\%$  at 30% identity ( $n = 2,261$  protein chains). Domains are scored with GDT and chains are

scored with IDDT-C $\alpha$ . The ablations are reported as a difference compared with the average of the three baseline seeds. Means (points) and 95% bootstrap percentile intervals (error bars) are computed using bootstrap estimates of 10,000 samples. **b**, Domain GDT trajectory over 4 recycling iterations and 48 Evoformer blocks on CASP14 targets LmrP (T1024) and Orf8 (T1064) where D1 and D2 refer to the individual domains as defined by the CASP assessment. Both T1024 domains obtain the correct structure early in the network, whereas the structure of T1064 changes multiple times and requires nearly the full depth of the network to reach the final structure. Note, 48 Evoformer blocks comprise one recycling iteration.

Additionally, we randomly mask out or mutate individual residues within the MSA and have a Bidirectional Encoder Representations from Transformers (BERT)-style<sup>37</sup> objective to predict the masked elements of the MSA sequences. This objective encourages the network to learn to interpret phylogenetic and covariation relationships without hardcoding a particular correlation statistic into the features. The BERT objective is trained jointly with the normal PDB structure loss on the same training examples and is not pre-trained, in contrast to recent independent work<sup>38</sup>.

## Interpreting the neural network

To understand how AlphaFold predicts protein structure, we trained a separate structure module for each of the 48 Evoformer blocks in the network while keeping all parameters of the main network frozen ([Supplementary Methods 1.14](#)). Including our recycling stages, this provides a trajectory of 192 intermediate structures—one per full Evoformer block—in which each intermediate represents the belief of the network of the most likely structure at that block. The resulting trajectories are surprisingly smooth after the first few blocks, showing that AlphaFold makes constant incremental improvements to the structure until it can no longer improve (see Fig. [4b](#) for a trajectory of accuracy). These trajectories also illustrate the role of network depth. For very challenging proteins such as ORF8 of SARS-CoV-2 (T1064), the network searches and rearranges secondary structure elements for many layers before settling on a good structure. For other proteins such as LmrP (T1024), the network finds the final structure within the first few layers. Structure trajectories of CASP14 targets T1024, T1044, T1064 and T1091 that demonstrate a clear iterative building process for a range of protein sizes and difficulties are shown in Supplementary Videos [1–4](#). In [Supplementary Methods 1.16](#) and Supplementary Figs. [12, 13](#), we interpret the attention maps produced by AlphaFold layers.

Figure [4a](#) contains detailed ablations of the components of AlphaFold that demonstrate that a variety of different mechanisms contribute to AlphaFold accuracy. Detailed descriptions of each ablation model, their training details, extended discussion of

ablation results and the effect of MSA depth on each ablation are provided in [Supplementary Methods 1.13](#) and [Supplementary Fig. 10](#).

## MSA depth and cross-chain contacts

Although AlphaFold has a high accuracy across the vast majority of deposited PDB structures, we note that there are still factors that affect accuracy or limit the applicability of the model. The model uses MSAs and the accuracy decreases substantially when the median alignment depth is less than around 30 sequences (see Fig. [5a](#) for details). We observe a threshold effect where improvements in MSA depth over around 100 sequences lead to small gains. We hypothesize that the MSA information is needed to coarsely find the correct structure within the early stages of the network, but refinement of that prediction into a high-accuracy model does not depend crucially on the MSA information. The other substantial limitation that we have observed is that AlphaFold is much weaker for proteins that have few intra-chain or homotypic contacts compared to the number of heterotypic contacts (further details are provided in a companion paper<sup>[39](#)</sup>). This typically occurs for bridging domains within larger complexes in which the shape of the protein is created almost entirely by interactions with other chains in the complex. Conversely, AlphaFold is often able to give high-accuracy predictions for homomers, even when the chains are substantially intertwined (Fig. [5b](#)). We expect that the ideas of AlphaFold are readily applicable to predicting full hetero-complexes in a future system and that this will remove the difficulty with protein chains that have a large number of hetero-contacts.

**Fig. 5: Effect of MSA depth and cross-chain contacts.**



- a.** Backbone accuracy (IDDT- $\text{Ca}$ ) for the redundancy-reduced set of the PDB after our training data cut-off, restricting to proteins in which at most 25% of the long-range contacts are between different heteromer chains. We further consider two groups of proteins based on template coverage at 30% sequence identity: covering more than 60% of the chain ( $n = 6,743$  protein chains) and covering less than 30% of the chain

( $n = 1,596$  protein chains). MSA depth is computed by counting the number of non-gap residues for each position in the MSA (using the  $N_{\text{eff}}$  weighting scheme; see [Methods](#) for details) and taking the median across residues. The curves are obtained through Gaussian kernel average smoothing (window size is 0.2 units in  $\log_{10}(N_{\text{eff}})$ ); the shaded area is the 95% confidence interval estimated using bootstrap of 10,000 samples. **b**, An intertwined homotrimer (PDB 6SK0) is correctly predicted without input stoichiometry and only a weak template (blue is predicted and green is experimental).

## Related work

The prediction of protein structures has had a long and varied development, which is extensively covered in a number of reviews [14,40,41,42,43](#). Despite the long history of applying neural networks to structure prediction [14,42,43](#), they have only recently come to improve structure prediction [10,11,44,45](#). These approaches effectively leverage the rapid improvement in computer vision systems [46](#) by treating the problem of protein structure prediction as converting an ‘image’ of evolutionary couplings [22,23,24](#) to an ‘image’ of the protein distance matrix and then integrating the distance predictions into a heuristic system that produces the final 3D coordinate prediction. A few recent studies have been developed to predict the 3D coordinates directly [47,48,49,50](#), but the accuracy of these approaches does not match traditional, hand-crafted structure prediction pipelines [51](#). In parallel, the success of attention-based networks for language processing [52](#) and, more recently, computer vision [31,53](#) has inspired the exploration of attention-based methods for interpreting protein sequences [54,55,56](#).

## Discussion

The methodology that we have taken in designing AlphaFold is a combination of the bioinformatics and physical approaches: we use a physical and geometric inductive bias to build components that learn from PDB data with minimal imposition of handcrafted features (for example, AlphaFold builds hydrogen bonds effectively without a hydrogen bond score function). This results in a network that learns far more efficiently from the limited data in the PDB but is able to cope with the complexity and variety of structural data.

In particular, AlphaFold is able to handle missing the physical context and produce accurate models in challenging cases such as intertwined homomers or proteins that only fold in the presence of an unknown haem group. The ability to handle underspecified structural conditions is essential to learning from PDB structures as the PDB represents the full range of conditions in which structures have been solved. In

general, AlphaFold is trained to produce the protein structure most likely to appear as part of a PDB structure. For example, in cases in which a particular stoichiometry, ligand or ion is predictable from the sequence alone, AlphaFold is likely to produce a structure that respects those constraints implicitly.

AlphaFold has already demonstrated its utility to the experimental community, both for molecular replacement<sup>57</sup> and for interpreting cryogenic electron microscopy maps<sup>58</sup>. Moreover, because AlphaFold outputs protein coordinates directly, AlphaFold produces predictions in graphics processing unit (GPU) minutes to GPU hours depending on the length of the protein sequence (for example, around one GPU minute per model for 384 residues; see [Methods](#) for details). This opens up the exciting possibility of predicting structures at the proteome-scale and beyond—in a companion paper<sup>39</sup>, we demonstrate the application of AlphaFold to the entire human proteome<sup>39</sup>.

The explosion in available genomic sequencing techniques and data has revolutionized bioinformatics but the intrinsic challenge of experimental structure determination has prevented a similar expansion in our structural knowledge. By developing an accurate protein structure prediction algorithm, coupled with existing large and well-curated structure and sequence databases assembled by the experimental community, we hope to accelerate the advancement of structural bioinformatics that can keep pace with the genomics revolution. We hope that AlphaFold—and computational approaches that apply its techniques for other biophysical problems—will become essential tools of modern biology.

## Methods

### Full algorithm details

Extensive explanations of the components and their motivations are available in [Supplementary Methods 1.1–1.10](#), in addition, pseudocode is available in [Supplementary Information Algorithms 1–32](#), network diagrams in Supplementary Figs. [1–8](#), input features in Supplementary Table [1](#) and additional details are provided in Supplementary Tables [2](#), [3](#). Training and inference details are provided in [Supplementary Methods 1.11–1.12](#) and Supplementary Tables [4](#), [5](#).

### IPA

The IPA module combines the pair representation, the single representation and the geometric representation to update the single representation (Supplementary Fig. [8](#)). Each of these representations contributes affinities to the shared attention weights and then uses these weights to map its values to the output. The IPA operates in 3D space.

Each residue produces query points, key points and value points in its local frame. These points are projected into the global frame using the backbone frame of the residue in which they interact with each other. The resulting points are then projected back into the local frame. The affinity computation in the 3D space uses squared distances and the coordinate transformations ensure the invariance of this module with respect to the global frame (see [Supplementary Methods 1.8.2](#) ‘Invariant point attention (IPA)’ for the algorithm, proof of invariance and a description of the full multi-head version). A related construction that uses classic geometric invariants to construct pairwise features in place of the learned 3D points has been applied to protein design<sup>59</sup>.

In addition to the IPA, standard dot product attention is computed on the abstract single representation and a special attention on the pair representation. The pair representation augments both the logits and the values of the attention process, which is the primary way in which the pair representation controls the structure generation.

## Inputs and data sources

Inputs to the network are the primary sequence, sequences from evolutionarily related proteins in the form of a MSA created by standard tools including jackhmmer<sup>60</sup> and HHBlits<sup>61</sup>, and 3D atom coordinates of a small number of homologous structures (templates) where available. For both the MSA and templates, the search processes are tuned for high recall; spurious matches will probably appear in the raw MSA but this matches the training condition of the network.

One of the sequence databases used, Big Fantastic Database (BFD), was custom-made and released publicly (see ‘Data availability’) and was used by several CASP teams. BFD is one of the largest publicly available collections of protein families. It consists of 65,983,866 families represented as MSAs and hidden Markov models (HMMs) covering 2,204,359,010 protein sequences from reference databases, metagenomes and metatranscriptomes.

BFD was built in three steps. First, 2,423,213,294 protein sequences were collected from UniProt (Swiss-Prot&TrEMBL, 2017-11)<sup>62</sup>, a soil reference protein catalogue and the marine eukaryotic reference catalogue<sup>7</sup>, and clustered to 30% sequence identity, while enforcing a 90% alignment coverage of the shorter sequences using MMseqs2/Linclus<sup>63</sup>. This resulted in 345,159,030 clusters. For computational efficiency, we removed all clusters with less than three members, resulting in 61,083,719 clusters. Second, we added 166,510,624 representative protein sequences from Metaclust NR (2017-05; discarding all sequences shorter than 150 residues)<sup>63</sup> by aligning them against the cluster representatives using MMseqs2<sup>64</sup>. Sequences that fulfilled the sequence identity and coverage criteria were assigned to the best scoring

cluster. The remaining 25,347,429 sequences that could not be assigned were clustered separately and added as new clusters, resulting in the final clustering. Third, for each of the clusters, we computed an MSA using FAMSA<sup>65</sup> and computed the HMMs following the Uniclust HH-suite database protocol<sup>36</sup>.

The following versions of public datasets were used in this study. Our models were trained on a copy of the PDB<sup>5</sup> downloaded on 28 August 2019. For finding template structures at prediction time, we used a copy of the PDB downloaded on 14 May 2020, and the PDB70<sup>66</sup> clustering database downloaded on 13 May 2020. For MSA lookup at both training and prediction time, we used Uniref90<sup>67</sup> v.2020\_01, BFD, Uniclust30<sup>36</sup> v.2018\_08 and MGnify<sup>6</sup> v.2018\_12. For sequence distillation, we used Uniclust30<sup>36</sup> v.2018\_08 to construct a distillation structure dataset. Full details are provided in [Supplementary Methods 1.2](#).

For MSA search on BFD + Uniclust30, and template search against PDB70, we used HHBlits<sup>61</sup> and HHSearch<sup>66</sup> from hh-suite v.3.0-beta.3 (version 14/07/2017). For MSA search on Uniref90 and clustered MGnify, we used jackhmmer from HMMER3<sup>68</sup>. For constrained relaxation of structures, we used OpenMM v.7.3.1<sup>69</sup> with the Amber99sb force field<sup>32</sup>. For neural network construction, running and other analyses, we used TensorFlow<sup>70</sup>, Sonnet<sup>71</sup>, NumPy<sup>72</sup>, Python<sup>73</sup> and Colab<sup>74</sup>.

To quantify the effect of the different sequence data sources, we re-ran the CASP14 proteins using the same models but varying how the MSA was constructed. Removing BFD reduced the mean accuracy by 0.4 GDT, removing MGnify reduced the mean accuracy by 0.7 GDT, and removing both reduced the mean accuracy by 6.1 GDT. In each case, we found that most targets had very small changes in accuracy but a few outliers had very large (20+ GDT) differences. This is consistent with the results in Fig. 5a in which the depth of the MSA is relatively unimportant until it approaches a threshold value of around 30 sequences when the MSA size effects become quite large. We observe mostly overlapping effects between inclusion of BFD and MGnify, but having at least one of these metagenomics databases is very important for target classes that are poorly represented in UniRef, and having both was necessary to achieve full CASP accuracy.

## Training regimen

To train, we use structures from the PDB with a maximum release date of 30 April 2018. Chains are sampled in inverse proportion to cluster size of a 40% sequence identity clustering. We then randomly crop them to 256 residues and assemble into batches of size 128. We train the model on Tensor Processing Unit (TPU) v3 with a batch size of 1 per TPU core, hence the model uses 128 TPU v3 cores. The model is

trained until convergence (around 10 million samples) and further fine-tuned using longer crops of 384 residues, larger MSA stack and reduced learning rate (see [Supplementary Methods 1.11](#) for the exact configuration). The initial training stage takes approximately 1 week, and the fine-tuning stage takes approximately 4 additional days.

The network is supervised by the FAPE loss and a number of auxiliary losses. First, the final pair representation is linearly projected to a binned distance distribution (distogram) prediction, scored with a cross-entropy loss. Second, we use random masking on the input MSAs and require the network to reconstruct the masked regions from the output MSA representation using a BERT-like loss<sup>37</sup>. Third, the output single representations of the structure module are used to predict binned per-residue IDDT-C $\alpha$  values. Finally, we use an auxiliary side-chain loss during training, and an auxiliary structure violation loss during fine-tuning. Detailed descriptions and weighting are provided in the [Supplementary Information](#).

An initial model trained with the above objectives was used to make structure predictions for a Uniclust dataset of 355,993 sequences with the full MSAs. These predictions were then used to train a final model with identical hyperparameters, except for sampling examples 75% of the time from the Uniclust prediction set, with sub-sampled MSAs, and 25% of the time from the clustered PDB set.

We train five different models using different random seeds, some with templates and some without, to encourage diversity in the predictions (see Supplementary Table 5 and [Supplementary Methods 1.12.1](#) for details). We also fine-tuned these models after CASP14 to add a pTM prediction objective ([Supplementary Methods 1.9.7](#)) and use the obtained models for Fig. 2d.

## Inference regimen

We inference the five trained models and use the predicted confidence score to select the best model per target.

Using our CASP14 configuration for AlphaFold, the trunk of the network is run multiple times with different random choices for the MSA cluster centres (see [Supplementary Methods 1.11.2](#) for details of the ensembling procedure). The full time to make a structure prediction varies considerably depending on the length of the protein. Representative timings for the neural network using a single model on V100 GPU are 4.8 min with 256 residues, 9.2 min with 384 residues and 18 h at 2,500 residues. These timings are measured using our open-source code, and the open-source code is notably faster than the version we ran in CASP14 as we now use the XLA compiler<sup>75</sup>.

Since CASP14, we have found that the accuracy of the network without ensembling is very close or equal to the accuracy with ensembling and we turn off ensembling for most inference. Without ensembling, the network is 8 $\times$  faster and the representative timings for a single model are 0.6 min with 256 residues, 1.1 min with 384 residues and 2.1 h with 2,500 residues.

Inferencing large proteins can easily exceed the memory of a single GPU. For a V100 with 16 GB of memory, we can predict the structure of proteins up to around 1,300 residues without ensembling and the 256- and 384-residue inference times are using the memory of a single GPU. The memory usage is approximately quadratic in the number of residues, so a 2,500-residue protein involves using unified memory so that we can greatly exceed the memory of a single V100. In our cloud setup, a single V100 is used for computation on a 2,500-residue protein but we requested four GPUs to have sufficient memory.

Searching genetic sequence databases to prepare inputs and final relaxation of the structures take additional central processing unit (CPU) time but do not require a GPU or TPU.

## Metrics

The predicted structure is compared to the true structure from the PDB in terms of IDDT metric<sup>34</sup>, as this metric reports the domain accuracy without requiring a domain segmentation of chain structures. The distances are either computed between all heavy atoms (IDDT) or only the C $\alpha$  atoms to measure the backbone accuracy (IDDT-C $\alpha$ ). As IDDT-C $\alpha$  only focuses on the C $\alpha$  atoms, it does not include the penalty for structural violations and clashes. Domain accuracies in CASP are reported as GDT<sup>33</sup> and the TM-score<sup>27</sup> is used as a full chain global superposition metric.

We also report accuracies using the r.m.s.d.<sub>95</sub> (C $\alpha$  r.m.s.d. at 95% coverage). We perform five iterations of (1) a least-squares alignment of the predicted structure and the PDB structure on the currently chosen C $\alpha$  atoms (using all C $\alpha$  atoms in the first iteration); (2) selecting the 95% of C $\alpha$  atoms with the lowest alignment error. The r.m.s.d. of the atoms chosen for the final iterations is the r.m.s.d.<sub>95</sub>. This metric is more robust to apparent errors that can originate from crystal structure artefacts, although in some cases the removed 5% of residues will contain genuine modelling errors.

## Test set of recent PDB sequences

For evaluation on recent PDB sequences (Figs. 2a–d, 4a, 5a), we used a copy of the PDB downloaded 15 February 2021. Structures were filtered to those with a release date after 30 April 2018 (the date limit for inclusion in the training set for AlphaFold).

Chains were further filtered to remove sequences that consisted of a single amino acid as well as sequences with an ambiguous chemical component at any residue position. Exact duplicates were removed, with the chain with the most resolved C $\alpha$  atoms used as the representative sequence. Subsequently, structures with less than 16 resolved residues, with unknown residues or solved by NMR methods were removed. As the PDB contains many near-duplicate sequences, the chain with the highest resolution was selected from each cluster in the PDB 40% sequence clustering of the data. Furthermore, we removed all sequences for which fewer than 80 amino acids had the alpha carbon resolved and removed chains with more than 1,400 residues. The final dataset contained 10,795 protein sequences.

The procedure for filtering the recent PDB dataset based on prior template identity was as follows. Hmmsearch was run with default parameters against a copy of the PDB SEQRES fasta downloaded 15 February 2021. Template hits were accepted if the associated structure had a release date earlier than 30 April 2018. Each residue position in a query sequence was assigned the maximum identity of any template hit covering that position. Filtering then proceeded as described in the individual figure legends, based on a combination of maximum identity and sequence coverage.

The MSA depth analysis was based on computing the normalized number of effective sequences ( $N_{\text{eff}}$ ) for each position of a query sequence. Per-residue  $N_{\text{eff}}$  values were obtained by counting the number of non-gap residues in the MSA for this position and weighting the sequences using the  $N_{\text{eff}}$  scheme<sup>76</sup> with a threshold of 80% sequence identity measured on the region that is non-gap in either sequence.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

All input data are freely available from public sources.

Structures from the PDB were used for training and as templates (<https://www.wwpdb.org/ftp/pdb-ftp-sites>; for the associated sequence data and 40% sequence clustering see also [https://ftp.wwpdb.org/pub/pdb/derived\\_data/](https://ftp.wwpdb.org/pub/pdb/derived_data/) and <https://cdn.rcsb.org/resources/sequence/clusters/bc-40.out>). Training used a version of the PDB downloaded 28 August 2019, while the CASP14 template search used a version downloaded 14 May 2020. The template search also used the PDB70 database, downloaded 13 May 2020 ([https://wwwuser.gwdg.de/~combiol/data/hhsuite/databases/hhsuite\\_dbs/](https://wwwuser.gwdg.de/~combiol/data/hhsuite/databases/hhsuite_dbs/)).

We show experimental structures from the PDB with accession numbers [6Y4F<sup>77</sup>](#), [6YJ1<sup>78</sup>](#), [6VR4<sup>79</sup>](#), [6SK0<sup>80</sup>](#), [6FES<sup>81</sup>](#), [6W6W<sup>82</sup>](#), [6T1Z<sup>83</sup>](#) and [7JTL<sup>84</sup>](#).

For MSA lookup at both the training and prediction time, we used UniRef90 v.2020\_01 ([https://ftp.ebi.ac.uk/pub/databases/uniprot/previous\\_releases/release-2020\\_01/uniref/](https://ftp.ebi.ac.uk/pub/databases/uniprot/previous_releases/release-2020_01/uniref/)), BFD (<https://bfd.mmseqs.com>), Uniclust30 v.2018\_08 ([https://wwwuser.gwdg.de/~comppbiol/uniclust/2018\\_08/](https://wwwuser.gwdg.de/~comppbiol/uniclust/2018_08/)) and MGnify clusters v.2018\_12 ([https://ftp.ebi.ac.uk/pub/databases/metagenomics/peptide\\_database/2018\\_12/](https://ftp.ebi.ac.uk/pub/databases/metagenomics/peptide_database/2018_12/)). Uniclust30 v.2018\_08 was also used as input for constructing a distillation structure dataset.

## Code availability

Source code for the AlphaFold model, trained weights and inference script are available under an open-source license at <https://github.com/deepmind/alphafold>.

Neural networks were developed with TensorFlow v.1 (<https://github.com/tensorflow/tensorflow>), Sonnet v.1 (<https://github.com/deepmind/sonnet>), JAX v.0.1.69 (<https://github.com/google/jax>) and Haiku v.0.0.4 (<https://github.com/deepmind/dm-haiku>). The XLA compiler is bundled with JAX and does not have a separate version number.

For MSA search on BFD+Uniclust30, and for template search against PDB70, we used HHBlits and HHSearch from hh-suite v.3.0-beta.3 release 14/07/2017 (<https://github.com/soedinglab/hh-suite>). For MSA search on UniRef90 and clustered MGnify, we used jackhmmer from HMMER v.3.3 (<http://eddylab.org/software/hmmer/>). For constrained relaxation of structures, we used OpenMM v.7.3.1 (<https://github.com/openmm/openmm>) with the Amber99sb force field.

Construction of BFD used MMseqs2 v.925AF (<https://github.com/soedinglab/MMseqs2>) and FAMSA v.1.2.5 (<https://github.com/refresh-bio/FAMSA>).

Data analysis used Python v.3.6 (<https://www.python.org/>), NumPy v.1.16.4 (<https://github.com/numpy/numpy>), SciPy v.1.2.1 (<https://www.scipy.org/>), seaborn v.0.11.1 (<https://github.com/mwaskom/seaborn>), Matplotlib v.3.3.4 (<https://github.com/matplotlib/matplotlib>), bokeh v.1.4.0 (<https://github.com/bokeh/bokeh>), pandas v.1.1.5 (<https://github.com/pandas-dev/pandas>), plotnine v.0.8.0 (<https://github.com/has2k1/plotnine>), statsmodels v.0.12.2 (<https://github.com/statsmodels/statsmodels>) and Colab

(<https://research.google.com/colaboratory>). TM-align v.20190822 (<https://zhanglab.dcmb.med.umich.edu/TM-align/>) was used for computing TM-scores. Structure visualizations were created in Pymol v.2.3.0 (<https://github.com/schrodinger/pymol-open-source>).

## References

1. 1.

Thompson, M. C., Yeates, T. O. & Rodriguez, J. A. Advances in methods for atomic resolution macromolecular structure determination. *F1000Res.* **9**, 667 (2020).

2. 2.

Bai, X.-C., McMullan, G. & Scheres, S. H. W. How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.* **40**, 49–57 (2015).

3. 3.

Jaskolski, M., Dauter, Z. & Wlodawer, A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS J.* **281**, 3985–4009 (2014).

4. 4.

Wüthrich, K. The way to NMR structures of proteins. *Nat. Struct. Biol.* **8**, 923–925 (2001).

5. 5.

wwPDB Consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **47**, D520–D528 (2018).

6. 6.

Mitchell, A. L. et al. MGnify: the microbiome analysis resource in 2020. *Nucleic Acids Res.* **48**, D570–D578 (2020).

7. 7.

Steinegger, M., Mirdita, M. & Söding, J. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nat. Methods* **16**, 603–

606 (2019).

8. 8.

Dill, K. A., Ozkan, S. B., Shell, M. S. & Weikl, T. R. The protein folding problem. *Annu. Rev. Biophys.* **37**, 289–316 (2008).

9. 9.

Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).

10. 10.

Senior, A. W. et al. Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020).

11. 11.

Wang, S., Sun, S., Li, Z., Zhang, R. & Xu, J. Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLOS Comput. Biol.* **13**, e1005324 (2017).

12. 12.

Zheng, W. et al. Deep-learning contact-map guided protein structure prediction in CASP13. *Proteins* **87**, 1149–1164 (2019).

13. 13.

Abriata, L. A., Tamò, G. E. & Dal Peraro, M. A further leap of improvement in tertiary structure prediction in CASP13 prompts new routes for future assessments. *Proteins* **87**, 1100–1112 (2019).

14. 14.

Pearce, R. & Zhang, Y. Deep learning techniques have significantly impacted protein structure prediction and protein design. *Curr. Opin. Struct. Biol.* **68**, 194–207 (2021).

15. 15.

Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T. & Topf, M. Critical assessment of techniques for protein structure prediction, fourteenth round. *CASP*

*14 Abstract Book*

[https://www.predictioncenter.org/casp14/doc/CASP14\\_Abstracts.pdf](https://www.predictioncenter.org/casp14/doc/CASP14_Abstracts.pdf) (2020).

16. 16.

Brini, E., Simmerling, C. & Dill, K. Protein storytelling through physics. *Science* **370**, eaaz3041 (2020).

17. 17.

Sippl, M. J. Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins. *J. Mol. Biol.* **213**, 859–883 (1990).

18. 18.

Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).

19. 19.

Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protocols* **5**, 725–738 (2010).

20. 20.

Altschuh, D., Lesk, A. M., Bloomer, A. C. & Klug, A. Correlation of co-ordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. *J. Mol. Biol.* **193**, 693–707 (1987).

21. 21.

Shindyalov, I. N., Kolchanov, N. A. & Sander, C. Can three-dimensional contacts in protein structures be predicted by analysis of correlated mutations? *Protein Eng.* **7**, 349–358 (1994).

22. 22.

Weigt, M., White, R. A., Szurmant, H., Hoch, J. A. & Hwa, T. Identification of direct residue contacts in protein–protein interaction by message passing. *Proc. Natl Acad. Sci. USA* **106**, 67–72 (2009).

23. 23.

Marks, D. S. et al. Protein 3D structure computed from evolutionary sequence variation. *PLoS ONE* **6**, e28766 (2011).

24. 24.

Jones, D. T., Buchan, D. W. A., Cozzetto, D. & Pontil, M. PSICOV: precise structural contact prediction using sparse inverse covariance estimation on large multiple sequence alignments. *Bioinformatics* **28**, 184–190 (2012).

25. 25.

Moult, J., Pedersen, J. T., Judson, R. & Fidelis, K. A large-scale experiment to assess protein structure prediction methods. *Proteins* **23**, ii–iv (1995).

26. 26.

Kryshtafovych, A., Schwede, T., Topf, M., Fidelis, K. & Moult, J. Critical assessment of methods of protein structure prediction (CASP)-round XIII. *Proteins* **87**, 1011–1020 (2019).

27. 27.

Zhang, Y. & Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins* **57**, 702–710 (2004).

28. 28.

Tu, Z. & Bai, X. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 1744–1757 (2010).

29. 29.

Carreira, J., Agrawal, P., Fragkiadaki, K. & Malik, J. Human pose estimation with iterative error feedback. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 4733–4742 (2016).

30. 30.

Mirabello, C. & Wallner, B. rawMSA: end-to-end deep learning using raw multiple sequence alignments. *PLoS ONE* **14**, e0220182 (2019).

31. 31.

Huang, Z. et al. CCNet: criss-cross attention for semantic segmentation. In *Proc. IEEE/CVF International Conference on Computer Vision* 603–612 (2019).

32. 32.

Hornak, V. et al. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).

33. 33.

Zemla, A. LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Res.* **31**, 3370–3374 (2003).

34. 34.

Mariani, V., Biasini, M., Barbato, A. & Schwede, T. IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **29**, 2722–2728 (2013).

35. 35.

Xie, Q., Luong, M.-T., Hovy, E. & Le, Q. V. Self-training with noisy student improves imagenet classification. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition* 10687–10698 (2020).

36. 36.

Mirdita, M. et al. Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Res.* **45**, D170–D176 (2017).

37. 37.

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* 1, 4171–4186 (2019).

38. 38.

Rao, R. et al. MSA transformer. In *Proc. 38th International Conference on Machine Learning* PMLR 139, 8844–8856 (2021).

39. 39.

Tunyasuvunakool, K. et al. Highly accurate protein structure prediction for the human proteome. *Nature* <https://doi.org/10.1038/s41586-021-03828-1> (2021).

40. 40.

Kuhlman, B. & Bradley, P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* **20**, 681–697 (2019).

41. 41.

Marks, D. S., Hopf, T. A. & Sander, C. Protein structure prediction from sequence variation. *Nat. Biotechnol.* **30**, 1072–1080 (2012).

42. 42.

Qian, N. & Sejnowski, T. J. Predicting the secondary structure of globular proteins using neural network models. *J. Mol. Biol.* **202**, 865–884 (1988).

43. 43.

Fariselli, P., Olmea, O., Valencia, A. & Casadio, R. Prediction of contact maps with neural networks and correlated mutations. *Protein Eng.* **14**, 835–843 (2001).

44. 44.

Yang, J. et al. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl Acad. Sci. USA* **117**, 1496–1503 (2020).

45. 45.

Li, Y. et al. Deducing high-accuracy protein contact-maps from a triplet of coevolutionary matrices through deep residual convolutional networks. *PLOS Comput. Biol.* **17**, e1008865 (2021).

46. 46.

He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 770–778 (2016).

47. 47.

AlQuraishi, M. End-to-end differentiable learning of protein structure. *Cell Syst.* **8**, 292–301 (2019).

48. 48.

Senior, A. W. et al. Protein structure prediction using multiple deep neural networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13). *Proteins* **87**, 1141–1148 (2019).

49. 49.

Ingraham, J., Riesselman, A. J., Sander, C. & Marks, D. S. Learning protein structure with a differentiable simulator. in *Proc. International Conference on Learning Representations* (2019).

50. 50.

Li, J. Universal transforming geometric network. Preprint at <https://arxiv.org/abs/1908.00723> (2019).

51. 51.

Xu, J., McPartlon, M. & Li, J. Improved protein structure prediction by deep learning irrespective of co-evolution information. *Nat. Mach. Intell.* **3**, 601–609 (2021).

52. 52.

Vaswani, A. et al. Attention is all you need. In *Advances in Neural Information Processing Systems* 5998–6008 (2017).

53. 53.

Wang, H. et al. Axial-deeplab: stand-alone axial-attention for panoptic segmentation. in *European Conference on Computer Vision* 108–126 (Springer, 2020).

54. 54.

Alley, E. C., Khimulya, G., Biswas, S., AlQuraishi, M. & Church, G. M. Unified rational protein engineering with sequence-based deep representation learning. *Nat. Methods* **16**, 1315–1322 (2019).

55. 55.

Heinzinger, M. et al. Modeling aspects of the language of life through transfer-learning protein sequences. *BMC Bioinformatics* **20**, 723 (2019).

56. 56.

Rives, A. et al. Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proc. Natl Acad. Sci. USA* **118**, e2016239118 (2021).

57. 57.

Pereira, J. et al. High-accuracy protein structure prediction in CASP14. *Proteins* <https://doi.org/10.1002/prot.26171> (2021).

58. 58.

Gupta, M. et al. CryoEM and AI reveal a structure of SARS-CoV-2 Nsp2, a multifunctional protein involved in key host processes. Preprint at <https://doi.org/10.1101/2021.05.10.443524> (2021).

59. 59.

Ingraham, J., Garg, V. K., Barzilay, R. & Jaakkola, T. Generative models for graph-based protein design. in *Proc. 33rd Conference on Neural Information Processing Systems* (2019).

60. 60.

Johnson, L. S., Eddy, S. R. & Portugaly, E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* **11**, 431 (2010).

61. 61.

Remmert, M., Biegert, A., Hauser, A. & Söding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods* **9**, 173–175 (2012).

62. 62.

The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2020).

63. 63.

Steinegger, M. & Söding, J. Clustering huge protein sequence sets in linear time. *Nat. Commun.* **9**, 2542 (2018).

64. 64.

Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).

65. 65.

Deorowicz, S., Debudaj-Grabysz, A. & Gudyś, A. FAMSA: fast and accurate multiple sequence alignment of huge protein families. *Sci. Rep.* **6**, 33964 (2016).

66. 66.

Steinegger, M. et al. HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* **20**, 473 (2019).

67. 67.

Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B. & Wu, C. H. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31**, 926–932 (2015).

68. 68.

Eddy, S. R. Accelerated profile HMM searches. *PLOS Comput. Biol.* **7**, e1002195 (2011).

69. 69.

Eastman, P. et al. OpenMM 7: rapid development of high performance algorithms for molecular dynamics. *PLOS Comput. Biol.* **13**, e1005659 (2017).

70. 70.

Ashish, A. M. A. et al. TensorFlow: large-scale machine learning on heterogeneous systems. Preprint at <https://arxiv.org/abs/1603.04467> (2015).

71. 71.

Reynolds, M. et al. Open sourcing Sonnet – a new library for constructing neural networks. *DeepMind* <https://deepmind.com/blog/open-sourcing-sonnet/> (7 April 2017).

72. 72.

Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).

73. 73.

Van Rossum, G. & Drake, F. L. *Python 3 Reference Manual* (CreateSpace, 2009).

74. 74.

Bisong, E. in *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners* 59–64 (Apress, 2019).

75. 75.

TensorFlow. XLA: Optimizing Compiler for TensorFlow.  
<https://www.tensorflow.org/xla> (2018).

76. 76.

Wu, T., Hou, J., Adhikari, B. & Cheng, J. Analysis of several key factors influencing deep learning-based inter-residue contact prediction. *Bioinformatics* **36**, 1091–1098 (2020).

77. 77.

Jiang, W. et al. MrpH, a new class of metal-binding adhesin, requires zinc to mediate biofilm formation. *PLoS Pathog.* **16**, e1008707 (2020).

78. 78.

Dunne, M., Ernst, P., Sobieraj, A., Pluckthun, A. & Loessner, M. J. The M23 peptidase domain of the Staphylococcal phage 2638A endolysin. *PDB*  
<https://doi.org/10.22110/pdb6YJ1/pdb> (2020).

79. 79.

Drobysheva, A. V. et al. Structure and function of virion RNA polymerase of a crAss-like phage. *Nature* **589**, 306–309 (2021).

80. 80.

Flaugnatti, N. et al. Structural basis for loading and inhibition of a bacterial T6SS phospholipase effector by the VgrG spike. *EMBO J.* **39**, e104129 (2020).

81. 81.

ElGamacy, M. et al. An interface-driven design strategy yields a novel, corrugated protein architecture. *ACS Synth. Biol.* **7**, 2226–2235 (2018).

82. 82.

Lim, C. J. et al. The structure of human CST reveals a decameric assembly bound to telomeric DNA. *Science* **368**, 1081–1085 (2020).

83. 83.

Debruycker, V. et al. An embedded lipid in the multidrug transporter LmrP suggests a mechanism for polyspecificity. *Nat. Struct. Mol. Biol.* **27**, 829–835 (2020).

84. 84.

Flower, T. G. et al. Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. *Proc. Natl Acad. Sci. USA* **118**, e2021785118 (2021).

## Acknowledgements

We thank A. Rrustemi, A. Gu, A. Guseynov, B. Hechtman, C. Beattie, C. Jones, C. Donner, E. Parisotto, E. Elsen, F. Popovici, G. Necula, H. Maclean, J. Menick, J. Kirkpatrick, J. Molloy, J. Yim, J. Stanway, K. Simonyan, L. Sifre, L. Martens, M. Johnson, M. O'Neill, N. Antropova, R. Hadsell, S. Blackwell, S. Das, S. Hou, S. Gouws, S. Wheelwright, T. Hennigan, T. Ward, Z. Wu, Ž. Avsec and the Research Platform Team for their contributions; M. Mirdita for his help with the datasets; M. Piovesan-Forster, A. Nelson and R. Kemp for their help managing the project; the JAX, TensorFlow and XLA teams for detailed support and enabling machine learning models of the complexity of AlphaFold; our colleagues at DeepMind, Google and Alphabet for their encouragement and support; and J. Moult and the CASP14 organizers, and the experimentalists whose structures enabled the assessment. M.S. acknowledges support from the National Research Foundation of Korea grant (2019R1A6A1A10073437, 2020M3A9G7103933) and the Creative-Pioneering Researchers Program through Seoul National University.

## Author information

### Author notes

1. These authors contributed equally: John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishabh Jain, Demis Hassabis

## Affiliations

### 1. DeepMind, London, UK

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishabh Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli & Demis Hassabis

### 2. School of Biological Sciences, Seoul National University, Seoul, South Korea

Martin Steinegger

### 3. Artificial Intelligence Institute, Seoul National University, Seoul, South Korea

Martin Steinegger

## Contributions

J.J. and D.H. led the research. J.J., R.E., A. Pritzel, M.F., O.R., R.B., A. Potapenko, S.A.A.K., B.R.-P., J.A., M.P., T. Berghammer and O.V. developed the neural network architecture and training. T.G., A.Ž., K.T., R.B., A.B., R.E., A.J.B., A.C., S.N., R.J., D.R., M.Z. and S.B. developed the data, analytics and inference systems. D.H., K.K., P.K., C.M. and E.C. managed the research. T.G. led the technical platform. P.K., A.W.S., K.K., O.V., D.S., S.P. and T. Back contributed technical advice and ideas. M.S. created the BFD genomics database and provided technical assistance on HHBlits. D.H., R.E., A.W.S. and K.K. conceived the AlphaFold project. J.J., R.E. and A.W.S. conceived the end-to-end approach. J.J., A. Pritzel, O.R., A. Potapenko, R.E., M.F., T.G., K.T., C.M. and D.H. wrote the paper.

## Corresponding authors

Correspondence to [John Jumper](#) or [Demis Hassabis](#).

## Ethics declarations

### Competing interests

J.J., R.E., A. Pritzel, T.G., M.F., O.R., R.B., A.B., S.A.A.K., D.R. and A.W.S. have filed non-provisional patent applications 16/701,070 and PCT/EP2020/084238, and provisional patent applications 63/107,362, 63/118,917, 63/118,918, 63/118,921 and 63/118,919, each in the name of DeepMind Technologies Limited, each pending, relating to machine learning for predicting protein structures. The other authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Mohammed AlQuraishi, Charlotte Deane and Yang Zhang for their contribution to the peer review of this work.

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

## Supplementary information

### [Supplementary Information](#)

Description of the method details of the AlphaFold system, model, and analysis, including data pipeline, datasets, model blocks, loss functions, training and inference details, and ablations. Includes Supplementary Methods, Supplementary Figures, Supplementary Tables and Supplementary Algorithms.

### [Reporting Summary](#)

### [Supplementary Video 1](#)

Video of the intermediate structure trajectory of the CASP14 target T1024 (LmrP) A two-domain target (408 residues). Both domains are folded early, while their packing is adjusted for a longer time.

### [Supplementary Video 2](#)

Video of the intermediate structure trajectory of the CASP14 target T1044 (RNA polymerase of crAss-like phage). A large protein (2180 residues), with multiple domains. Some domains are folded quickly, while others take a considerable amount of time to fold.

### **Supplementary Video 3**

Video of the intermediate structure trajectory of the CASP14 target T1064 (Orf8). A very difficult single-domain target (106 residues) that takes the entire depth of the network to fold.

### **Supplementary Video 4**

Video of the intermediate structure trajectory of the CASP14 target T1091. A multi-domain target (863 residues). Individual domains' structure is determined early, while the domain packing evolves throughout the network. The network is exploring unphysical configurations throughout the process, resulting in long 'strings' in the visualization.

## **Rights and permissions**

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

### **Reprints and Permissions**

## **About this article**

### **Cite this article**

Jumper, J., Evans, R., Pritzel, A. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). <https://doi.org/10.1038/s41586-021-03819-2>

- Received: 11 May 2021
- Accepted: 12 July 2021
- Published: 15 July 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03819-2>

## **Highly accurate protein structure prediction for the human proteome**

- Kathryn Tunyasuvunakool
- Jonas Adler
- Demis Hassabis

Article Open Access 22 Jul 2021

## **Protein-structure prediction revolutionized**

- Mohammed AlQuraishi

News & Views 23 Aug 2021

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03819-2>

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

- Article
- Open Access
- [Published: 22 July 2021](#)

# Highly accurate protein structure prediction for the human proteome

- [Kathryn Tunyasuvunakool](#) ORCID: [orcid.org/0000-0002-8594-1074](https://orcid.org/0000-0002-8594-1074)<sup>1</sup>,
- [Jonas Adler](#) ORCID: [orcid.org/0000-0001-9928-3407](https://orcid.org/0000-0001-9928-3407)<sup>1</sup>,
- [Zachary Wu](#)<sup>1</sup>,
- [Tim Green](#) ORCID: [orcid.org/0000-0002-3227-1505](https://orcid.org/0000-0002-3227-1505)<sup>1</sup>,
- [Michał Zielinski](#)<sup>1</sup>,
- [Augustin Žídek](#)<sup>1</sup>,
- [Alex Bridgland](#)<sup>1</sup>,
- [Andrew Cowie](#)<sup>1</sup>,
- [Clemens Meyer](#)<sup>1</sup>,
- [Agata Laydon](#)<sup>1</sup>,
- [Sameer Velankar](#) ORCID: [orcid.org/0000-0002-8439-5964](https://orcid.org/0000-0002-8439-5964)<sup>2</sup>,
- [Gerard J. Kleywegt](#) ORCID: [orcid.org/0000-0002-4670-0331](https://orcid.org/0000-0002-4670-0331)<sup>2</sup>,
- [Alex Bateman](#) ORCID: [orcid.org/0000-0002-6982-4660](https://orcid.org/0000-0002-6982-4660)<sup>2</sup>,
- [Richard Evans](#) ORCID: [orcid.org/0000-0003-4675-8469](https://orcid.org/0000-0003-4675-8469)<sup>1</sup>,
- [Alexander Pritzel](#)<sup>1</sup>,
- [Michael Figurnov](#)<sup>1</sup>,
- [Olaf Ronneberger](#)<sup>1</sup>,
- [Russ Bates](#)<sup>1</sup>,
- [Simon A. A. Kohl](#) ORCID: [orcid.org/0000-0003-4271-4418](https://orcid.org/0000-0003-4271-4418)<sup>1</sup>,
- [Anna Potapenko](#)<sup>1</sup>,
- [Andrew J. Ballard](#)<sup>1</sup>,
- [Bernardino Romera-Paredes](#)<sup>1</sup>,
- [Stanislav Nikolov](#)<sup>1</sup>,
- [Rishabh Jain](#)<sup>1</sup>,
- [Ellen Clancy](#)<sup>1</sup>,
- [David Reiman](#)<sup>1</sup>,
- [Stig Petersen](#)<sup>1</sup>,
- [Andrew W. Senior](#) ORCID: [orcid.org/0000-0002-2401-5691](https://orcid.org/0000-0002-2401-5691)<sup>1</sup>,

- [Koray Kavukcuoglu<sup>1</sup>](#),
- [Ewan Birney ORCID: orcid.org/0000-0001-8314-8497<sup>2</sup>](#),
- [Pushmeet Kohli<sup>1</sup>](#),
- [John Jumper ORCID: orcid.org/0000-0001-6169-6580<sup>1 na1</sup>](#) &
- [Demis Hassabis ORCID: orcid.org/0000-0003-2812-9917<sup>1 na1</sup>](#)

*Nature* volume **596**, pages 590–596 (2021)

- 132k Accesses
- 1 Citations
- 1332 Altmetric
- [Metrics details](#)

## Subjects

- [Machine learning](#)
- [Protein structure predictions](#)
- [Proteomic analysis](#)
- [Structural biology](#)

## Abstract

Protein structures can provide invaluable information, both for reasoning about biological processes and for enabling interventions such as structure-based drug development or targeted mutagenesis. After decades of effort, 17% of the total residues in human protein sequences are covered by an experimentally determined structure<sup>1</sup>. Here we markedly expand the structural coverage of the proteome by applying the state-of-the-art machine learning method, AlphaFold<sup>2</sup>, at a scale that covers almost the entire human proteome (98.5% of human proteins). The resulting dataset covers 58% of residues with a confident prediction, of which a subset (36% of all residues) have very high confidence. We introduce several metrics developed by building on the AlphaFold model and use them to interpret the dataset, identifying strong multi-domain predictions as well as regions that are likely to be disordered. Finally, we provide some case studies to illustrate how high-quality predictions could be used to generate biological hypotheses. We are making our predictions freely available to the community and anticipate that routine large-scale and high-accuracy structure prediction will become an important tool that will allow new questions to be addressed from a structural perspective.

[Download PDF](#)

## Main

The monumental success of the human genome project revealed new worlds of protein-coding genes, and many researchers set out to map these proteins to their structures<sup>3,4</sup>. Thanks to the efforts of individual laboratories and dedicated structural genomics initiatives, more than 50,000 human protein structures have now been deposited, making *Homo sapiens* by far the best represented species in the Protein Data Bank (PDB)<sup>5</sup>. Despite this intensive study, only 35% of human proteins map to a PDB entry, and in many cases the structure covers only a fragment of the sequence<sup>6</sup>. Experimental structure determination requires overcoming many time-consuming hurdles: the protein must be produced in sufficient quantities and purified, appropriate sample preparation conditions chosen and high-quality datasets collected. A target may prove intractable at any stage, and depending on the chosen method, properties such as protein size, the presence of transmembrane regions, presence of disorder or susceptibility to conformational change can be a hindrance<sup>7,8</sup>. As such, full structural coverage of the proteome remains an outstanding challenge.

Protein structure prediction contributes to closing this gap by providing actionable structural hypotheses quickly and at scale. Previous large-scale structure prediction studies have addressed protein families<sup>9,10,11,12</sup>, specific functional classes<sup>13,14</sup>, domains identified within whole proteomes<sup>15</sup> and, in some cases, full chains or complexes<sup>16,17</sup>. In particular, projects such as the SWISS-MODEL Repository, Genome3D and ModBase have made valuable contributions by providing access to large numbers of structures and encouraging their free use by the community<sup>17,18,19</sup>. Related protein bioinformatics fields have developed alongside structure prediction, including protein design<sup>20,21</sup>, function annotation<sup>22,23,24</sup>, disorder prediction<sup>25</sup>, and domain identification and classification<sup>26,27,28</sup>. Although some of our analyses are inspired by these previous studies, here we focus mainly on structural investigations for which scale and accuracy are particularly beneficial.

Structure prediction has seen substantial progress in recent years, as evidenced by the results of the biennial Critical Assessment of protein Structure Prediction (CASP)<sup>29,30</sup>. In particular, the latest version of AlphaFold was entered in CASP14 under the team name ‘AlphaFold2’. This system used a completely different model from our CASP13 entry<sup>31</sup>, and demonstrated a considerable improvement over previous methods in terms of providing routinely high accuracy<sup>29,30</sup>. Backbone predictions with sub-Ångström root mean square deviation ( $\text{Ca}$  r.m.s.d.) are now common, and side chains are increasingly accurate<sup>2</sup>. Good results can often be achieved even for challenging proteins without a template structure in the PDB, or with relatively few related

sequences to build a multiple sequence alignment (MSA)<sup>2</sup>. These improvements are important, because more accurate models permit a wider range of applications: not only homology search and putative function assignment, but also molecular replacement and druggable pocket detection, for instance<sup>32,33,34</sup>. In light of this, we applied the current state-of-the-art method—AlphaFold—to the human proteome. All of our predictions can be accessed freely at <https://alphafold.ebi.ac.uk/>, hosted by the European Bioinformatics Institute.

## Model confidence and added coverage

We predicted structures for the UniProt human reference proteome (one representative sequence per gene), with an upper length limit of 2,700 residues<sup>6</sup>. The final dataset covers 98.5% of human proteins with a full chain prediction.

For the resulting predictions to be practically useful, they must come with a well-calibrated and sequence-resolved confidence measure. The latter point is particularly important when predicting full chains, as we expect to see high confidence on domains but low confidence on linkers and unstructured regions (Extended Data Fig. 1). To this end, AlphaFold produces a per-residue confidence metric called predicted local distance difference test (pLDDT) on a scale from 0 to 100. pLDDT estimates how well the prediction would agree with an experimental structure based on the local distance difference test  $\text{Ca}$  (IDDT- $\text{Ca}$ )<sup>35</sup>. It has been shown to be well-calibrated (Fig. 1a, Extended Data Fig. 2 and Extended Data Table 1) and full details on how the pLDDT is produced are given in the supplementary information of the companion AlphaFold paper<sup>2</sup>.

**Fig. 1: Model confidence and added coverage.**

---

 **figure1**

**a**, Correlation between per-residue pLDDT and IDDT- $\text{C}\alpha$ . Data are based on a held-out set of recent PDB chains (Methods) filtered to those with a reported resolution of  $<3.5\text{ \AA}$  ( $n = 10,215$  chains and 2,756,569 residues). The scatterplot shows a subsample (1% of residues), with the blue line showing a least-squares linear fit and the shaded region a 95% confidence interval estimated with 1,000 bootstrap samples. The black line shows  $x = y$ , for comparison. The smaller plot is a magnified region of the larger one. On the full dataset, the Pearson's  $r = 0.73$  and the least-squares linear fit is  $y = (0.967 \pm 0.001) \times x + (1.9 \pm 0.1)$ . **b**, AlphaFold prediction and experimental structure for a CASP14 target (PDB: 6YJ1)<sup>64</sup>. The prediction is coloured by model confidence band, and the N terminus is an expression tag included in CASP but unresolved in the PDB structure. **c**, AlphaFold model confidence on all residues for which a prediction was produced ( $n = 10,537,122$  residues). Residues covered by a template at the specified identity level are shown in a lighter colour and a heavy dashed line separates these from residues without a template. **d**, Added residue-level coverage of the proteome for high-level GO terms, on top of residues covered by a template with sequence identity of more than 50%. Based on the same human proteome dataset as in **c** ( $n = 10,537,122$  residues).

We consider a prediction highly accurate when—in addition to a good backbone prediction—the side chains are frequently correctly oriented. On this basis, pLDDT > 90 is taken as the high accuracy cut-off, above which AlphaFold  $\chi_1$  rotamers are 80% correct for a recent PDB test dataset (Extended Data Fig. 3). A lower cut-off of pLDDT > 70 corresponds to a generally correct backbone prediction (Extended Data Table 2). The accuracy of AlphaFold within a number of pLDDT bands is illustrated for an example protein in Fig. 1b.

Of the human proteome, 35.7% of total residues fall within the highest accuracy band (corresponding to 38.6% of residues for which a prediction was produced) (Fig. 1c). This is double the number of residues covered by an experimental structure. In total, 58.0% of residues were predicted confidently (pLDDT > 70), indicating that we also add substantial coverage for sequences without a good template in PDB (with a sequence identity below 30%). At the per-protein level, 43.8% of proteins have a confident prediction on at least three quarters of their sequence, while 1,290 proteins contain a substantial region (more than 200 residues) with  $p\text{LDDT} \geq 70$  and no good template.

The dataset adds high-quality structural models across a broad range of Gene Ontology (GO) terms<sup>36,37</sup>, including pharmaceutically relevant classes such as enzymes and membrane proteins<sup>38</sup> (Fig. 1d). Membrane proteins, in particular, are generally underrepresented in the PDB because they have historically been challenging experimental targets. This shows that AlphaFold is able to produce confident predictions even for protein classes that are not abundant within its training set.

We note that the accuracy of AlphaFold was validated in CASP14<sup>2</sup>, which focuses on challenging proteins that are dissimilar to structures already available in the PDB. By contrast, many human proteins have templates with high sequence identity. To evaluate the applicability of AlphaFold to this collection, we predicted structures for 1 year of targets from the Continuous Automated Model Evaluation (CAMEO) benchmark<sup>39,40</sup>—a structure-prediction assessment that measures a wider range of difficulties. We find that AlphaFold adds substantial accuracy over the BestSingleStructuralTemplate baseline of CAMEO across a wide range of levels of template identity (Extended Data Fig. 4).

## Prediction of full-length protein chains

Many previous large-scale structure prediction efforts have focused on domains—regions of the sequence that fold independently<sup>9,10,11,15</sup>. Here we process full-length protein chains. There are several motivations for this. Restricting the prediction to pre-identified domains risks missing structured regions that have yet to be annotated. It

also discards contextual information from the rest of the sequence, which might be useful in cases in which two or more domains interact substantially. Finally, the full chain approach lets the model attempt an inter-domain packing prediction.

Inter-domain accuracy was assessed at CASP14, and AlphaFold outperformed other methods<sup>41</sup>. However, the assessment was based on a small target set. To further evaluate AlphaFold on long multi-domain proteins, we compiled a test dataset of recent PDB chains that were not in the training set of the model. Only chains with more than 800 resolved residues were included, and a template filter was applied (Methods). Performance on this set was evaluated using the template modelling score (TM-score<sup>42</sup>), which should better reflect global, as opposed to per-domain, accuracy. The results were encouraging, with 70% of predictions having a TM-score > 0.7 (Fig. 2a).

**Fig. 2: Full chain structure prediction.**

 figure2

**a**, TM-score distribution for AlphaFold evaluated on a held-out set of template-filtered, long PDB chains ( $n = 151$  chains). Includes recent PDB proteins with more than 800 resolved residues and best 50% coverage template below 30% identity. **b**, Correlation between full chain TM-score and pTM on the same set ( $n = 151$  chains), Pearson's  $r = 0.84$ . The ground truth and predicted structure are shown for the most over-optimistic outlier (PDB: 6OFS, chain A). **c**, pTM distribution on a subset of the

human proteome that we expect to be enriched for structurally novel multidomain proteins ( $n = 1,165$  chains). Human proteome predictions comprise more than 600 confident residues (more than 50% coverage) and no proteins with 50% coverage templates. **d**, Four of the top hits from the set shown in **c**, filtering by  $pTM > 0.8$  and sorting by number of confident residues. Proteins are labelled by their UniProt accession. For clarity, regions with  $pLDDT < 50$  are hidden, as are isolated smaller regions that were left after this cropping.

The supplementary information of the companion AlphaFold paper<sup>2</sup> describes how a variety of useful predictors can be built on top of the main model. In particular, we can predict the residues that are likely to be experimentally resolved, and use them to produce a predicted TM-score ( $pTM$ ), in which the contribution of each residue is weighted by the probability of it being resolved ([Supplementary Methods 1](#)). The motivation for the weighting is to downweight unstructured parts of the prediction, producing a metric that better reflects the confidence of the model about the packing of the structured domains that are present. On the same recent PDB test dataset,  $pTM$  correlates well with the actual TM-score (Pearson's  $r = 0.84$ ) (Fig. [2b](#)). Notably, although some outliers in this plot are genuine failure cases, others appear to be plausible alternate conformations (for example, 6OFS chain A<sup>43</sup> in Fig. [2b](#)).

We computed  $pTM$  scores for the human proteome, in an effort to identify multi-domain predictions that could feature novel domain packings. The criteria applied were a  $pLDDT > 70$  on at least 600 residues constituting over half the sequence, with no template hit covering more than half the sequence. The distribution of  $pTM$  scores after applying the above filters is shown in Fig. [2c](#). Note that we would not expect uniformly high TM-scores to be achievable for this set, as some proteins will contain domains that are mobile relative to each other, with no fixed packing. Of the set, 187 proteins have  $pTM > 0.8$  and 343 have  $pTM > 0.7$ . Although we expect the inter-domain accuracy of AlphaFold to be lower than its within-domain accuracy, this set should nonetheless be enriched for interesting multi-domain predictions, suggesting that the dataset provides on the order of hundreds of these. Four examples—the predictions with the highest number of confident residues subject to  $pTM > 0.8$ —are shown in Fig. [2d](#).

## Highlighted predictions

We next discuss some case study predictions and the insights that they may provide. All predictions presented are de novo, lacking any template with 25% sequence identity or more covering 20% of the sequence. Our discussion concerns biological hypotheses, which would ultimately need to be confirmed by experimental studies.

### Glucose-6-phosphatase

G6Pase- $\alpha$  (UniProt P35575) is a membrane-bound enzyme that catalyses the final step in glucose synthesis; it is therefore of critical importance to maintaining blood sugar levels. To our knowledge, no experimental structure exists, but previous studies have attempted to characterize the transmembrane topology<sup>44</sup> and active site<sup>45</sup>. Our prediction has very high confidence (median pLDDT of 95.5) and gives a nine-helix topology with the putative active site accessible via an entry tunnel that is roughly in line with the surface of the endoplasmic reticulum (Fig. 3a and [Supplementary Video 1](#)). Positively charged residues in our prediction (median pLDDT of 96.6) align closely with the previously identified active site homologue in a fungal vanadium chloroperoxidase (PDB 1IDQ; r.m.s.d. of 0.56 Å; 49 out of 51 aligned atoms)<sup>46</sup>. As these enzymes have distinct functions, we investigated our prediction for clues about substrate specificity. In the G6Pase- $\alpha$  binding pocket face, opposite the residues shared with the chloroperoxidase, we predict a conserved glutamate (Glu110) that is also present in our G6Pase- $\beta$  prediction (Glu105) but not in the chloroperoxidase (Fig. 3a). The glutamate could stabilize the binding pocket in a closed conformation, forming salt bridges with positively charged residues there. It is also the most solvent-exposed residue of the putative active site, suggesting a possible gating function. To our knowledge, this residue has not been discussed previously and illustrates the novel mechanistic hypotheses that can be obtained from high-quality structure predictions.

**Fig. 3: Highlighted structure predictions.**

---

 **figure3**

**a**, Left, comparison of the active sites of two G6Pases (G6Pase- $\alpha$  and G6Pase- $\beta$ ) and a chloroperoxidase (PDB 1IDQ). The G6Pases are glucose-forming enzymes that contain a conserved, solvent-accessible glutamate (red; right) opposite the shared active-site residues (middle). **b**, Left, pocket prediction (P2Rank<sup>65</sup>) identifies a putative binding pocket for DGAT2, which is involved in body-fat synthesis. Red and green spheres represent the ligandability scores by P2Rank of 1 and 0, respectively. Middle, a proposed mechanism for DGAT1<sup>51</sup> activates the substrate with Glu416 and His415, which have analogous residues in the DGAT2 pocket. The docked inhibitor is well placed for polar interactions with His163 and Thr194 (right). The chemical structure (middle) is adapted from ref. <sup>51</sup>. **c**, Predicted structure of wolframin, mutations in which cause Wolfram syndrome. Although there are regions in wolframin with low pLDDT (left), we could identify an OB-fold region (green/yellow), with a comparable core to a prototypical OB-fold (grey; middle). However, the most similar PDB chain (magenta; right) lacks the conserved cysteine-rich region (yellow) of our prediction. This region forms the characteristic  $\beta$ 1 strand and an extended L12 loop, and is predicted to contain three disulfide bridges (yellow mesh).

## Diacylglycerol *O*-acyltransferase 2

Triacylglycerol synthesis is responsible for storing excess metabolic energy as fat in adipose tissue. DGAT2 (UniProt Q96PD7) is one of two essential acyltransferases catalysing the final acyl addition in this pathway, and inhibiting DGAT2 has been shown to improve liver function in mouse models of liver disease<sup>47</sup>. With our highly confident predicted structure (median pLDDT of 95.9), we set out to identify the binding pocket for a known inhibitor, PF-06424439 (ref. <sup>48</sup>). We identified a pocket (median pLDDT of 93.7) in which we were able to dock the inhibitor and observe specific interactions (Fig. <sup>3b</sup>) that were not recapitulated in a negative example<sup>49</sup> (Extended Data Fig. <sup>5</sup> and [Supplementary Methods 2](#)). DGAT2 has an evolutionarily divergent but biochemically similar analogue, diacylglycerol *O*-acyltransferase 1 (DGAT1)<sup>50</sup>. Within the binding pocket of DGAT2, we identified residues (Glu243 and His163) (Fig. <sup>3b</sup>) that are analogous to the proposed catalytic residues in DGAT1 (His415 and Glu416)<sup>51</sup>, although we note that the nearby Ser244 in DGAT2 may present an alternative mechanism through an acyl-enzyme intermediate. Previous experimental research with DGAT2 has shown that mutating His163 has a stronger deleterious effect than mutating a histidine that is two residues away<sup>52</sup>. Additionally, Glu243 and His163 are conserved across species<sup>50</sup>, supporting this hypothesized catalytic geometry.

## Wolframin

Wolframin (UniProt O76024) is a transmembrane protein localized to the endoplasmic reticulum. Mutations in the *WFS1* gene are associated with Wolfram syndrome 1, a neurodegenerative disease characterized by early onset diabetes, gradual visual and hearing loss, and early death<sup>53,54</sup>. Given the lower confidence in our full prediction (median pLDDT of 81.7) (Fig. <sup>3c</sup>), we proposed identifying regions that are unique to this structure. A recent evolutionary analysis suggested domains for wolframin, which our prediction largely supports<sup>55</sup>. An interesting distinction is the incorporation of a cysteine-rich domain (Fig. <sup>3c</sup>, yellow) to the oligonucleotide binding (OB) fold (Fig. <sup>3c</sup>, green and yellow) as the characteristic  $\beta$ 1 strand<sup>56</sup>. The cysteine-rich region then forms an extended L12 loop with two predicted disulfide bridges, before looping back to the prototypical  $\beta$ 2 strand. Comparing our prediction for this region (median pLDDT of 86.0) to existing PDB chains using TM-align<sup>42,57</sup> identified 3F1Z<sup>58</sup> as the most similar known chain (TM-score of 0.472) (Fig. <sup>3c</sup>, magenta). Despite being the most similar chain, 3F1Z lacks the cysteines that are present in wolframin, which could form disulfide cross-links in the endoplasmic reticulum<sup>59</sup>. As this region is hypothesized to recruit other proteins<sup>55</sup>, these structural insights are probably important to understanding its partners.

## Regions without a confident prediction

As we are applying AlphaFold to the entire human proteome, we would expect a considerable percentage of residues to be contained in regions that are always or sometimes disordered in solution. Disorder is common in the proteomes of eukaryotes<sup>60,61</sup>, and one previous study<sup>62</sup> estimated that the percentage of disordered residues in the human proteome is between 37% and 50%. Thus disorder will have a large role when we consider a comprehensive set of predictions that covers an entire proteome.

Furthermore, we observed a large difference in the pLDDT distribution between resolved and unresolved residues in PDB sequences (Fig. 4a). To investigate this connection, we evaluated pLDDT as a disorder predictor on the Critical Assessment of protein Intrinsic Disorder prediction (CAID) benchmark dataset<sup>25</sup>. The results showed pLDDT to be a competitive disorder predictor compared with the current state of the art (SPOT-Disorder2<sup>63</sup>), with an area under the curve (AUC) of 0.897 (Fig. 4b). Moreover, the supplementary information of the companion AlphaFold paper<sup>2</sup> describes an ‘experimentally resolved head’, which is specifically trained for the task of predicting whether a residue will be resolved in an experimental structure. The experimentally resolved head performed even better on the CAID benchmark, with an AUC of 0.921.

**Fig. 4: Low-confidence regions.**

---

 **figure4**

**a**, pLDDT distribution of the resolved parts of PDB sequences ( $n = 3,440,359$  residues), the unresolved parts of PDB sequences ( $n = 589,079$  residues) and the human proteome ( $n = 10,537,122$  residues). **b**, Performance of pLDDT and the experimentally resolved head of AlphaFold as disorder predictors on the CAID Disprot-PDB benchmark dataset ( $n = 178,124$  residues). **c**, An example low-confidence prediction aligned to the corresponding PDB submission (7KPx chain C)<sup>66</sup>. The globular domain is well-predicted but the extended interface exhibits low pLDDT and is incorrect apart from some of the secondary structure. a.a., amino acid. **d**, A high ratio of heterotypic contacts is associated with a lower AlphaFold accuracy on the recent PDB dataset, restricted to proteins with fewer than 40% of residues with template identity above 30% ( $n = 3,007$  chains) (Methods). The ratio of heterotypic contacts is defined as: heterotypic/(intra-chain + homomeric + heterotypic).

These disorder prediction results suggest that a considerable percentage of low-confidence residues may be explained by some form of disorder, but we caution that this could encompass both regions that are intrinsically disordered and regions that are structured only in complex. A potential example of the latter scenario drawn from a

recent PDB structure is shown in Fig. 4c; chain C interacts extensively with the rest of the complex, such that the interface region would be unlikely to adopt the same structure outside of this context. In a systematic analysis of recent PDB chains, we observed that AlphaFold has much lower accuracy for regions in which the chain has a high percentage of heterotypic, cross-chain contacts (Fig. 4d).

In summary, our current interpretation of regions in which AlphaFold exhibits low pLDDT is that they have high likelihood of being unstructured in isolation. In the current dataset, long regions with  $p\text{LDDT} < 50$  adopt a readily identifiable ribbon-like appearance, and should not be interpreted as structures but rather as a prediction of disorder.

## Discussion

In this study, we generated comprehensive, state-of-the-art structure predictions for the human proteome. The resulting dataset makes a large contribution to the structural coverage of the proteome; particularly for tasks in which high accuracy is advantageous, such as molecular replacement or the characterization of binding sites. We also applied several metrics produced by building on the AlphaFold architecture—pLDDT, pTM and the experimentally resolved head—to demonstrate how they can be used to interpret our predictions.

Although we present several case studies to illustrate the type of insights that may be gained from these data, we recognize that there is still much more to uncover. By making our predictions available to the community via <https://alphafold.ebi.ac.uk/>, we hope to enable exploration of new directions in structural bioinformatics.

The parts of the human proteome that are still without a confident prediction represent directions for future research. Some proportion of these will be genuine failures, in which a fixed structure exists but the current version of AlphaFold does not predict it. In many other cases, in which the sequence is unstructured in isolation, the problem arguably falls outside the scope of single-chain structure prediction. It will be crucial to develop new methods that can address the biology of these regions—for example, by predicting the structure in complex or by predicting a distribution over possible states in the cellular milieu.

Finally, we note that the importance of the human proteome for health and medicine has led to it being intensively studied from a structural perspective. Other organisms are much less well represented in the PDB, including biologically important, medically relevant or economically important species. Structure prediction may have a more profound effect on the study of these organisms, for which fewer experimental structures are available. Looking beyond the proteome scale, the UniProt database

contains hundreds of millions of proteins that have so far been addressed mainly by sequence-based methods, and for which the easy availability of structures could open up entirely new avenues of investigation. By providing scalable structure prediction with very high accuracy, AlphaFold could enable an exciting shift towards structural bioinformatics, further illuminating protein space.

## Methods

### Structure prediction (human proteome)

Sequences for the human reference proteome were obtained from UniProt release 2021\_02<sup>6</sup>. Structure prediction was attempted for all sequences with 16–2,700 amino acids; sequences with residue codes B, J, O, U, Z or X were excluded. The length ceiling of 2,700 residues does not represent an absolute limit for the method, but was chosen to keep run times manageable. The structure prediction process was largely as described in the AlphaFold paper<sup>2</sup>, consisting of five steps: MSA construction, template search, inference with five models, model ranking based on mean pLDDT and constrained relaxation of the predicted structures. The following differences were introduced for the proteome-scale pipeline. First, the search against the metagenomics database Big Fantastic Database (BFD) was replaced with a search against ‘Reduced BFD’ using Jackhmmer from HMMER3<sup>67,68</sup>. Reduced BFD consists of a multiline FASTA file containing the first non-consensus sequence from each BFD a3m alignment. Second, the amount of ensembling was reduced by a factor of eight. At least four relaxed full chain models were successfully produced for 20,296 sequences out of 20,614 FASTA entries, covering 98.5% of proteins. Sequences with more than 2,700 residues account for the majority of exclusions. This amounts to 10,537,122 residues (92.5% of residues).

### Structure prediction (recent PDB dataset)

For structure predictions of recent PDB sequences, we used a copy of the PDB downloaded on 15 February 2021. Structures were filtered to those with a release date after 30 April 2018 (the date limit for inclusion in the training set). Chains were then further filtered to remove sequences that consisted of a single amino acid, sequences with an ambiguous chemical component at any residue position and sequences without a PDB 40% sequence clustering. Exact duplicates were removed by choosing the chain with the most resolved Cα atoms as the representative sequence. Then, structures with fewer than 16 resolved residues, with unknown residues and structures solved by NMR methods were filtered out. Structure prediction then followed the same procedure as for the human proteome with the same length and residue limits, except that templates with a release date after 30 April 2018 were disallowed. Finally,

the dataset was redundancy reduced, by taking the chain with the best non-zero resolution from each cluster in the PDB 40% sequence clustering, producing a dataset of 12,494 chains. This is referred to as the recent PDB dataset.

## Computational resources

Inference was run on V100 graphics processing units (GPUs), with each sequence inferenced five times to produce five inputs to model selection. To prevent out-of-memory errors, long sequences were assigned to multi-GPU workers. Specifically, sequences of length 1,401–2,000 residues were processed by workers with two GPUs, and those of length 2,001–2,700 residues by workers with four GPUs (further details of unified memory on longer proteins are provided in the companion paper<sup>2</sup>; it is possible higher memory workers could be used without additional GPUs).

The total resources used for inference were logged and amounted to 930 GPU days. This accounts for generating five models per protein; around 190 GPU days would be sufficient to inference each protein once. Long sequences had a disproportionate effect owing to the multi-GPU workers described above. Approximately 250 GPU days would have been sufficient to produce five models for all proteins shorter than 1,400 residues. For reference, Extended Data Fig. 6 shows the relationship between sequence length and inference time.

All other stages of the pipeline (MSA search, template search and constrained relaxation) ran on the central processing unit (CPU) and used standard tools. Our human proteome run made use of some cached intermediates (for example, stored MSA search results). However, we estimate the total cost of running these stages from scratch at 510 core days. This estimate is based on taking a sample of 240 human proteins stratified by length, timing each stage when run with empty caches, fitting a quadratic relationship between sequence length and run time, then applying that relationship to the sequences in the human proteome. Extended Data Figure 7 shows the data used to make this estimate.

## Template coverage

Except where otherwise noted, template coverage was estimated on a per-residue basis as follows. Hmmsearch was run against a copy of the PDB SEQRES (downloaded on 15 February 2021) using default flags<sup>67</sup>. The prior template coverage at residue  $i$  is the maximum percentage sequence identity of all hits covering residue  $i$ , regardless of whether the hit residue is experimentally resolved. For the recent PDB analysis, only template hits corresponding to a structure released before 30 April 2018 were accepted.

In the section on full chain prediction, template filtering is based on the highest sequence identity of any single Hmmsearch hit with more than 50% coverage. This is because high-coverage templates are particularly relevant when considering whether a predicted domain packing is novel.

## GO term breakdown

GO annotations were taken from the XML metadata for the UniProt human reference proteome and were matched to the Gene Ontology in obo format<sup>36,37</sup>. One erroneous is\_a relationship was manually removed (GO:0071702 is\_a GO:0006820, see change log <https://www.ebi.ac.uk/QuickGO/term/GO:0071702>). The ontology file was used to propagate the GO annotations using is\_a and part\_of relations to assign parent-child relationships, and accounting for alternative IDs.

GO terms were then filtered to a manageable number for display, first by filtering for terms with more than 3,000 annotations, and from those selecting only moderately specific terms (a term cannot have a child with more than 3,000 annotations). The remaining terms in the ‘molecular function’ and ‘cellular component’ ontologies are shown in Fig. 1d.

## Structure analysis

Structure images were created in PyMOL<sup>69</sup>, and PyMOL align was used to compute r.m.s.d.s (outlier rejection is described in the text where applicable).

For docking against DGAT2, P2Rank<sup>65</sup> was used to identify ligand-binding pockets in the AlphaFold structure. AutoDockTools<sup>70</sup> was used to convert the AlphaFold prediction to PDBQT format. For the ligands, DGAT2-specific inhibitor (CAS number 1469284-79-4) and DGAT1-specific inhibitor (CAS number 942999-61-3) were also prepared in PDBQT format using AutoDockTools. AutoDock Vina<sup>71</sup> was run with an exhaustiveness parameter of 32, a seed of 0 and a docking search space of  $25 \times 25 \times 25 \text{ \AA}^3$  centred at the point identified by P2Rank.

For identifying the most similar structure to wolframin, TM-align<sup>42</sup> was used to compare against all PDB chains (downloaded 15 February 2021) with our prediction as the reference. This returned 3F1Z with a TM-score of 0.472.

## Additional metrics

The implementation of pTM is described in supplementary information section 1.9.7 of the companion AlphaFold paper<sup>2</sup> and the implementation of the experimentally resolved head is described in supplementary information section 1.9.10 of the

companion AlphaFold paper<sup>2</sup>. The weighted version of pTM is described in [Supplementary Methods 1](#).

## Analysis of low-confidence regions

For evaluation on CAID, the target sequences and ground-truth labels for the Disprot-PDB dataset were downloaded from <https://idpcentral.org/>. Structure prediction was performed as described above for the recent PDB dataset, with a template cut-off of 30 April 2018. To enable complete coverage, two sequences containing non-standard residues (X, U) had these remapped to G (glycine). Sequences longer than 2,000 residues were split into two segments: 1–2,000 and 2,000–end, and the pLDDT and experimentally resolved head arrays were concatenated for evaluation. The two evaluated disorder predictors were taken to be  $1 - 0.01 \times \text{pLDDT}$  and  $1 - \text{predicted resolvability for C}\alpha\text{ atoms}$ .

To obtain the ratio of heterotypic contacts to all contacts (Fig. 4d), two residues are considered in contact if their C $\beta$  atoms (or C $\alpha$  for glycine) are within 8 Å and if they are separated in primary sequence by at least three other residues (to exclude contacts within an  $\alpha$ -helix). Heteromers are identified as protein entities with a different entity\_id in the structure mmCIF file.

## Comparison with BestSingleStructuralTemplate

CAMEO data for the period 21 March 2020 to 13 March 2021 were downloaded from the CAMEO website. AlphaFold predictions were produced for all sequences in the target.fasta files, using the same procedure detailed above but with a maximum template date of 1 March 2020. Predictions were scored against the CAMEO ground truth using IDDT-C $\alpha$ . For BestSingleStructuralTemplate, IDDT-C $\alpha$  scores were taken from the CAMEO JavaScript Object Notation (JSON) files provided. Structures solved by solution NMR and solid-state NMR were filtered out at the analysis stage. To determine the template identity, templates were drawn from a copy of the PDB downloaded on 15 February 2021 with a template search performed using Hmmsearch. Templates were filtered to those with at least 70% coverage of the sequence and a release date before the query. The template with the highest *e*-value after filtering was used to compute the template identity. Targets were binned according to template identity, with width 10 bins ranging from 30 to 90. Extended Data Figure 4 shows the distribution of IDDT-C $\alpha$  for each model within each bin as a box plot (horizontal line at the median, box spanning from the lower to the upper quartile, whiskers extending to the minimum and maximum. In total 428 targets were included in the analysis.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

Structure predictions by AlphaFold for the human proteome are available under a CC-BY-4.0 license at <https://alphafold.ebi.ac.uk/>. All input data are freely available from public sources. The human reference proteome together with its XML annotations was obtained from UniProt v.2021\_02

([https://ftp.ebi.ac.uk/pub/databases/uniprot/previous\\_releases/release-2021\\_02/knowledgebase/](https://ftp.ebi.ac.uk/pub/databases/uniprot/previous_releases/release-2021_02/knowledgebase/)). At prediction time, MSA search was performed against UniRef90 v.2020\_03

([https://ftp.ebi.ac.uk/pub/databases/uniprot/previous\\_releases/release-2020\\_03/uniref/](https://ftp.ebi.ac.uk/pub/databases/uniprot/previous_releases/release-2020_03/uniref/)), MGnify clusters v.2018\_12

([https://ftp.ebi.ac.uk/pub/databases/metagenomics/peptide\\_database/2018\\_12/](https://ftp.ebi.ac.uk/pub/databases/metagenomics/peptide_database/2018_12/)) and a reduced version of BFD (produced as outlined in the Methods using the BFD (<https://bfd.mmseqs.com/>)). Template structures, the SEQRES fasta file and the 40% sequence clustering were taken from a copy of the PDB downloaded on 15 February 2021 (<https://www.wwpdb.org/ftp/pdb-ftp-sites>; see also

[https://ftp.wwpdb.org/pub/pdb/derived\\_data/](https://ftp.wwpdb.org/pub/pdb/derived_data/) and

<https://cdn.rcsb.org/resources/sequence/clusters/bc-40.out> for sequence data).

Experimental structures were drawn from the same copy of the PDB; we show structures with accessions [6YJ1<sup>64</sup>](#), [6QFS<sup>43</sup>](#), [1IDQ<sup>46</sup>](#), [1PRT<sup>72</sup>](#), [3F1Z<sup>58</sup>](#), [7KPx<sup>66</sup>](#) and [6VP0<sup>51</sup>](#). The template search used PDB70, downloaded on 10 February 2021 ([http://wwwuser.gwdg.de/~comppbiol/data/hhsuite/databases/hhsuite\\_dbs/](http://wwwuser.gwdg.de/~comppbiol/data/hhsuite/databases/hhsuite_dbs/)). The CAID dataset was downloaded from <https://idpcentral.org/caid/data/1/reference/disprot-disorder-pdb-atleast.txt>. CAMEO data was accessed on 17 March 2021 at [https://www.cameo3d.org/static/downloads/modeling/1-year/raw\\_targets-1-year\\_public.tar.gz](https://www.cameo3d.org/static/downloads/modeling/1-year/raw_targets-1-year_public.tar.gz). A copy of the current Gene Ontology database was downloaded on 29 April 2021 from <http://current.geneontology.org/ontology/go.obo>. Source data are provided with this paper.

## Code availability

Source code for the AlphaFold model, trained weights and an inference script are available under an open-source license at <https://github.com/deepmind/alphafold>.

Neural networks were developed with TensorFlow v.1

(<https://github.com/tensorflow/tensorflow>), Sonnet v.1

(<https://github.com/deepmind/sonnet>), JAX v.0.1.69 (<https://github.com/google/jax/>) and Haiku v.0.0.4 (<https://github.com/deepmind/dm-haiku>).

For MSA search on UniRef90, MGnify clusters and the reduced BFD, we used jackhmmer and for the template search on the PDB SEQRES we used hmmsearch, both from HMMER v.3.3 (<http://eddylab.org/software/hmmer/>). For the template search against PDB70, we used HHsearch from HH-suite v.3.0-beta.3 14/07/2017 (<https://github.com/soedinglab/hh-suite>). For constrained relaxation of structures, we used OpenMM v.7.3.1 (<https://github.com/openmm/openmm>) with the Amber99sb force field.

Docking analysis on DGAT used P2Rank v.2.1 (<https://github.com/rdk/p2rank>), MGLTools v.1.5.6 (<https://ccsb.scripps.edu/mgltools>) and AutoDockVina v.1.1.2 (<http://vina.scripps.edu/download/>) on a workstation running Debian GNU/Linux rodete 5.10.40-1rodete1-amd64 x86\_64.

Data analysis used Python v.3.6 (<https://www.python.org/>), NumPy v.1.16.4 (<https://github.com/numpy/numpy>), SciPy v.1.2.1 (<https://www.scipy.org/>), seaborn v.0.11.1 (<https://github.com/mwaskom/seaborn>), scikit-learn v.0.24.0 (<https://github.com/scikit-learn/>), Matplotlib v.3.3.4 (<https://github.com/matplotlib/matplotlib>), pandas v.1.1.5 (<https://github.com/pandas-dev/pandas>) and Colab (<https://research.google.com/colaboratory>). TM-align v.20190822 (<https://zhanglab.dcmb.med.umich.edu/TM-align>) was used for computing TM-scores. Structure analysis used Pymol v.2.3.0 (<https://github.com/schrodinger/pymol-open-source>).

## References

1. 1.

SWISS-MODEL. *Homo sapiens* (human).  
<https://swissmodel.expasy.org/repository/species/9606> (2021).

2. 2.

Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* <https://doi.org/10.1038/s41586-021-03819-2> (2021).

3. 3.

International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

4. 4.

Venter, J. C. et al. The sequence of the human genome. *Science* **291**, 1304–1351 (2001).

5. 5.

wwPDB Consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **47**, D520–D528 (2018).

6. 6.

The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).

7. 7.

Slabinski, L. et al. The challenge of protein structure determination—lessons from structural genomics. *Protein Sci.* **16**, 2472–2482 (2007).

8. 8.

Elmlund, D., Le, S. N. & Elmlund, H. High-resolution cryo-EM: the nuts and bolts. *Curr. Opin. Struct. Biol.* **46**, 1–6 (2017).

9. 9.

Yang, J. et al. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl Acad. Sci. USA* **117**, 1496–1503 (2020).

10. 10.

Greener, J. G., Kandathil, S. M. & Jones, D. T. Deep learning extends de novo protein modelling coverage of genomes using iteratively predicted structural constraints. *Nat. Commun.* **10**, 3977 (2019).

11. 11.

Michel, M., Menéndez Hurtado, D., Uziela, K. & Elofsson, A. Large-scale structure prediction by improved contact predictions and model quality assessment. *Bioinformatics* **33**, i23–i29 (2017).

12. 12.

Ovchinnikov, S. et al. Large-scale determination of previously unsolved protein structures using evolutionary information. *eLife* **4**, e09248 (2015).

13. 13.  
Zhang, J., Yang, J., Jang, R. & Zhang, Y. GPCR-I-TASSER: a hybrid approach to G protein-coupled receptor structure modeling and the application to the human genome. *Structure* **23**, 1538–1549 (2015).
14. 14.  
Bender, B. J., Marlow, B. & Meiler, J. Improving homology modeling from low-sequence identity templates in Rosetta: a case study in GPCRs. *PLOS Comput. Biol.* **16**, e1007597 (2020).
15. 15.  
Drew, K. et al. The Proteome Folding Project: proteome-scale prediction of structure and function. *Genome Res.* **21**, 1981–1994 (2011).
16. 16.  
Xu, D. & Zhang, Y. Ab initio structure prediction for *Escherichia coli*: towards genome-wide protein structure modeling and fold assignment. *Sci. Rep.* **3**, 1895 (2013).
17. 17.  
Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
18. 18.  
Sillitoe, I. et al. Genome3D: integrating a collaborative data pipeline to expand the depth and breadth of consensus protein structure annotation. *Nucleic Acids Res.* **48**, D314–D319 (2020).
19. 19.  
Pieper, U. et al. ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* **42**, D336–D346 (2014).
20. 20.  
Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).

21. 21.  
Kuhlman, B. & Bradley, P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* **20**, 681–697 (2019).
22. 22.  
The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
23. 23.  
Zhou, N. et al. The CAFA challenge reports improved protein function prediction and new functional annotations for hundreds of genes through experimental screens. *Genome Biol.* **20**, 244 (2019).
24. 24.  
Gligorijević, V. et al. Structure-based protein function prediction using graph convolutional networks. *Nat. Commun.* **12**, 3168 (2021).
25. 25.  
Necci, M., Piovesan, D. CAID Predictors, DisProt Curators & Tosatto, S. C. E. Critical assessment of protein intrinsic disorder prediction. *Nat. Methods* **18**, 472–481 (2021).
26. 26.  
Sillitoe, I. et al. CATH: expanding the horizons of structure-based functional annotations for genome sequences. *Nucleic Acids Res.* **47**, D280–D284 (2019).
27. 27.  
Andreeva, A., Kulesha, E., Gough, J. & Murzin, A. G. The SCOP database in 2020: expanded classification of representative family and superfamily domains of known protein structures. *Nucleic Acids Res.* **48**, D376–D382 (2020).
28. 28.  
Mistry, J. et al. Pfam: the protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
29. 29.

Kryshtafovych, A., Schwede, T., Topf, M., Fidelis, K. & Moult, J. Critical assessment of methods of protein structure prediction (CASP)-round XIII. *Proteins* **87**, 1011–1020 (2019).

30. 30.

Pereira, J. et al. High-accuracy protein structure prediction in CASP14. *Proteins* <https://doi.org/10.1002/prot.26171> (2021).

31. 31.

Senior, A. W. et al. Improved protein structure prediction using potentials from deep learning. *Nature* **577**, 706–710 (2020).

32. 32.

Zhang, Y. Protein structure prediction: when is it useful? *Curr. Opin. Struct. Biol.* **19**, 145–155 (2009).

33. 33.

Flower, T. G. & Hurley, J. H. Crystallographic molecular replacement using an in silico-generated search model of SARS-CoV-2 ORF8. *Protein Sci.* **30**, 728–734 (2021).

34. 34.

Egbert, M. et al. Functional assessment.  
[https://predictioncenter.org/casp14/doc/presentations/2020\\_12\\_03\\_Function\\_Assessment\\_VajdaLab\\_KozakovLab.pdf](https://predictioncenter.org/casp14/doc/presentations/2020_12_03_Function_Assessment_VajdaLab_KozakovLab.pdf) (2020).

35. 35.

Mariani, V., Biasini, M., Barbato, A. & Schwede, T. IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **29**, 2722–2728 (2013).

36. 36.

The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

37. 37.

The Gene Ontology Consortium. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).

38. 38.

Hopkins, A. L. & Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727–730 (2002).

39. 39.

Haas, J. et al. Introducing “best single template” models as reference baseline for the Continuous Automated Model Evaluation (CAMEO). *Proteins* **87**, 1378–1387 (2019).

40. 40.

Haas, J. et al. Continuous Automated Model Evaluation (CAMEO) complementing the critical assessment of structure prediction in CASP12. *Proteins* **86**, 387–398 (2018).

41. 41.

Schaeffer, R. D., Kinch, L. & Grishin, N. CASP14: InterDomain Performance. [https://predictioncenter.org/casp14/doc/presentations/2020\\_12\\_02\\_Interdomain\\_assessment1\\_Schaeffer.pdf](https://predictioncenter.org/casp14/doc/presentations/2020_12_02_Interdomain_assessment1_Schaeffer.pdf) (2020).

42. 42.

Zhang, Y. & Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins* **57**, 702–710 (2004).

43. 43.

Grinter, R. et al. Protease-associated import systems are widespread in Gram-negative bacteria. *PLoS Genet.* **15**, e1008435 (2019).

44. 44.

Pan, C.-J., Lei, K.-J., Annabi, B., Hemrika, W. & Chou, J. Y. Transmembrane topology of glucose-6-phosphatase. *J. Biol. Chem.* **273**, 6144–6148 (1998).

45. 45.

van Schaftingen, E. & Gerin, I. The glucose-6-phosphatase system. *Biochem. J.* **362**, 513–532 (2002).

46. 46.

Messerschmidt, A., Prade, L. & Wever, R. Implications for the catalytic mechanism of the vanadium-containing enzyme chloroperoxidase from the fungus *Curvularia inaequalis* by X-ray structures of the native and peroxide form. *Biol. Chem.* **378**, 309–315 (1997).

47. 47.

Amin, N. B. et al. Targeting diacylglycerol acyltransferase 2 for the treatment of nonalcoholic steatohepatitis. *Sci. Transl. Med.* **11**, eaav9701 (2019).

48. 48.

Futatsugi, K. et al. Discovery and optimization of imidazopyridine-based inhibitors of diacylglycerol acyltransferase 2 (DGAT2). *J. Med. Chem.* **58**, 7173–7185 (2015).

49. 49.

Birch, A. M. et al. Discovery of a potent, selective, and orally efficacious pyrimidinooxazinyl bicyclooctaneacetic acid diacylglycerol acyltransferase-1 inhibitor. *J. Med. Chem.* **52**, 1558–1568 (2009).

50. 50.

Cao, H. Structure-function analysis of diacylglycerol acyltransferase sequences from 70 organisms. *BMC Res. Notes* **4**, 249 (2011).

51. 51.

Wang, L. et al. Structure and mechanism of human diacylglycerol *O*-acyltransferase 1. *Nature* **581**, 329–332 (2020).

52. 52.

Stone, S. J., Levin, M. C. & Farese, R. V. Jr. Membrane topology and identification of key functional amino acid residues of murine acyl-CoA:diacylglycerol acyltransferase-2. *J. Biol. Chem.* **281**, 40273–40282 (2006).

53. 53.

Rigoli, L., Lombardo, F. & Di Bella, C. Wolfram syndrome and *WFS1* gene. *Clin. Genet.* **79**, 103–117 (2011).

54. 54.

Urano, F. Wolfram syndrome: diagnosis, management, and treatment. *Curr. Diab. Rep.* **16**, 6 (2016).

55. 55.

Schäffer, D. E., Iyer, L. M., Burroughs, A. M. & Aravind, L. Functional innovation in the evolution of the calcium-dependent system of the eukaryotic endoplasmic reticulum. *Front. Genet.* **11**, 34 (2020).

56. 56.

Guardino, K. M., Sheftic, S. R., Slattery, R. E. & Alexandrescu, A. T. Relative stabilities of conserved and non-conserved structures in the OB-fold superfamily. *Int. J. Mol. Sci.* **10**, 2412–2430 (2009).

57. 57.

Zhang, Y. & Skolnick, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **33**, 2302–2309 (2005).

58. 58.

Das, D. et al. The structure of KPN03535 (gi|152972051), a novel putative lipoprotein from *Klebsiella pneumoniae*, reveals an OB-fold. *Acta Crystallogr. F* **66**, 1254–1260 (2010).

59. 59.

Fass, D. & Thorpe, C. Chemistry and enzymology of disulfide cross-linking in proteins. *Chem. Rev.* **118**, 1169–1198 (2018).

60. 60.

Basile, W., Salvatore, M., Bassot, C. & Elofsson, A. Why do eukaryotic proteins contain more intrinsically disordered regions? *PLOS Comput. Biol.* **15**, e1007186 (2019).

61. 61.

Bhowmick, A. et al. Finding our way in the dark proteome. *J. Am. Chem. Soc.* **138**, 9730–9742 (2016).

62. 62.

Oates, M. E. et al. D<sup>2</sup>P<sup>2</sup>: database of disordered protein predictions. *Nucleic Acids Res.* **41**, D508–D516 (2013).

63. 63.

Hanson, J., Paliwal, K. K., Litfin, T. & Zhou, Y. SPOT-Disorder2: improved protein intrinsic disorder prediction by ensembled deep learning. *Genomics Proteomics Bioinformatics* **17**, 645–656 (2019).

64. 64.

Dunne, M., Ernst, P., Sobieraj, A., Pluckthun, A. & Loessner, M. J. The M23 peptidase domain of the *Staphylococcal* phage 2638A endolysin. <https://doi.org/10.2210/pdb6YJ1/pdb> (2020).

65. 65.

Krivák, R. & Hoksza, D. P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. *J. Cheminform.* **10**, 39 (2018).

66. 66.

Li, Y.-C. et al. Structure and noncanonical Cdk8 activation mechanism within an Argonaute-containing Mediator kinase module. *Sci. Adv.* **7**, eabd4484 (2021).

67. 67.

Eddy, S. R. A new generation of homology search tools based on probabilistic inference. *Genome Inform.* **23**, 205–211 (2009).

68. 68.

Steinegger, M., Mirdita, M. & Söding, J. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. *Nat. Methods* **16**, 603–606 (2019).

69. 69.

Schrödinger. The PyMOL Molecular Graphics System v.1.8 (2015).

70. 70.

Morris, G. M. et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).

71. 71.

Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).

72. 72.

Stein, P. E. et al. The crystal structure of pertussis toxin. *Structure* **2**, 45–57 (1994).

73. 73.

Necci, M., Piovesan, D., Clementel, D., Dosztányi, Z. & Tosatto, S. C. E. MobiDB-lite 3.0: fast consensus annotation of intrinsic disorder flavours in proteins. *Bioinformatics* **36**, 5533–5534 (2020).

74. 74.

Dyson, H. J. Roles of intrinsic disorder in protein–nucleic acid interactions. *Mol. Biosyst.* **8**, 97–104 (2012).

75. 75.

Dunbrack, R. L. Jr & Karplus, M. Backbone-dependent rotamer library for proteins. Application to side-chain prediction. *J. Mol. Biol.* **230**, 543–574 (1993).

## Acknowledgements

We thank A. Paterson, C. Low, C. Donner, D. Evans, F. Yang, J. Stanway, J. Stanton, L. Deason, N. Latysheva, N. Hobbs, R. Hadsell, R. Green, S. Brown, V. Bolina, Ž. Avsec and the Research Platform Team for their contributions; R. Kemp for help in managing the project and our colleagues at DeepMind, Google and Alphabet for their encouragement and support; E. van Schaftingen, M. Zhou and F. Urano for reading and commenting on our discussion of glucose-6-phosphatase, diacylglycerol *O*-acyltransferase 2 and wolframin, respectively; the team at EMBL-EBI for their work

on making AlphaFold structure predictions available, in particular M. Varadi, M. Deshpande, S. Sasidharan Nair, S. Anyango, G. Yordanova, C. Natassia, D. Yuan and E. Heard.

## Author information

### Author notes

1. These authors contributed equally: John Jumper, Demis Hassabis

### Affiliations

1. DeepMind, London, UK

Kathryn Tunyasuvunakool, Jonas Adler, Zachary Wu, Tim Green, Michal Zielinski, Augustin Žídek, Alex Bridgland, Andrew Cowie, Clemens Meyer, Agata Laydon, Richard Evans, Alexander Pritzel, Michael Figurnov, Olaf Ronneberger, Russ Bates, Simon A. A. Kohl, Anna Potapenko, Andrew J. Ballard, Bernardino Romera-Paredes, Stanislav Nikolov, Rishabh Jain, Ellen Clancy, David Reiman, Stig Petersen, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, John Jumper & Demis Hassabis

2. European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, UK

Sameer Velankar, Gerard J. Kleywegt, Alex Bateman & Ewan Birney

### Contributions

K.T., J.J. and D.H. led the research. D.H., K.K., P.K., C.M. and E.C. managed the research. T.G. developed the proteome-scale inference system. K.T., J.A., Z.W., M.Z., R.E., M.F., A. Bridgland and A.C. generated and analysed the structure predictions. J.J., M.F., S.A.A.K. and O.R. developed the metrics used to interpret predictions. A.Ž., S.P., T.G., A.C. and K.T. developed the data-processing pipelines to produce the AlphaFold protein structure database. S.V., A.L., A. Bateman, G.J.K., D.H. and E.B. managed the work to make AlphaFold predictions available via EMBL-EBI-hosted resources. S.V., G.J.K. and A. Bateman provided scientific advice on how predictions should be displayed. J.J., R.E., A. Pritzel, M.F., O.R., R.B., A. Potapenko, S.A.A.K., B.R.-P., J.A., A.W.S., T.G., A.Ž., K.T., A. Bridgland, A.J.B., A.C., S.N., R.J., D.R. and M.Z. developed the network and associated infrastructure used in inferencing the proteome. K.T., J.A., Z.W., J.J., M.F., M.Z., C.M. and D.H. wrote the paper.

## Corresponding authors

Correspondence to [Kathryn Tunyasuvunakool](#) or [John Jumper](#) or [Demis Hassabis](#).

## Ethics declarations

### Competing interests

J.J., R.E., A. Pritzel, T.G., M.F., O.R., R.B., A. Bridgland, S.A.A.K., D.R. and A.W.S. have filed non-provisional patent applications 16/701,070, PCT/EP2020/084238, and provisional patent applications 63/107,362, 63/118,917, 63/118,918, 63/118,921 and 63/118,919, each in the name of DeepMind Technologies Limited, each pending, relating to machine learning for predicting protein structures. E.B. is a paid consultant to Oxford Nanopore and Dovetail Inc, which are genomics companies. The other authors declare no competing interests.

## Additional information

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

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

## Extended data figures and tables

### [Extended Data Fig. 1 Example full chain outputs containing both high- and low-confidence regions.](#)

Q06787 (synaptic functional regulator FMR1) and P54725 (UV excision repair protein RAD23 homologue A) are predicted to be disordered outside the experimentally determined regions by MobiDB<sup>73</sup>. Q92664 (transcription factor IIIA) has been described as ‘beads on a string’, consisting of zinc-finger domains joined by flexible linkers<sup>74</sup>.

### [Extended Data Fig. 2 Distribution of per-residue IDDT-Cα within eight pLDDT bins.](#)

This represents an alternative visualization to Fig. 1a that does not sample the data. It uses the recent PDB dataset (Methods), which is restricted to structures with a reported

resolution of  $<3.5$  Å ( $n = 2,756,569$  residues). Residues were assigned to bins of width 10 based on their pLDDT (minimum, 20; maximum, 100). Markers show the mean IDDT-C $\alpha$  within each bin, while the IDDT-C $\alpha$  distribution is visualized as a Matplotlib violin plot (kernel density estimate bandwidth, 0.2). The smallest sample size for the corresponding violin is 5,655 residues for the left-most bin.

### [Extended Data Fig. 3 Relationship between pLDDT and side-chain \$\chi\_1\$ correctness.](#)

Evaluated on the recent PDB dataset (Methods), which is restricted to structures with a reported resolution of  $<2.5$  Å ( $n = 5,983$  chains) and residues with a  $B$ -factor of  $<30$  Å $^2$  ( $n = 609,623$  residues). Residues are binned by pLDDT, with bin width 5 between 20 and 70 pLDDT and bin width 2 above 70 pLDDT. A  $\chi_1$  angle is considered correct if it is within 40° of its value in the PDB structure<sup>75</sup>. Markers show the proportion of correct  $\chi_1$  angles within each bin; error bars indicate the 95% confidence interval (two-sided Student's  $t$ -test). The smallest sample size for the error bars is 193 residues for the left-most bin.

### [Extended Data Fig. 4 AlphaFold performance at a range of template sequence identities.](#)

IDDT-C $\alpha$  for AlphaFold and BestSingleStructuralTemplate on 1 year of CAMEO targets<sup>39</sup>. Targets are binned according to the sequence identity of the best template covering at least 70% of the target, and a box plot is shown for each bin. The horizontal line indicates the median, boxes range from the lower to the upper quartile, and the whiskers extend from the minimum to the maximum. In total, 428 targets are included (see Source Data); the smallest number of targets in any bin is 18

[Source data.](#)

### [Extended Data Fig. 5 Docking poses for a DGAT1-specific inhibitor in DGAT2.](#)

**a**, Top binding pose from Autodock Vina for a DGAT1-specific inhibitor in DGAT2, which does not match the predicted binding pocket for a DGAT2-specific inhibitor. **b**, Next best binding pose, which matches the binding pocket for the DGAT2-specific inhibitor, but does not contain components that satisfy the polar side chains His163 and Thr194. **c**, Relative positions of both binding poses.

## [Extended Data Fig. 6 Relationship between sequence length and inference time.](#)

On the basis of logs from our human proteome set. All of the processed proteins are shown ( $n = 20,296$ ). Each point indicates the mean inference time for the protein over the models produced. Vertical lines show the length cut-offs above which sequences were processed by multi-GPU workers.

## [Extended Data Fig. 7 Relationship between sequence length and run time for the non-inference stages of the pipeline.](#)

On the basis of 240 human protein sequences, chosen by stratified sampling from the length buckets: [16, 500), [500, 1,000), [1,000, 1,500), [1,500, 2,000), [2,000, 2,500) and [2,500, 2,700]. The relax plot shows five times more points, since five relaxed models are generated per protein. Coefficients for the quadratic lines of best fit were computed with Numpy polyfit.

**Extended Data Table 1 IDDT-C $\alpha$  distribution in various pLDDT bins**  
**Extended Data Table 2 Relationship between pLDDT and TM-score**

## **Supplementary information**

### [Supplementary Methods](#)

This file contains (1) Predicted TM-score weighting, and (2) DGAT docking scores.

### [Reporting Summary](#)

### [Supplementary Video 1](#)

Solvent accessibility of putative active site for G6Pase- $\alpha$ .

## **Source data**

### [Source Data Extended Data Fig. 4](#)

## **Rights and permissions**

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and

reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

### Reprints and Permissions

## About this article

### Cite this article

Tunyasuvunakool, K., Adler, J., Wu, Z. *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596 (2021).  
<https://doi.org/10.1038/s41586-021-03828-1>

- Received: 11 May 2021
- Accepted: 16 July 2021
- Published: 22 July 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03828-1>

## Further reading

- [DeepMind's AI predicts structures for a vast trove of proteins](#)
  - Ewen Callaway

*Nature* (2021)

## [Highly accurate protein structure prediction with AlphaFold](#)

- John Jumper
- Richard Evans
- Demis Hassabis

Article Open Access 15 Jul 2021

## Protein-structure prediction revolutionized

- Mohammed AlQuraishi

News & Views 23 Aug 2021

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03828-1>

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

- Article
- [Published: 18 August 2021](#)

# Molecular basis for DarT ADP-ribosylation of a DNA base

- [Marion Schuller](#)<sup>1</sup>,
- [Rachel E. Butler](#)<sup>2</sup>,
- [Antonio Ariza](#) [ORCID: orcid.org/0000-0003-4364-823X](#)<sup>1</sup>,
- [Callum Tromans-Coia](#)<sup>1</sup>,
- [Gytis Jankevicius](#)<sup>1,3</sup>,
- [Tim D. W. Claridge](#) [ORCID: orcid.org/0000-0001-5583-6460](#)<sup>4</sup>,
- [Sharon L. Kendall](#) [ORCID: orcid.org/0000-0003-3277-3035](#)<sup>5</sup>,
- [Shan Goh](#)<sup>5</sup>,
- [Graham R. Stewart](#) [ORCID: orcid.org/0000-0002-6867-6248](#)<sup>2</sup> &
- [Ivan Ahel](#) [ORCID: orcid.org/0000-0002-9446-3756](#)<sup>1</sup>

[Nature](#) volume **596**, pages 597–602 (2021)

- 3524 Accesses
- 68 Altmetric
- [Metrics details](#)

## Subjects

- [DNA damage and repair](#)
- [DNA repair enzymes](#)
- [Pathogens](#)
- [X-ray crystallography](#)

## Abstract

ADP-ribosyltransferases use NAD<sup>+</sup> to catalyse substrate ADP-ribosylation<sup>1</sup>, and thereby regulate cellular pathways or contribute to toxin-mediated pathogenicity of

bacteria<sup>2,3,4</sup>. Reversible ADP-ribosylation has traditionally been considered a protein-specific modification<sup>5</sup>, but recent in vitro studies have suggested nucleic acids as targets<sup>6,7,8,9</sup>. Here we present evidence that specific, reversible ADP-ribosylation of DNA on thymidine bases occurs in cellulo through the DarT–DarG toxin–antitoxin system, which is found in a variety of bacteria (including global pathogens such as *Mycobacterium tuberculosis*, enteropathogenic *Escherichia coli* and *Pseudomonas aeruginosa*)<sup>10</sup>. We report the structure of DarT, which identifies this protein as a diverged member of the PARP family. We provide a set of high-resolution structures of this enzyme in ligand-free and pre- and post-reaction states, which reveals a specialized mechanism of catalysis that includes a key active-site arginine that extends the canonical ADP-ribosyltransferase toolkit. Comparison with PARP–HPF1, a well-established DNA repair protein ADP-ribosylation complex, offers insights into how the DarT class of ADP-ribosyltransferases evolved into specific DNA-modifying enzymes. Together, our structural and mechanistic data provide details of this PARP family member and contribute to a fundamental understanding of the ADP-ribosylation of nucleic acids. We also show that thymine-linked ADP-ribose DNA adducts reversed by DarG antitoxin (functioning as a noncanonical DNA repair factor) are used not only for targeted DNA damage to induce toxicity, but also as a signalling strategy for cellular processes. Using *M. tuberculosis* as an exemplar, we show that DarT–DarG regulates growth by ADP-ribosylation of DNA at the origin of chromosome replication.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

**Access all Nature Journals**

Immediate access to more than 50 journals, including  
*Nature*

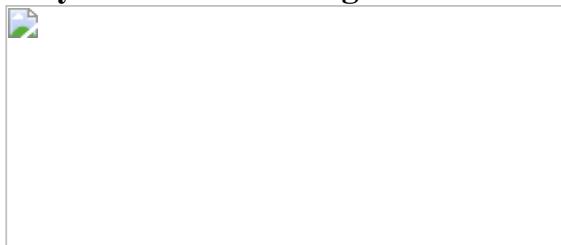
\$19.99/month

[Learn more](#)

**Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

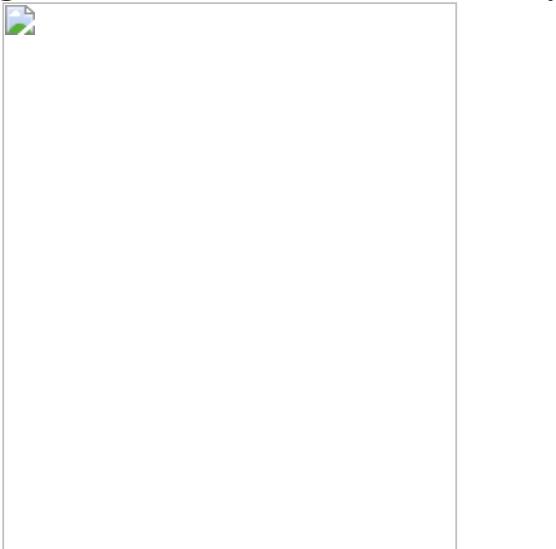
**Fig. 1: DarT is a PARP-like protein that catalyses ADP-ribosylation of DNA at the thymidine base nitrogen.**



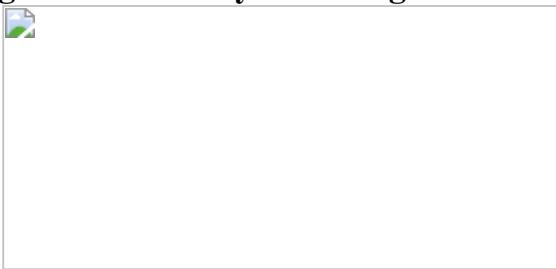
**Fig. 2: DNA binding is highly coordinated by DarT for site-specific ADP-ribosylation of DNA.**



**Fig. 3: Mechanism of DNA ADP-ribosylation.**



**Fig. 4: ADP-ribosylation of gDNA in *M. tuberculosis*.**



## Data availability

Crystallography atomic coordinates and structure factors are deposited in the PDB (<https://www.rcsb.org>) under the following accession codes: [7OMV](#), [7OMW](#), [7OMX](#), [7OMY](#), [7OMU](#), [7OMZ](#) and [7ON0](#). RNA-sequencing sequence files are deposited at the NCBI Gene Expression Omnibus GEO under the accession code [GSE174526](#). Tn-seq sequence files are deposited at the NCBI Sequence Read Archive, SRA accession number [PRJNA532518](#) run [SRR8886987](#). All data supporting the findings of this study are available within Article and any further information will be provided upon request to the corresponding authors.

## References

1. 1.  
Barkauskaite, E., Jankevicius, G. & Ahel, I. Structures and mechanisms of enzymes employed in the synthesis and degradation of PARP-dependent protein ADP-ribosylation. *Mol. Cell* **58**, 935–946 (2015).
2. 2.  
Gibson, B. A. & Kraus, W. L. New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nat. Rev. Mol. Cell Biol.* **13**, 411–424 (2012).
3. 3.  
Cohen, M. S. & Chang, P. Insights into the biogenesis, function, and regulation of ADP-ribosylation. *Nat. Chem. Biol.* **14**, 236–243 (2018).
4. 4.  
Perina, D. et al. Distribution of protein poly(ADP-ribosyl)ation systems across all domains of life. *DNA Repair* **23**, 4–16 (2014).
5. 5.  
Kraus, W. L. PARPs and ADP-ribosylation: 50 years ... and counting. *Mol. Cell* **58**, 902–910 (2015).
6. 6.  
Talhaoui, I. et al. Poly(ADP-ribose) polymerases covalently modify strand break termini in DNA fragments in vitro. *Nucleic Acids Res.* **44**, 9279–9295 (2016).
7. 7.  
Munnur, D. & Ahel, I. Reversible mono-ADP-ribosylation of DNA breaks. *FEBS J.* **284**, 4002–4016 (2017).
8. 8.  
Munnur, D. et al. Reversible ADP-ribosylation of RNA. *Nucleic Acids Res.* **47**, 5658–5669 (2019).
9. 9.

Groslambert, J., Prokhorova, E. & Ahel, I. ADP-ribosylation of DNA and RNA. *DNA Repair (Amst.)* **105**, 103144 (2021).

10. 10.

Jankevicius, G., Ariza, A., Ahel, M. & Ahel, I. The toxin-antitoxin system DarTG catalyzes reversible ADP-ribosylation of DNA. *Mol. Cell* **64**, 1109–1116 (2016).

11. 11.

Aravind, L., Zhang, D., de Souza, R. F., Anand, S. & Iyer, L. M. in *Endogenous ADP-ribosylation* (ed. Koch-Nolte, F.) 3–32 (Springer International, 2014).

12. 12.

Bell, C. E. & Eisenberg, D. Crystal structure of diphtheria toxin bound to nicotinamide adenine dinucleotide. *Adv. Exp. Med. Biol.* **419**, 35–43 (1997).

13. 13.

Steffen, J. D., Brody, J. R., Armen, R. S. & Pascal, J. M. Structural implications for selective targeting of PARPs. *Front. Oncol.* **3**, 301 (2013).

14. 14.

Vyas, S. et al. Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nat. Commun.* **5**, 4426 (2014).

15. 15.

Hottiger, M. O., Hassa, P. O., Lüscher, B., Schüler, H. & Koch-Nolte, F. Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem. Sci.* **35**, 208–219 (2010).

16. 16.

Belousova, E. A., Ishchenko, A. A. & Lavrik, O. I. DNA is a new target of Parp3. *Sci. Rep.* **8**, 4176 (2018).

17. 17.

Zarkovic, G. et al. Characterization of DNA ADP-ribosyltransferase activities of PARP2 and PARP3: new insights into DNA ADP-ribosylation. *Nucleic Acids*

*Res.* **46**, 2417–2431 (2018).

18. 18.

Ueda, K. & Hayaishi, O. ADP-ribosylation. *Annu. Rev. Biochem.* **54**, 73–100 (1985).

19. 19.

Lawarée, E. et al. DNA ADP-ribosylation stalls replication and is reversed by RecF-mediated homologous recombination and nucleotide excision repair. *Cell Rep.* **30**, 1373–1384 (2020).

20. 20.

Mendum, T. A. et al. Transposon libraries identify novel *Mycobacterium bovis* BCG genes involved in the dynamic interactions required for BCG to persist during in vivo passage in cattle. *BMC Genomics* **20**, 431 (2019).

21. 21.

Ahel, I. et al. The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates. *Nature* **443**, 713–716 (2006).

22. 22.

Tumbale, P. et al. Structure of an aprataxin–DNA complex with insights into AOA1 neurodegenerative disease. *Nat. Struct. Mol. Biol.* **18**, 1189–1195 (2011).

23. 23.

Harms, A., Brodersen, D. E., Mitarai, N. & Gerdes, K. Toxins, targets, and triggers: an overview of toxin-antitoxin biology. *Mol. Cell* **70**, 768–784 (2018).

24. 24.

Gamulin, V., Cetkovic, H. & Ahel, I. Identification of a promoter motif regulating the major DNA damage response mechanism of *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* **238**, 57–63 (2004).

25. 25.

Suskiewicz, M. J. et al. HPF1 completes the PARP active site for DNA damage-induced ADP-ribosylation. *Nature* **579**, 598–602 (2020).

26. 26.

Singh, A. K. et al. Investigating essential gene function in *Mycobacterium tuberculosis* using an efficient CRISPR interference system. *Nucleic Acids Res.* **44**, e143 (2016).

27. 27.

Larson, M. H. et al. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protocols* **8**, 2180–2196 (2013).

28. 28.

St-Pierre, F. et al. One-step cloning and chromosomal integration of DNA. *ACS Synth. Biol.* **2**, 537–541 (2013).

29. 29.

Thu, K. L. et al. Methylated DNA immunoprecipitation. *J. Vis. Exp.* **935**, 935 (2009).

30. 30.

Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).

31. 31.

Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D* **60**, 432–438 (2004).

32. 32.

Cowtan, K. Recent developments in classical density modification. *Acta Crystallogr. D* **66**, 470–478 (2010).

33. 33.

Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62**, 1002–1011 (2006).

34. 34.

Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

35. 35.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).

36. 36.

Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

37. 37.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

38. 38.

Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2019).

39. 39.

Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).

40. 40.

Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).

41. 41.

Tsurumura, T. et al. Arginine ADP-ribosylation mechanism based on structural snapshots of iota-toxin and actin complex. *Proc. Natl Acad. Sci. USA* **110**, 4267–4272 (2013).

## Acknowledgements

We thank C. Tang for guidance and discussion of this study and sharing reagents; J. Rack and M. Suskiewicz for discussions and critically revising the manuscript; M. Suskiewicz and E. Lowe for help with X-ray diffraction data collection and processing; S. Meyer for help with biochemical experiments; J. Elkins for discussions; J. Holder and T. Agnew for proofreading of the manuscript; T. Mendum, H. Wu and A. Smith for help with transposon mutagenesis; S. Hingley-Wilson for discussions on toxin–antitoxin systems; and Diamond Light Source for access to and assistance at beamlines I03, I04, I04-1, and I24 throughout the project (proposal numbers mx18069). Work in the laboratory of I.A. is supported by the Wellcome Trust (101794 and 210634), Biotechnology and Biological Sciences Research Council (BB/R007195/1) and Cancer Research United Kingdom (C35050/A22284). The G.R.S. and S.L.K. laboratories were supported by Biotechnology and Biological Sciences Research Council grants BB/R006393/1 and BB/N004590/1, respectively.

## Author information

### Affiliations

1. Sir William Dunn School of Pathology, University of Oxford, Oxford, UK

Marion Schuller, Antonio Ariza, Callum Tromans-Coia, Gytis Jankevicius & Ivan Ahel

2. Department of Microbial Sciences, School of Biosciences and Medicine, University of Surrey, Guildford, UK

Rachel E. Butler & Graham R. Stewart

3. Biozentrum, University of Basel, Basel, Switzerland

Gytis Jankevicius

4. Department of Chemistry, Chemistry Research Laboratory, University of Oxford, Oxford, UK

Tim D. W. Claridge

5. Centre for Emerging, Endemic and Exotic Disease, Pathology and Population Sciences, The Royal Veterinary College, Hatfield, UK

Sharon L. Kendall & Shan Goh

## Contributions

I.A. and G.R.S. conceived the project and conceptualized experiments with input from M.S., R.E.B. and G.J. M.S. conducted biochemical and crystallographic studies, including structure and data analysis and interpretation, with assistance of other authors; A.A. solved *T. africanus* DarTG structure and refined structural data; C.T.-C. and G.J. established method for detection of ADP-ribosylated DNA and supported strain construction; T.D.W.C. conducted NMR experiments and analysis; R.E.B. and G.R.S. performed mycobacteria experiments with assistance from S.L.K. and S.G. for DarG knockdown. M.S., I.A., G.R.S. and R.E.B. wrote the manuscript with support of all other authors.

## Corresponding authors

Correspondence to [Graham R. Stewart](#) or [Ivan Ahel](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

### [Extended Data Fig. 1 DarT structure reveals a PARP-like ADP-ribosyltransferase.](#)

Related to Fig. 1. a, Crystal structure of *T. africanus* DarTG(E152A) fusion protein served as model for solving DarT of *Thermus* sp. 2.9 by molecular replacement. The

fused DarG macrodomain is coloured in light orange with the bound ADP-ribose molecule shown as an atom-coloured stick model in black. **b**, **c**, Comparison of DarT with eukaryotic ARTD (that is, PARP) and bacterial ARTD fold. **b**, Secondary structure analysis shows the close similarity of DarT to ARTD family members, and PARPs in particular, with its fold-stabilizing central 6-stranded  $\beta$ -sheet core and the ARTD-conserved helices between strands  $\beta$ 1 and  $\beta$ 2 ( $\beta$ 1-2) and  $\beta$ 2 and  $\beta$ 3 ( $\beta$ 2-3). The crystal structure of *Thermus* sp. 2.9 DarT(E160A) in ligand-free state (apo) was overlaid with diphtheria toxin (PDB code 1TOX), PARP1 (PDB code 6BHV) and PARP13 (PDB code 2X5Y). For clarity, only central secondary structure elements showing the similarity between the folds are depicted from diphtheria toxin, PARP1 and PARP13. **c**, The N-terminal extension of the  $\beta$ -sheet core found in PARPs (that is, a strand–helix–strand arrangement next to  $\beta$ 6) is spatially replaced in DarT with a shorter C-terminal helix–strand extension. The crystal structure of *Thermus* sp. 2.9 DarT(E160A) in ADPr–DNA-bound state was overlaid with the crystal structure of PARP1 (PDB code 6BHV) (left) and PARP13 (PDB code 2X5Y) (right). For clarity, only the secondary structure elements showing this difference between the folds are depicted from PARP1 and PARP13. **d**, *Thermus* sp. 2.9 DarT(E160A) in complex with NAD<sup>+</sup> and carba-NAD<sup>+</sup>. Overlay of the crystal structures with *Thermus* sp. 2.9 DarT(E160A) in ligand-free state (apo) is shown on the right. Unresolved regions of the NAD<sup>+</sup>-binding loop-helical element including the ART donor loop (purple) in the DarT–NAD<sup>+</sup> and DarT–carba-NAD<sup>+</sup> structures are marked with asterisks.

## Extended Data Fig. 2 Structural details of ADPr–DNA bound to DarT.

Related to Fig. 1c, d. **a**, Crystal structure of *Thermus* sp. 2.9 DarT(E160A) in complex with ADPr–DNA. Co-crystallization with ADP-ribosylated DNA 5-mer after in vitro modification by *T. aquaticus* DarT and purification. The substrate-binding ARTT loop is highlighted in green, the NAD<sup>+</sup>-binding loop–helix element in purple (set, for clarity, in higher transparency). The catalytic glutamate E160, which is conserved in ARTs, is modelled as red sticks. **b**, Overlay of the ADP-ribosylated DNA products of the ADPr–DNA- and NAM-bound and the ADPr–DNA-bound DarT structures shows their perfect overlap, apart from a slight tilting of the bonds that connect the NAM-ribose with the  $\beta$ -phosphate. **c**, **d**, The ADPr–DNA ligands are highly resolved in the *Thermus* sp. 2.9 DarT(E160A) co-crystal structures, revealing the ADP-ribose linkage to the thymidine base nitrogen N3 in  $\alpha$  configuration. The  $2F_o - F_c$  electron density maps contoured at  $1.0\sigma$  around the ligands are shown in blue. ADPr, linked ADP-ribose; NAM, nicotinamide; ssDNA, single-stranded DNA. **c**, The ADPr–DNA ligand in the ADPr–DNA co-crystal structure of 1.46 Å resolution. **d**, The ADPr–DNA ligand in the ADPr–DNA- and NAM-bound co-crystal structure at 1.66 Å resolution. The NAM ligand left in the protein after ADP-ribosylation of the DNA is also clearly resolved. **e**, ADP-

ribosylation activity of the *T. aquaticus* DarT(E160A) mutant can also be observed in in vitro assays at low DNA (50 nM) and high protein concentrations under long incubation times. Modification of the Cy3-labelled oligonucleotide (DarT-ADPr-27mer-Cy3) was visualized after separation of the reaction products on denaturing polyacrylamide gel. Representative of two independent experiments.

### **Extended Data Fig. 3 Structural features of DarT for ssDNA binding and catalysis.**

Related to Figs. 1–3. **a**, The ADP-ribosylating turn–turn (ARTT) loop of *Thermus* sp. 2.9 DarT(E160A) in the ADPr–DNA- and NAM-bound structure is shown in green with its stabilized DNA substrate in magenta. Several loops form together with the few short  $\alpha$ -helices a stable scaffold, which is held in position by a network of over 100 interactions between main chains, side chains and water molecules. Left, cartoon representation; middle, the atom-coloured stick model of the ARTT loop. Interactions are indicated with grey dashes and water molecules as red spheres. Right, table comparing the ARTT loop length of DarT with other human and bacterial ARTDs<sup>15</sup>. **b**, *Thermus* sp. 2.9 DarT preferentially modifies a TNTC motif in ssDNA, which was verified by testing permutations of the motif. In vitro ADP-ribosylation activity of *Thermus* sp. 2.9 DarT was assessed by visualizing the modification of the oligonucleotides under UV light after separation and ethidium bromide staining of the reaction products on denaturing polyacrylamide gel. Representative of three independent experiments. **c**, Close views on the nucleotide recognition of DarT, rationalizing its preferred modification of DNA over RNA. Cartoon–stick models of the *Thermus* sp. 2.9 DarT(E160A) structure in the ADPr–DNA- and NAM-bound state are shown. Left, middle, additional 2' hydroxyl groups as in RNA strands may lead to clashes with parts of the proteins (that is, W147 (1st nucleotide) and the  $\alpha$ -helix between  $\beta$ 2 and  $\beta$ 3 (2nd nucleotide)). Right, as shown in previous studies<sup>10</sup>, the methyl group on the modified thymine (circle) increases thymidine base modification, probably by locking the base in optimal conformation for the ADP-ribosylation reaction. Interactions are indicated with grey dashes and water molecules as red spheres. **d**, Modelling of possible rotamers of glutamate E160 into the *Thermus* sp. 2.9 DarT apo structure. Several conformations of the glutamate would allow a proton transfer from arginine R51 (green) to glutamate E160 (red). Possible interactions are shown with dashes in magenta. **e**, Arginine R51 flexibility observed among the *Thermus* sp. 2.9 DarT apo, substrate- and product-bound states. The NAD<sup>+</sup>- and DNA-bound and as ADP-ribose unlinked state (fourth imagine from left) is modelled by superimposing the NAD<sup>+</sup> molecule with the NAD<sup>+</sup> co-crystal structure onto the carba-NAD<sup>+</sup>- and DNA-bound structure. R51 and ligands are shown as atom-coloured stick models, with R51 in green, NAD<sup>+</sup> in cyan, carba-

$\text{NAD}^+$  in brown and higher transparency and DNA (thymine only) in magenta. Interactions are indicated with grey dashes.

### Extended Data Fig. 4 DarT sequence alignments.

Related to Figs. 2c, d, 3. **a**, Sequence alignment of *Thermus* sp. 2.9 DarT with DarT of *T. aquaticus*. Numbers on top of the alignments refer to *Thermus* sp. 2.9 DarT. Table provides a residue identifier comparison for functional relevant residues. **b**, Multiple sequence alignment of DarT sequences, representing five main phylogenetically diverging branches. Numbers on top of the residues refer to *Thermus* sp. 2.9 DarT. Active site residues are highlighted in green, DNA-binding residues in magenta, with functionally similar residues as the reference in lower opacity. Shared sequence identities compared to *Thermus* sp. 2.9 DarT: group 1: 60%, group 2: 40–45%, group 3 and 4: 31–38%, group 5: 20–27%.

### Extended Data Fig. 5 $\text{NAD}^\pm$ coordination in the active site of DarT.

Related to Fig. 3, Supplementary Discussion. **a**, Comparison of the  $\text{NAD}^+$ -binding sites in the carba- $\text{NAD}^+$ -bound structure and the  $\text{NAD}^+$ -bound structure of *Thermus* sp. 2.9 DarT(E160A). Top, Overlay of the carba- $\text{NAD}^+$ - and DNA-bound structure (grey) with the carba- $\text{NAD}^+$ -bound structure (brown), of which just the ligand and the side chains are shown as an atom-coloured stick model. The carba- $\text{NAD}^+$  ligands of both structures perfectly overlay and DarT-interacting side chains show same positioning. DNA binding does not induce conformational changes upon the  $\text{NAD}^+$  ligand. Bottom, overlay of the carba- $\text{NAD}^+$ - and DNA-bound structure (grey) with the  $\text{NAD}^+$ -bound structure (cyan) shows slight differences in the ligand and the side chains positioning around the pyrophosphate-ribose moiety of the  $\text{NAD}^+$  molecule, which needs to be considered for analysis of  $\text{NAD}^+$  polarization. **b**, Molecular structures of  $\text{NAD}^+$  and carba- $\text{NAD}^+$ . **c**, Cartoon–stick model showing the coordination of the NAM side (left) and the adenine side (right) of the carba- $\text{NAD}^+$  ligand in the *Thermus* sp. 2.9 DarT(E160A) carba- $\text{NAD}^+$ - and DNA-bound structure with side- and main-chain interactions (dashed lines), including water (red spheres) contacts. **d**, Integrated thermogram obtained by isothermal titration calorimetry giving  $\text{NAD}^+$ -binding parameters for *Thermus* sp. 2.9 DarT(E160Q). A representative result from three independent experiments is shown, with the number of binding sites ( $n$ ) and the dissociation constant ( $K_D$ ) calculated from the repeats with mean  $\pm$  s.d. See Supplementary Fig. 3 for raw titration curves and additional thermodynamic parameters. **e**, Autoradiography of thin layer chromatography plate analysing the reaction products after incubation of *T. aquaticus* DarT and

DarT(E160A) with NAD<sup>+</sup> and DNA. NADase from pig brain was used as control for monitoring NADase activity. Representative of three independent experiments.

### Extended Data Fig. 6 Visualization of the ADP-ribosylation activity of DarT in cells.

Related to Fig. 4. **a**, Validation of the antibody identified for detection of ADP-ribosylated DNA. ADP-ribosylation of the oligonucleotide by *T. aquaticus* DarT was verified by analysis of the reaction product on denaturing polyacrylamide gel (top) and visualized by immunoblotting using the poly/mono-ADP ribose antibody, E6F6A (Cell Signaling Technology) (bottom). Immunodetection of ssDNA (using the DSHB autoanti-ssDNA antibody) served as loading control (middle).

Representative result of four independent experiments with three individually purified *T. aquaticus* DarT-ADP-ribosylated oligonucleotides. **b**, Dot blot showing DNA ADP-ribosylation activity by *T. aquaticus* and EPEC wild-type DarT and mutants on gDNA, its physiological target, (row 1 and 2 from top), which consequently induces DNA damage (RecA marker) in cells. EPEC DarT(G49D) is a characterized DarT mutant that retains ssDNA ADP-ribosylation activity, albeit to a lesser extent than the wild-type protein, and EPEC DarT(E170) is its respective catalytically inactive mutant<sup>19</sup>. See **c** for the structural basis of this lower activity of EPEC DarT(G49D). **c**, G49D mutation in EPEC DarT reduces ssDNA ADP-ribosylation activity. Overlay of a homology model of EPEC DarT with the structure of *Thermus* sp. 2.9 DarT(E160A) in complex with ADPr-DNA indicates that the EPEC DarT mutation G49D reduces DarT ssDNA activity owing to an aspartate side chain pointing into the NAD<sup>+</sup>-binding site towards the second phosphate group. This may sterically, but also owing to its negative charge, impair NAD<sup>+</sup> binding, resulting in a less efficient ADP-ribosylation reaction. **d**, DNA ADP-ribosylation by *T. aquaticus* DarT (dot blot, row 1 and 2 from top) and induction of DNA damage (RecA marker) is suppressed by *T. aquaticus* DarG with its macrodomain (MD), including by DarG macrodomains from non-cognate species (EPEC and *M. tuberculosis*). **e**, Dot blot showing ADP-ribose removal from *T. aquaticus* DarT-ADP-ribosylated gDNA by *T. aquaticus* DarG antitoxin with its macrodomain, and macrodomains from non-cognate species (EPEC and *M. tuberculosis*) in contrast to the human hydrolases MacroD1, PARG and ARH3. In **b**, **d**, **e**, cell lysates were prepared and gDNA was purified from samples before (+ glucose) and after (+ arabinose/IPTG) induction of protein expression and subjected to immunodetection. EV, empty vector. The N22A/K80A double mutation in *T. aquaticus* DarG results in loss of catalytic activity of the macrodomain. For gel source data, see Supplementary Fig. 1. Results are representative for three biologically independent experiments.

## Extended Data Fig. 7 Characterization of DarT gDNA ADP-ribosylation in *M. tuberculosis*.

Related to Fig. 4. **a**, Unregulated DarT activity (*darG* silencing) and induction of DNA damage (MMC treatment) led to a marked DNA damage response and induces expression of *dnaB-darT*. Gene transcription was compared by qPCR with reverse transcription of *M. bovis* BCG *darG* sgRNA un-induced, ATC-induced and MMC-treated samples. Data are mean  $\pm$  s.d. of three biologically independent replicates. **b**, Knockdown of *darG* expression in *M. tuberculosis* induces expression of RecA. *Mycobacterium tuberculosis* was treated with 200 ng ml<sup>-1</sup> ATC to induce dCas9 and *darG* sgRNA or non-targeting control sgRNA for 48 h or with MMC for 24 h. Cell-free bacterial lysates were probed by western blotting with an anti-RecA antiserum or anti-Hsp70 (DnaK) antibody as loading control. Representative of two biologically independent experiments. **c**, *darT* and *darG* are transcriptionally linked to *dnaB*. PCR products were generated with the indicated set of primers (Supplementary Table 2) and visualized by gel electrophoresis. The presence of PCR products across the *dnaB-darT* and *darT-darG* junctions demonstrates the transcriptional linkage of *dnaB*, *darT* and *darG* as a polycistronic mRNA. Representative of three independent experiments. **d**, *Mycobacterium tuberculosis* DarT preferentially modifies a TTTW motif in ssDNA. Screening of 40 ssDNA oligonucleotide sequences with potential four-base motifs for ADP-ribosylation by DarT (data not shown) identified TTTW as targeted sequence, which was verified by testing permutations of the TTTT motif. In vitro ADP-ribosylation activity of *M. tuberculosis* DarT was assessed by visualizing the modification of the oligonucleotides under UV light after separation and ethidium bromide staining of the reaction products on denaturing polyacrylamide gel. Representative of three independent experiments. **e**, *Mycobacterium tuberculosis* DarT ADP-ribosylates the OriC in vitro with preference for the lower strand at the TTTW motifs. ADP-ribosylation activity was assessed by visualizing the modification of the oligonucleotides under UV light after separation and ethidium bromide staining of the reaction products on denaturing polyacrylamide gel. Representative of three independent experiments. For gel source data, see Supplementary Fig. 1.

## Extended Data Fig. 8 DarT is a PARP-like protein that evolved novel features that allow its specialized function as a DNA ADP-ribosyltransferase.

Related to ‘Discussion’ in the Article. **a**, Schematics of the interactions between the NAD<sup>+</sup> substrate and the residues of the class-defining (H-Y-E) motif in ARTD members, including PARPs, compared to DarT. Conserved motif residues (purple)

and additional active site residues (green) essential for catalysis with their relative position to the NAD<sup>+</sup> substrate are compared. **b**, ARTs seem to share the spatial position and orientation of mechanistically relevant residues. Overlay of crystal structures of *Thermus* sp. 2.9 DarT(E160A) in ADPr-DNA- and NAM-bound state with *Clostridium perfringens* iota-toxin (Ia)-actin complex (left) (PDB code 4H0T) and PARP2 in the PARP2-HPF1 complex (right) (PDB code 6TX3). H119 in DarT takes spatially the same position as Y375 in the iota-toxin, which was suggested to have a role in target protein (that is, actin) recognition<sup>41</sup>. Both Y375 and H119 are accommodated in the ARTT loops, which do not show any similarity in either residue length or structural make up. The approximate position of DarT H119 is occupied by E284 of HPF1 in the PARP2-HPF1 complex, in which HPF1 sits on the ARTT loop of PARP2. This leads to the formation of a composite active site with the catalytic glutamate residues E284 and E545 for catalysing serine ADP-ribosylation<sup>25</sup>. Enlarged views of the active sites are below the respective cartoon models. For clarity in the enlarged views, only the ARTT loop from iota-toxin and PARP2 and only a fragment of the respective binding partner (that is, actin and HPF1) are shown as cartoon model. The substrate-coordinating and catalytic residues as well as the ADP-ribose products and complex-bound ligands are shown as stick models. Root mean square deviation (r.m.s.d.) of DarT-iota-toxin overlay, 2.71 Å; r.m.s.d. of DarT-PARP2 overlay, 2.58 Å.

**Extended Data Table 1 Data collection and refinement statistics for crystal structures of *T. a fricanus* DarTG and *Thermus* sp. 2.9 DarT described in this Article**

**Extended Data Table 2 Data collection and refinement statistics for DNA co-crystal structures of *Thermus* sp. 2.9 DarT described in this Article**

## Supplementary information

### Supplementary Information

This file contains Supplementary Figs 1-4, Supplementary Tables 1-3 and Supplementary References.

### Reporting Summary

### Supplementary Data

RNA-Seq Differential Expression Analysis.

## Peer Review File

## Rights and permissions

### Reprints and Permissions

## About this article

### Cite this article

Schuller, M., Butler, R.E., Ariza, A. *et al.* Molecular basis for DarT ADP-ribosylation of a DNA base. *Nature* **596**, 597–602 (2021).  
<https://doi.org/10.1038/s41586-021-03825-4>

- Received: 29 October 2020
- Accepted: 15 July 2021
- Published: 18 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03825-4>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03825-4>

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

- Article
- [Published: 11 August 2021](#)

# Cryo-EM structures of full-length *Tetrahymena* ribozyme at 3.1 Å resolution

- [Zhaoming Su](#) ORCID: [orcid.org/0000-0002-9279-1721](#)<sup>1,2 na1</sup>,
- [Kaiming Zhang](#)<sup>2,3 na1</sup>,
- [Kalli Kappel](#) ORCID: [orcid.org/0000-0002-2129-199X](#)<sup>4 na1</sup>,
- [Shanshan Li](#) ORCID: [orcid.org/0000-0002-7041-5960](#)<sup>2,3</sup>,
- [Michael Z. Palo](#)<sup>4</sup>,
- [Grigore D. Pintilie](#)<sup>2</sup>,
- [Ramya Rangan](#) ORCID: [orcid.org/0000-0002-0960-0825](#)<sup>4</sup>,
- [Bingnan Luo](#)<sup>1</sup>,
- [Yuquan Wei](#)<sup>1</sup>,
- [Rhiju Das](#) ORCID: [orcid.org/0000-0001-7497-0972](#)<sup>4,5</sup> &
- [Wah Chiu](#) ORCID: [orcid.org/0000-0002-8910-3078](#)<sup>2,6</sup>

[Nature](#) volume 596, pages 603–607 (2021)

- 5484 Accesses
- 1 Citations
- 114 Altmetric
- [Metrics details](#)

## Subjects

- [Cryoelectron microscopy](#)
- [Molecular conformation](#)
- [RNA](#)
- [RNA folding](#)

## Abstract

Single-particle cryogenic electron microscopy (cryo-EM) has become a standard technique for determining protein structures at atomic resolution<sup>1,2,3</sup>. However, cryo-EM studies of protein-free RNA are in their early days. The *Tetrahymena thermophila* group I self-splicing intron was the first ribozyme to be discovered and has been a prominent model system for the study of RNA catalysis and structure–function relationships<sup>4</sup>, but its full structure remains unknown. Here we report cryo-EM structures of the full-length *Tetrahymena* ribozyme in substrate-free and bound states at a resolution of 3.1 Å. Newly resolved peripheral regions form two coaxially stacked helices; these are interconnected by two kissing loop pseudoknots that wrap around the catalytic core and include two previously unforeseen (to our knowledge) tertiary interactions. The global architecture is nearly identical in both states; only the internal guide sequence and guanosine binding site undergo a large conformational change and a localized shift, respectively, upon binding of RNA substrates. These results provide a long-sought structural view of a paradigmatic RNA enzyme and signal a new era for the cryo-EM-based study of structure–function relationships in ribozymes.

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

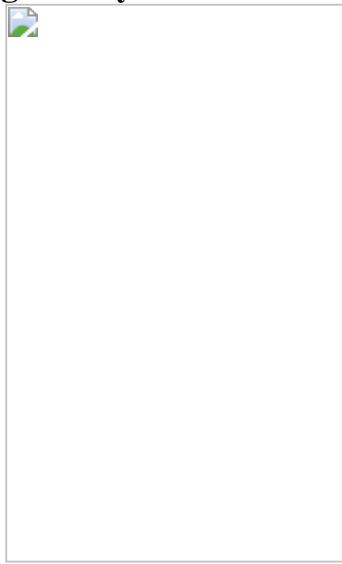
\$19.99/month

[Learn more](#)

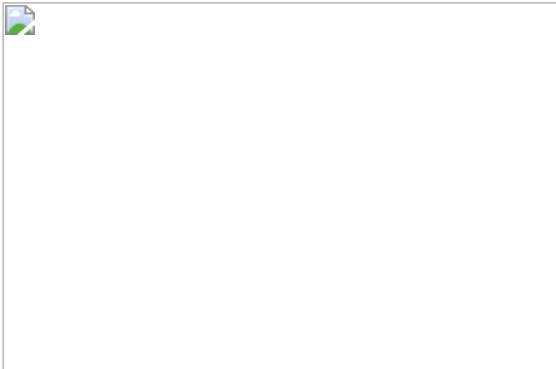
**Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

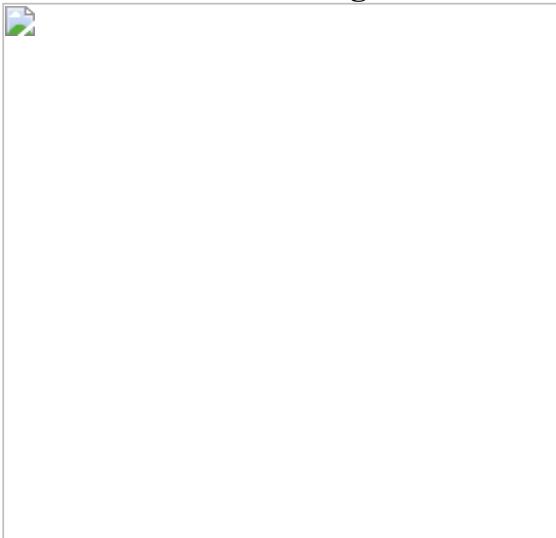
**Fig. 1: Cryo-EM reconstruction of apo L-21 ScaI ribozyme.**



**Fig. 2: Structural insights in the peripheral regions.**



**Fig. 3: Cryo-EM structure of holo L-16 ScaI ribozyme reveals docked P1–P10 and conformational change of IGS.**



**Fig. 4: Localized shifts and mechanistic insights.**

## Data availability

The cryo-EM maps and associated atomic coordinate models of the apo L-21 and holo L-16 ScaI ribozymes have been deposited in the wwPDB OneDep System under EMD accession codes [EMD-31385](#) and [EMD-31386](#) and PDB ID codes [7EZ0](#) and [7EZ2](#), respectively.

## References

1. 1.

Nakane, T. et al. Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).

2. 2.

Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).

3. 3.

Zhang, K., Pintilie, G. D., Li, S., Schmid, M. F. & Chiu, W. Resolving individual atoms of protein complex by cryo-electron microscopy. *Cell Res.* **30**, 1136–1139 (2020).

4. 4.

Kruger, K. et al. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**, 147–157 (1982).

5. 5.

Cech, T. R. The RNA worlds in context. *Cold Spring Harb. Perspect. Biol.* **4**, a006742 (2012).

6. 6.

Cech, T. R. & Steitz, J. A. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* **157**, 77–94 (2014).

7. 7.

Zhang, H. & Keane, S. C. Advances that facilitate the study of large RNA structure and dynamics by nuclear magnetic resonance spectroscopy. *Wiley Interdiscip. Rev. RNA* **10**, e1541 (2019).

8. 8.

Zhang, J. & Ferré-D’Amaré, A. R. New molecular engineering approaches for crystallographic studies of large RNAs. *Curr. Opin. Struct. Biol.* **26**, 9–15 (2014).

9. 9.

Zhang, K. et al. Cryo-EM structure of a 40 kDa SAM-IV riboswitch RNA at 3.7 Å resolution. *Nat. Commun.* **10**, 5511 (2019).

10. 10.

Kappel, K. et al. Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **17**, 699–707 (2020).

11. 11.

Cech, T. R. Ribozymes, the first 20 years. *Biochem. Soc. Trans.* **30**, 1162–1166 (2002).

12. 12.

Cate, J. H. et al. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273**, 1678–1685 (1996).

13. 13.

Golden, B. L., Gooding, A. R., Podell, E. R. & Cech, T. R. A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science* **282**, 259–264 (1998).

14. 14.

Guo, F., Gooding, A. R. & Cech, T. R. Structure of the *Tetrahymena* ribozyme: base triple sandwich and metal ion at the active site. *Mol. Cell* **16**, 351–362 (2004).

15. 15.

Juneau, K., Podell, E., Harrington, D. J. & Cech, T. R. Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA–solvent interactions. *Structure* **9**, 221–231 (2001).

16. 16.

Piccirilli, J. A., Vyle, J. S., Caruthers, M. H. & Cech, T. R. Metal ion catalysis in the *Tetrahymena* ribozyme reaction. *Nature* **361**, 85–88 (1993).

17. 17.

Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. & Strobel, S. A. Crystal structure of a self-splicing group I intron with both exons. *Nature* **430**, 45–50 (2004).

18. 18.

Golden, B. L., Kim, H. & Chase, E. Crystal structure of a phage Twort group I ribozyme-product complex. *Nat. Struct. Mol. Biol.* **12**, 82–89 (2005).

19. 19.

Shan, S., Yoshida, A., Sun, S., Piccirilli, J. A. & Herschlag, D. Three metal ions at the active site of the *Tetrahymena* group I ribozyme. *Proc. Natl Acad. Sci. USA* **96**, 12299–12304 (1999).

20. 20.

Stahley, M. R. & Strobel, S. A. Structural evidence for a two-metal-ion mechanism of group I intron splicing. *Science* **309**, 1587–1590 (2005).

21. 21.

Steitz, T. A. & Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl Acad. Sci. USA* **90**, 6498–6502 (1993).

22. 22.

Forconi, M., Piccirilli, J. A. & Herschlag, D. Modulation of individual steps in group I intron catalysis by a peripheral metal ion. *RNA* **13**, 1656–1667 (2007).

23. 23.

Beaudry, A. A. & Joyce, G. F. Minimum secondary structure requirements for catalytic activity of a self-splicing group I intron. *Biochemistry* **29**, 6534–6539 (1990).

24. 24.

Benz-Moy, T. L. & Herschlag, D. Structure-function analysis from the outside in: long-range tertiary contacts in RNA exhibit distinct catalytic roles. *Biochemistry* **50**, 8733–8755 (2011).

25. 25.

Doherty, E. A. & Doudna, J. A. Ribozyme structures and mechanisms. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 457–475 (2001).

26. 26.

Joyce, G. F., van der Horst, G. & Inoue, T. Catalytic activity is retained in the *Tetrahymena* group I intron despite removal of the large extension of element P5. *Nucleic Acids Res.* **17**, 7879–7889 (1989).

27. 27.

Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem. Biol.* **3**, 993–1009 (1996).

28. 28.

Russell, R. et al. The paradoxical behavior of a highly structured misfolded intermediate in RNA folding. *J. Mol. Biol.* **363**, 531–544 (2006).

29. 29.

Cate, J. H. et al. RNA tertiary structure mediation by adenosine platforms. *Science* **273**, 1696–1699 (1996).

30. 30.

Cate, J. H., Hanna, R. L. & Doudna, J. A. A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol.* **4**, 553–558 (1997).

31. 31.

Shoffner, G. M., Wang, R., Podell, E., Cech, T. R. & Guo, F. In crystallo selection to establish new RNA crystal contacts. *Structure* **26**, 1275–1283.e3 (2018).

32. 32.

Pintilie, G. et al. Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020).

33. 33.

Shi, X. et al. Roles of long-range tertiary interactions in limiting dynamics of the *Tetrahymena* group I ribozyme. *J. Am. Chem. Soc.* **136**, 6643–6648 (2014).

34. 34.

Downs, W. D. & Cech, T. R. A tertiary interaction in the *Tetrahymena* intron contributes to selection of the 5' splice site. *Genes Dev.* **8**, 1198–1211 (1994).

35. 35.

Zarrinkar, P. P. & Williamson, J. R. The P9.1-P9.2 peripheral extension helps guide folding of the *Tetrahymena* ribozyme. *Nucleic Acids Res.* **24**, 854–858 (1996).

36. 36.

Correll, C. C., Freeborn, B., Moore, P. B. & Steitz, T. A. Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell* **91**, 705–712 (1997).

37. 37.

Wang, J. F., Downs, W. D. & Cech, T. R. Movement of the guide sequence during RNA catalysis by a group I ribozyme. *Science* **260**, 504–508 (1993).

38. 38.

Lipchock, S. V. & Strobel, S. A. A relaxed active site after exon ligation by the group I intron. *Proc. Natl Acad. Sci. USA* **105**, 5699–5704 (2008).

39. 39.

Karbstein, K. & Herschlag, D. Extraordinarily slow binding of guanosine to the *Tetrahymena* group I ribozyme: implications for RNA preorganization and function. *Proc. Natl Acad. Sci. USA* **100**, 2300–2305 (2003).

40. 40.

Karbstein, K., Lee, J. & Herschlag, D. Probing the role of a secondary structure element at the 5'- and 3'-splice sites in group I intron self-splicing: the *Tetrahymena* L-16 ScaI ribozyme reveals a new role of the G.U pair in self-splicing. *Biochemistry* **46**, 4861–4875 (2007).

41. 41.

Shi, X., Mollova, E. T., Pljevaljcić, G., Millar, D. P. & Herschlag, D. Probing the dynamics of the P1 helix within the *Tetrahymena* group I intron. *J. Am. Chem. Soc.* **131**, 9571–9578 (2009).

42. 42.

Shan, S., Kravchuk, A. V., Piccirilli, J. A. & Herschlag, D. Defining the catalytic metal ion interactions in the *Tetrahymena* ribozyme reaction. *Biochemistry* **40**, 5161–5171 (2001).

43. 43.

Shan, S. O. & Herschlag, D. An unconventional origin of metal-ion rescue and inhibition in the *Tetrahymena* group I ribozyme reaction. *RNA* **6**, 795–813 (2000).

44. 44.

Sjögren, A. S., Pettersson, E., Sjöberg, B. M. & Strömberg, R. Metal ion interaction with cosubstrate in self-splicing of group I introns. *Nucleic Acids Res.* **25**, 648–653 (1997).

45. 45.

Weinstein, L. B., Jones, B. C., Cosstick, R. & Cech, T. R. A second catalytic metal ion in group I ribozyme. *Nature* **388**, 805–808 (1997).

46. 46.

Forconi, M., Lee, J., Lee, J. K., Piccirilli, J. A. & Herschlag, D. Functional identification of ligands for a catalytic metal ion in group I introns. *Biochemistry* **47**, 6883–6894 (2008).

47. 47.

Hougland, J. L., Kravchuk, A. V., Herschlag, D. & Piccirilli, J. A. Functional identification of catalytic metal ion binding sites within RNA. *PLoS Biol.* **3**, e277 (2005).

48. 48.

Zaug, A. J., Grosshans, C. A. & Cech, T. R. Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes. *Biochemistry* **27**, 8924–8931 (1988).

49. 49.

Kladwang, W., Hum, J. & Das, R. Ultraviolet shadowing of RNA can cause significant chemical damage in seconds. *Sci. Rep.* **2**, 517 (2012).

50. 50.

Rajagopal, J., Doudna, J. A. & Szostak, J. W. Stereochemical course of catalysis by the *Tetrahymena* ribozyme. *Science* **244**, 692–694 (1989).

51. 51.

Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

52. 52.

Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

53. 53.

Chen, M. et al. Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. *Nat. Methods* **14**, 983–985 (2017).

54. 54.

Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

55. 55.

Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**, 5–17 (2019).

56. 56.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

57. 57.

Kappel, K. et al. De novo computational RNA modeling into cryo-EM maps of large ribonucleoprotein complexes. *Nat. Methods* **15**, 947–954 (2018).

58. 58.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

59. 59.

Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).

60. 60.

Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

61. 61.

Frederiksen, J. K., Li, N. S., Das, R., Herschlag, D. & Piccirilli, J. A. Metal-ion rescue revisited: biochemical detection of site-bound metal ions important for RNA folding. *RNA* **18**, 1123–1141 (2012).

62. 62.

Strauss-Soukup, J. K. & Strobel, S. A. A chemical phylogeny of group I introns based upon interference mapping of a bacterial ribozyme. *J. Mol. Biol.* **302**, 339–358 (2000).

63. 63.

Ortoleva-Donnelly, L., Szewczak, A. A., Gutell, R. R. & Strobel, S. A. The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA* **4**, 498–519 (1998).

64. 64.

Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).

65. 65.

Green, R. & Szostak, J. W. In vitro genetic analysis of the hinge region between helical elements P5-P4-P6 and P7-P3-P8 in the sunY group I self-splicing intron. *J. Mol. Biol.* **235**, 140–155 (1994).

66. 66.

Karbstein, K., Tang, K. H. & Herschlag, D. A base triple in the *Tetrahymena* group I core affects the reaction equilibrium via a threshold effect. *RNA* **10**, 1730–1739 (2004).

67. 67.

Tanner, M. A., Anderson, E. M., Gutell, R. R. & Cech, T. R. Mutagenesis and comparative sequence analysis of a base triple joining the two domains of group I ribozymes. *RNA* **3**, 1037–1051 (1997).

68. 68.

Tanner, M. A. & Cech, T. R. Joining the two domains of a group I ribozyme to form the catalytic core. *Science* **275**, 847–849 (1997).

69. 69.

Ikawa, Y., Yoshimura, T., Hara, H., Shiraishi, H. & Inoue, T. Two conserved structural components, A-rich bulge and P4 XJ6/7 base-triples, in activating the group I ribozymes. *Genes Cells* **7**, 1205–1215 (2002).

70. 70.

Ikawa, Y., Naito, D., Shiraishi, H. & Inoue, T. Structure–function relationships of two closely related group IC3 intron ribozymes from *Azoarcus* and *Synechococcus* pre-tRNA. *Nucleic Acids Res.* **28**, 3269–3277 (2000).

71. 71.

Michel, F., Ellington, A. D., Couture, S. & Szostak, J. W. Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns.

*Nature* **347**, 578–580 (1990).

72. 72.

Szewczak, A. A. et al. An important base triple anchors the substrate helix recognition surface within the *Tetrahymena* ribozyme active site. *Proc. Natl Acad. Sci. USA* **96**, 11183–11188 (1999).

73. 73.

Rangan, P., Masquida, B., Westhof, E. & Woodson, S. A. Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA* **100**, 1574–1579 (2003).

74. 74.

Green, R., Ellington, A. D. & Szostak, J. W. In vitro genetic analysis of the *Tetrahymena* self-splicing intron. *Nature* **347**, 406–408 (1990).

75. 75.

Mitchell, D., III, Jarmoskaite, I., Seval, N., Seifert, S. & Russell, R. The long-range P3 helix of the *Tetrahymena* ribozyme is disrupted during folding between the native and misfolded conformations. *J. Mol. Biol.* **425**, 2670–2686 (2013).

## Acknowledgements

We thank R. N. Sengupta, D. Herschlag and I. Zheludev for discussions. Cryo-EM data were collected at SLAC and Stanford, and processed at SLAC, Stanford and Duyu High Performance Computing Center in Sichuan University. This work was supported by the National Institutes of Health (P41GM103832, R01GM079429, P01AI120943, and S10OD021600 to W.C.; R35GM112579 and R21AI145647 to R.D.); National Science Foundation (DGE-114747 to K.K. and DGE-1656518 to R.R.); Gabilan Stanford Graduate Fellowship to K.K.; Sichuan University start-up funding 20822041D4057 and Natural Science Foundation of China (NSFC82041016 and 32070049 to Z.S). We thank N. Lawless for manuscript editing.

## Author information

Author notes

1. These authors contributed equally: Zhaoming Su, Kaiming Zhang, Kalli Kappel

## Affiliations

1. The State Key Laboratory of Biotherapy and Cancer Center, Department of Geriatrics and National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, Chengdu, China

Zhaoming Su, Bingnan Luo & Yuquan Wei

2. Department of Bioengineering and James H. Clark Center, Stanford University, Stanford, CA, USA

Zhaoming Su, Kaiming Zhang, Shanshan Li, Grigore D. Pintilie & Wah Chiu

3. MOE Key Laboratory for Membraneless Organelles and Cellular Dynamics, Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, China

Kaiming Zhang & Shanshan Li

4. Biophysics Program, Stanford University, Stanford, CA, USA

Kalli Kappel, Michael Z. Palo, Ramya Rangan & Rhiju Das

5. Department of Biochemistry and Department of Physics, Stanford University, Stanford, CA, USA

Rhiju Das

6. Division of CryoEM and Bioimaging, SSRL, SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Wah Chiu

## Contributions

Z.S., K.Z., K.K., R.D., W.C. conceived the project; K.K. and M.Z.P. prepared RNA samples; Z.S. and K.Z. collected cryo-EM data; Z.S., K.Z., S.L., and Y.W. processed cryo-EM data; Z.S., K.K., S.L., B.L. and G.D.P. built and refined atomic models; G.D.P, B.L., S.L. and R.R. validated the models and map–model correlation. All authors contributed to the preparation of the manuscript.

## Corresponding authors

Correspondence to [Zhaoming Su](#) or [Rhiju Das](#) or [Wah Chiu](#).

## Ethics declarations

### Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Thomas Cech, Amedee des Georges and Andrej Lupták for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

### [Extended Data Fig. 1 Cryo-EM single-particle reconstruction of the apo L-21 Scal ribozyme.](#)

Related to Fig. 1. **a**, Single-particle pipeline yields the final cryo-EM reconstruction with the corresponding angular distribution and local resolution map. The local resolution map shows more flexibility and lower resolution in the 4-nt bulge that connects stem P2.1 and P13, and towards the end of stem P6 and P9.2. **b**, FSC curve shows 3.1 Å resolution according to the 0.143 cutoff. **c**, Cryo-EM *B* factor<sup>64</sup>, which relates the number of particles to the map resolution attributed to cumulative experimental and computational factors that affect the final reconstruction.

### [Extended Data Fig. 2 Focused 3D classification of apo L-21 Scal ribozyme reveals local conformational dynamics.](#)

The regions of low local resolution, P9.2 (blue), P9.2–P9.1–P13 (orange) and P13 (green), were extracted. Focused 3D classification was performed and different classes were superimposed to show rotational and translational motions on P9.2 (left), P9.2–P9.1–P13 (middle) and P13 (right).

## Extended Data Fig. 3 Q-score analyses of cryo-EM maps and models of both the apo L-21 and holo L-16 ScaI ribozymes.

Related to Fig. 1. **a**,  $Q$ -score analyses per residue of the apo L-21 (grey) and holo L-16 (blue) ScaI ribozyme cryo-EM models and maps. Black dashed line indicates average  $Q$ -score from nucleic acid cryo-EM models and maps at 3.1 Å resolution in the PDB. **b**, Cryo-EM model of apo L-21 ScaI ribozyme coloured according to  $Q$ -score per residue. Dashed boxes (black, blue and green) correspond to regions in the cryo-EM model with low  $Q$ -scores in **a**.

## Extended Data Fig. 4 Detailed tertiary interactions in the core region of the apo L-21 ScaI ribozyme.

**a–f**, The P5–J5/5a–L9 region has a highly structured J5–5a junction in previous structures<sup>12,15</sup>. The cryo-EM structure shows tertiary interactions of C197 (**a**), C124 and A125 (**b**), G126, A324 and A325 (**d, e**) and G327 (**f**) in the minor groove of P5. **g–k**, Previous studies have shown that A183 and A184 in the A-rich bulge of the metal core, U259 and C260 from J6/J7 and U305 from J8/J7 are conserved and essential for catalytic site formation and splicing reactions<sup>65,66,67,68,69,70,71</sup>. The cryo-EM structure shows that U259 (**h**), C260 (**i**) and U305 (**j, k**) stack continuously and interact with P4 base triples. **k**, U168 from P5c stacks on the A-rich bulge and interacts with P4 in the minor groove, while pairing with G188 in P5a. **l**, The Hoogsteen base triple U277–A97–U300 is essential for substrate helix recognition<sup>72</sup>. **m–q**, In the P3–J3/4–P6–J7/3 region, the A-rich J3/4 and J7/3 were previously found to interact with P6 in the *Azoarcus* ribozyme<sup>17,73</sup>. Base triples formed by J3/4 were critical for catalysis<sup>65,70,71,74</sup>, and alterations in these regions result in RNA misfolding<sup>28,75</sup>. In the cryo-EM structure, A104 and A105 form A-minor interactions with P6, whereas A103 and A104 join A269 and A270 from J7/3 to form an adenosine cluster (**m, n**). The proposed A103–U271 reverse-Hoogsteen pair is not found; instead we observed a noncanonical A103–A270 pair<sup>74</sup>. **p, q**, The previously observed A-platform of A218–A219 is disrupted in the cryo-EM structure with P3 present<sup>12,29</sup>. A218 forms two A-minor interactions with C102–G272 and U273–U101 from P3, which also supports the conservation of this C–G pair in group Ib introns<sup>74</sup>. Black dashed lines indicate hydrogen bonds. The cryo-EM maps of all subpanels are visualized at  $1\sigma$  threshold except for **c, g, o** ( $1.5\sigma$ ).

## Extended Data Fig. 5 Comparison between the previous 3.8 Å crystal structure of the mutated *Tetrahymena* ribozyme catalytic core (green).

**and the cryo-EM structure of the wild-type apo L-21 ScaI ribozyme construct (grey) shows minor differences.**

Related to Extended Data Fig. 4. The overall r.m.s.d. for the catalytic core region (stem P3–P9) is 6.6 Å. **a**, The same view of the P5–J5/5a–P9 region as in Extended Data Fig. 4c. The nucleotide conformations generally agree well between two models; three mutations (U322C, U323G and U326A) are not involved in tertiary interactions. The r.m.s.d. for this region is 4.9 Å. **b**, The same view of the P4–P5a–J6/7–J8/7 region as in Extended Data Fig. 4g. In the crystal structure, U259A is slightly moved away from the G108–C213 base pair and disrupts this base triple interaction. A210G is moved far away from the wild-type position of A210, because there is no A46 in stem P2 to interact with in the crystal structure. The very top base quartet is much more compact in the cryo-EM structure compared to the crystal structure, probably owing to the presence of the peripheral domain that wraps around the catalytic core to make it more compact. The r.m.s.d. for this region is 5.7 Å. **c**, The same view of the P3–J3/4–P6–J7/3 region as in Extended Data Fig. 4o. The overall nucleotide conformations agree very well between the two models, except that A269G and A270 in the crystal structure are completely moved away and disrupt their interactions with A103, which is observed in the cryo-EM structure. The r.m.s.d. for this region is 1.7 Å. **d**, The same view as in Extended Data Fig. 4l. U277C disrupts the U277–A97–U300 base triple. The r.m.s.d. for this base triple is 2.7 Å. See also Supplementary Table 1.

**Extended Data Fig. 6 Superposition of the apo L-21 ScaI ribozyme cryo-EM structure (grey) with previous crystal structures of the truncated and/or mutated *Tetrahymena* ribozyme, other group I introns and 5S rRNA loop E show global and local structural similarities.**

**a–d**, Overlays of the cryo-EM structure (grey) with the *Tetrahymena* ribozyme P4–P6 Delta C209 (**a**; blue, PDB 1HR2); the mutated *Tetrahymena* ribozyme P3–P9 (**b**; green, PDB 1X8W); the *Azoarcus* ribozyme (**c**; violet, PDB 1U6B) and the phage Twort ribozyme (**d**; yellow, PDB 1Y0Q). **e**, P5c region of the wild type P4–P6 crystal structure (blue, PDB 1GID). **f**, 5S rRNA loop E crystal structure (red, PDB 354D).

**Extended Data Fig. 7 Cryo-EM single-particle reconstruction of the holo L-16 ScaI ribozyme.**

Related to Fig. 3. **a**, Single-particle pipeline yields the final cryo-EM reconstruction with the corresponding angular distribution and local resolution map. The local resolution map shows more flexibility and lower resolution in the 4-nt bulge that connects stem P2.1 and P13, and towards the end of stem P6 and P9.2. **b**, FSC curve shows 3.1 Å resolution according to the 0.143 cutoff. **c**, Cryo-EM *B* factor.

**Extended Data Fig. 8 Comparisons of apo L-21 and holo L-16 ScaI ribozyme cryo-EM models with previous crystal structures show structural conservation and metal ion shifts in the guanosine binding site among group I introns.**

Related to Fig. 4. **a–c**, The apo L-21 ScaI ribozyme adopts a preorganized guanosine binding site (grey) that superimposes with previous crystal structures of mutated P3–P9 of the *Tetrahymena* ribozyme (**a**; green, PDB 1X8W), the *Azoarcus* ribozyme (**b**; violet, PDB 1U6B), and the phage Twort ribozyme (**c**; yellow, PDB 1y0q). **d–g**, The holo L-16 ScaI ribozyme (sky blue) superimposes with apo L-21 ScaI ribozyme (**d**; grey), the *Azoarcus* ribozyme (**e**; violet), the phage Twort ribozyme (**f**; yellow), and mutated P3–P9 of the *Tetrahymena* ribozyme (**g**; green). M<sub>C</sub> is absent in the apo L-21 ScaI ribozyme, whereas M<sub>C</sub> in the *Azoarcus* ribozyme is shifted compared to the holo L-16 ScaI ribozyme. Dash line indicates metal ion coordination with surrounding atoms.

**Extended Data Fig. 9 Metal ion validations by distance and *Q*-score analysis, and illustrations of the apo L-21 and holo L-16 ScaI ribozyme cryo-EM structures compared with previous crystal structures.**

**a**, Distances between metal ions and other atoms in the apo L-21 ScaI ribozyme model. **b**, Distances between metal ions and other atoms in the holo L-16 ScaI ribozyme model. **c**, *Q*-score analysis per metal ion of the apo L-21 and holo L-16 ScaI ribozyme cryo-EM models and maps. **d**, Metal core region of the holo L-16 ScaI ribozyme, visualized at 1.1 $\sigma$  threshold. **e**, Comparisons of the apo L-21 (grey) and holo L-16 (sky blue) ScaI ribozyme cryo-EM models with P4–P6 Delta C209 (blue, PDB 1HR2) and mutated P3–P9 of the *Tetrahymena* ribozyme (green, PDB 1X8W) in the same view as **d**. **f**, Catalytic site of the holo L-16 ScaI ribozyme, visualized at 1.4 $\sigma$  threshold. **g**, Comparisons of the apo L-21 (grey) and holo L-16 (sky blue) ScaI ribozyme cryo-EM models with the *Azoarcus* ribozyme (violet, PDB 1U6B), mutated P3–P9 of the *Tetrahymena* ribozyme (green, PDB 1X8W), and the phage Twort ribozyme (yellow, PDB 1Y0Q) in the same view as **f**. See also Supplementary Table 2.

## **Extended Data Table 1 Cryo-EM data collection, processing, and model refinement statistics of the apo L-21 and holo L-16 ScaI ribozymes**

## **Supplementary information**

### **Supplementary Tables**

This file contains Supplementary Tables 1-2.

### **Reporting Summary**

### **Peer Review File**

### **Video 1**

Cryo-EM map and model of the apo L-21 ScaI *Tetrahymena* ribozyme with zoom in views of newly discovered tertiary interactions and representative metal ions shown in Figure 2, accompanied with the updated secondary structure.

### **Video 2**

Comparison of the apo L-21 and holo L-16 *Tetrahymena* ribozyme cryo-EM maps and models reveals substantial conformational change upon substrate binding.

## **Rights and permissions**

### **Reprints and Permissions**

## **About this article**

### **Cite this article**

Su, Z., Zhang, K., Kappel, K. *et al.* Cryo-EM structures of full-length *Tetrahymena* ribozyme at 3.1 Å resolution. *Nature* **596**, 603–607 (2021).  
<https://doi.org/10.1038/s41586-021-03803-w>

- Received: 13 January 2021
- Accepted: 06 July 2021

- Published: 11 August 2021
  - Issue Date: 26 August 2021
  - DOI: <https://doi.org/10.1038/s41586-021-03803-w>
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03803-w>

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

- Matters Arising
- [Published: 25 August 2021](#)

# Unusual width of the superconducting transition in a hydride

- [J. E. Hirsch](#) <sup>1</sup> [ORCID: orcid.org/0000-0001-7175-3497](#)<sup>1</sup> &
- [F. Marsiglio](#) <sup>2</sup>

*Nature* volume **596**, pages E9–E10 (2021)

- 1643 Accesses
- 6 Altmetric
- [Metrics details](#)

## Subjects

- [Electronic properties and materials](#)
- [Superconducting properties and materials](#)

The [Original Article](#) was published on 14 October 2020

arising from E. Snider et al. *Nature* <https://doi.org/10.1038/s41586-020-2801-z> (2020)

## Access options

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

## Rent or Buy

All prices are NET prices.

## Access all Nature Journals

Immediate access to more than 50 journals, including  
*Nature*

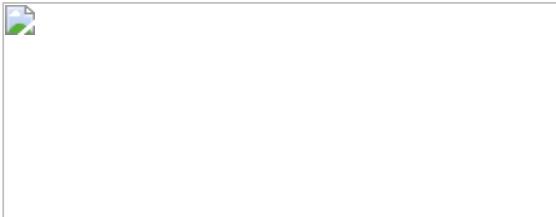
\$19.99/month

Learn more

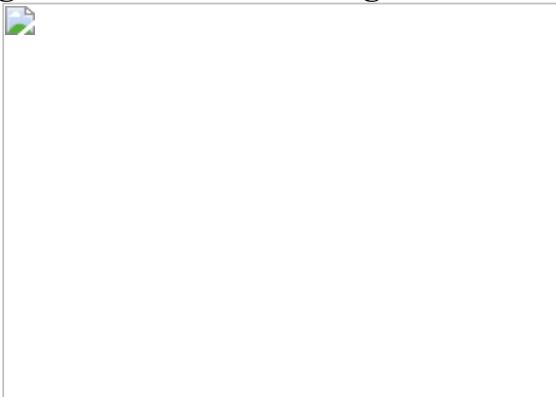
## Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Resistive transition in a magnetic field for MgB<sub>2</sub> and YBCO.**



**Fig. 2: Relative broadening of the resistive transition versus magnetic field.**



## References

1. 1.

Snider, E. et al. Room-temperature superconductivity in a carbonaceous sulfur hydride. *Nature* **586**, 373–377 (2020).

2. 2.

Tinkham, M. *Introduction to Superconductivity* (McGraw Hill, 1996).

3. 3.

Pickard, C. J., Errea, I. & Eremets, M. I. Superconducting hydrides under pressure. *Annu. Rev. Condens. Matter Phys.* **11**, 57–76 (2020).

4. 4.

Kim, Y. B., Hempstead, C. F. & Strnad, A. R. Flux-flow resistance in type-II superconductors. *Phys. Rev.* **139**, A1163 (1965).

5. 5.

Anderson, P. W. & Kim, Y. B. Hard superconductivity: theory of the motion of Abrikosov flux lines. *Rev. Mod. Phys.* **36**, 39 (1964).

6. 6.

Fisher, D. S., Fisher, M. P. A. & Huse, D. A. Thermal fluctuations, quenched disorder, phase transitions, and transport in type-II superconductors. *Phys. Rev. B* **43**, 130–159 (1991).

7. 7.

Canfield, P. C., Bud'ko, S. L. & Finnemore, D. K. An overview of the basic physical properties of MgB<sub>2</sub>. *Physica C* **385**, 1–7 (2003).

8. 8.

Gupta, A. et al. Resistivity broadening, upper critical fields and irreversibility lines in bulk PbMo<sub>6</sub>S<sub>8</sub> and SnMo<sub>6</sub>S<sub>8</sub> Chevrel phase superconductors. *Physica C* **235–240**, 2541–2542 (1994).

9. 9.

Guo, J. et al. Robust zero resistance in a superconducting high-entropy alloy at pressures up to 190 GPa. *Proc. Natl Acad. Sci. USA* **114**, 13144–13147 (2017).

10. 10.

Kumar, D. et al. Flux pinning and improved critical current density in superconducting boron doped diamond films. *J. Phys. Commun.* **2**, 045015 (2018).

11. 11.

Eisaki, H. et al. Competition between magnetism and superconductivity in rare-earth nickel boride carbides. *Phys. Rev. B* **50**, 647(R) (1994).

12. 12.

Iye, Y. et al. The anisotropic superconductivity of RBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub> (R: Y, Gd and Ho) single crystals. *Physica C* **153–155**, 26–31 (1988).

13. 13.

Kitazawa, K. et al. Broadening mechanism of resistive transition under magnetic field in single crystalline (La<sub>1-x</sub>Sr<sub>x</sub>)<sub>2</sub>CuO<sub>4</sub>. *Jpn. J. Appl. Phys.* **28**, L555 (1989).

14. 14.

Ito, H. et al. Resistive superconducting transition of  $\kappa$ -type BEDT-TTF organic superconductors in a magnetic field. *J. Supercond.* **7**, 667–669 (1994).

15. 15.

Chen, X. H. et al. Superconductivity at 43 K in SmFeAsO<sub>1-x</sub>F<sub>x</sub>. *Nature* **453**, 761–762 (2008).

16. 16.

Jung, S.-G. et al. Influence of carbon-ion irradiation on the superconducting critical properties of MgB<sub>2</sub> thin films. *Supercond. Sci. Technol.* **32**, 025006 (2019).

17. 17.

Song, J., Fabbris, G., Bi, W., Haskel, D. & Schilling, J. S. Pressure-induced superconductivity in elemental ytterbium metal. *Phys. Rev. Lett.* **121**, 037004 (2018).

## Acknowledgements

We acknowledge clarifying correspondence with the authors of ref. [1](#). F.M. was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by an MIF from the Province of Alberta.

## Author information

### Affiliations

1. Department of Physics, University of California, San Diego, La Jolla, CA, USA

J. E. Hirsch

2. Department of Physics, University of Alberta, Edmonton, Alberta, Canada

F. Marsiglio

### Contributions

J.E.H. and F.M. contributed equally to all aspects of the preparation of this work.

## Corresponding authors

Correspondence to [J. E. Hirsch](#) or [F. Marsiglio](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

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

## Rights and permissions

[Reprints and Permissions](#)

## About this article

## Cite this article

Hirsch, J.E., Marsiglio, F. Unusual width of the superconducting transition in a hydride. *Nature* **596**, E9–E10 (2021). <https://doi.org/10.1038/s41586-021-03595-z>

- Received: 20 October 2020
- Accepted: 23 April 2021
- Published: 25 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03595-z>

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

# Amendments & Corrections

- [\*\*Author Correction: Molecular logic of cellular diversification in the mouse cerebral cortex\*\*](#) [ 02 August 2021]  
Author Correction •
- [\*\*Author Correction: Quantum-enhanced nonlinear microscopy\*\*](#) [ 02 August 2021]  
Author Correction •
- [\*\*Author Correction: DHODH-mediated ferroptosis defence is a targetable vulnerability in cancer\*\*](#) [ 02 August 2021]  
Author Correction •
- [\*\*Publisher Correction: Observation of first and second sound in a BKT superfluid\*\*](#) [ 02 August 2021]  
Publisher Correction •
- [\*\*Publisher Correction: A highly magnetized and rapidly rotating white dwarf as small as the Moon\*\*](#) [ 02 August 2021]  
Publisher Correction •

- Author Correction
- [Published: 02 August 2021](#)

# Author Correction: Molecular logic of cellular diversification in the mouse cerebral cortex

- [Daniela J. Di Bella](#) ORCID: [orcid.org/0000-0001-6141-3704](https://orcid.org/0000-0001-6141-3704)<sup>1,2 na2</sup>,
- [Ehsan Habibi](#) ORCID: [orcid.org/0000-0002-7208-7999](https://orcid.org/0000-0002-7208-7999)<sup>1,3 na1</sup>,
- [Robert R. Stickels](#) ORCID: [orcid.org/0000-0003-4326-4084](https://orcid.org/0000-0003-4326-4084)<sup>2</sup>,
- [Gabriele Scalia](#) ORCID: [orcid.org/0000-0003-3305-9220](https://orcid.org/0000-0003-3305-9220)<sup>3</sup>,
- [Juliana Brown](#)<sup>1,2</sup>,
- [Payman Yadollahpour](#) ORCID: [orcid.org/0000-0003-1984-5014](https://orcid.org/0000-0003-1984-5014)<sup>3</sup>,
- [Sung Min Yang](#) ORCID: [orcid.org/0000-0003-4318-2011](https://orcid.org/0000-0003-4318-2011)<sup>1,2</sup>,
- [Catherine Abbate](#)<sup>1,2</sup>,
- [Tommaso Biancalani](#) ORCID: [orcid.org/0000-0001-9104-9755](https://orcid.org/0000-0001-9104-9755)<sup>3 nAff5</sup>,
- [Evan Z. Macosko](#) ORCID: [orcid.org/0000-0002-2794-5165](https://orcid.org/0000-0002-2794-5165)<sup>2</sup>,
- [Fei Chen](#)<sup>1,2</sup>,
- [Aviv Regev](#) ORCID: [orcid.org/0000-0003-3293-3158](https://orcid.org/0000-0003-3293-3158)<sup>3,4 na2 nAff5</sup> &
- [Paola Arlotta](#) <sup>1,2 na2</sup>

*Nature* volume **596**, page E11 (2021)

- 510 Accesses
- 2 Altmetric
- [Metrics details](#)

## Subjects

- [Cell fate and cell lineage](#)
- [Developmental neurogenesis](#)

The [Original Article](#) was published on 23 June 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03670-5>

Published online 23 June 2021

In this Article, the first name of author Tommaso Biancalani was incorrectly spelled ‘Tommasso’. The original Article has been corrected online.

## Author information

### Author notes

1. Tommaso Biancalani & Aviv Regev

Present address: Genentech, South San Francisco, CA, USA

2. These authors contributed equally: Daniela J. Di Bella, Ehsan Habibi
3. These authors jointly supervised this work: Aviv Regev, Paola Arlotta

## Affiliations

1. Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA

Daniela J. Di Bella, Ehsan Habibi, Juliana Brown, Sung Min Yang, Catherine Abbate, Fei Chen & Paola Arlotta

2. Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Daniela J. Di Bella, Robert R. Stickels, Juliana Brown, Sung Min Yang, Catherine Abbate, Evan Z. Macosko, Fei Chen & Paola Arlotta

3. Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Ehsan Habibi, Gabriele Scalia, Payman Yadollahpour, Tommaso Biancalani & Aviv Regev

4. Howard Hughes Medical Institute, Koch Institute of Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

Aviv Regev

## Corresponding authors

Correspondence to [Aviv Regev](#) or [Paola Arlotta](#).

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Di Bella, D.J., Habibi, E., Stickels, R.R. *et al.* Author Correction: Molecular logic of cellular diversification in the mouse cerebral cortex. *Nature* **596**, E11 (2021). <https://doi.org/10.1038/s41586-021-03797-5>

- Published: 02 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03797-5>

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03797-5>

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

- Author Correction
- [Published: 02 August 2021](#)

# Author Correction: Quantum-enhanced nonlinear microscopy

- [Catxere A. Casacio](#)<sup>1</sup>,
- [Lars S. Madsen](#)<sup>1</sup>,
- [Alex Terrasson](#)<sup>1</sup>,
- [Muhammad Waleed](#)<sup>1</sup>,
- [Kai Barnscheidt](#) [ORCID: orcid.org/0000-0001-6783-7372](#)<sup>2</sup>,
- [Boris Hage](#)<sup>2</sup>,
- [Michael A. Taylor](#) [ORCID: orcid.org/0000-0003-4393-7760](#)<sup>3</sup> &
- [Warwick P. Bowen](#) [ORCID: orcid.org/0000-0001-8127-1715](#)<sup>1</sup>

[Nature](#) volume **596**, page E12 (2021)

- 483 Accesses
- 2 Altmetric
- [Metrics details](#)

## Subjects

- [Multiphoton microscopy](#)
- [Quantum optics](#)

The [Original Article](#) was published on 09 June 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03528-w>  
Published online 09 June 2021

In the Acknowledgements section of this Article, the sentence “This work was supported primarily by the Air Force Office of Scientific Research (AFOSR) grant FA2386-14-1-4046.” should have been “This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-17-1-0397.”. The original Article has been corrected online.

## Author information

### Affiliations

1. ARC Centre of Excellence for Engineered Quantum Systems,  
University of Queensland, St Lucia, Queensland, Australia

Catxere A. Casacio, Lars S. Madsen, Alex Terrasson, Muhammad  
Waleed & Warwick P. Bowen

2. Institut für Physik, Universität Rostock, Rostock, Germany

Kai Barnscheidt & Boris Hage

3. Australian Institute for Bioengineering and Nanotechnology, The  
University of Queensland, St Lucia, Queensland, Australia

Michael A. Taylor

### Corresponding author

Correspondence to [Warwick P. Bowen](#).

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Casacio, C.A., Madsen, L.S., Terrasson, A. *et al.* Author Correction: Quantum-enhanced nonlinear microscopy. *Nature* **596**, E12 (2021). <https://doi.org/10.1038/s41586-021-03808-5>

- Published: 02 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03808-5>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03808-5>

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

- Author Correction
- [Published: 02 August 2021](#)

# Author Correction: DHODH-mediated ferroptosis defence is a targetable vulnerability in cancer

- [Chao Mao](#) ORCID: [orcid.org/0000-0002-8685-8539<sup>1</sup>](https://orcid.org/0000-0002-8685-8539),
- [Xiaoguang Liu](#) ORCID: [orcid.org/0000-0001-6977-5128<sup>1</sup>](https://orcid.org/0000-0001-6977-5128),
- [Yilei Zhang](#) ORCID: [orcid.org/0000-0003-1407-0301<sup>1</sup>](https://orcid.org/0000-0003-1407-0301),
- [Guang Lei](#) ORCID: [orcid.org/0000-0002-3282-0666<sup>1</sup>](https://orcid.org/0000-0002-3282-0666),
- [Yuelong Yan](#) ORCID: [orcid.org/0000-0001-8495-6445<sup>1</sup>](https://orcid.org/0000-0001-8495-6445),
- [Hyemin Lee<sup>1</sup>](#),
- [Pranavi Koppula](#) ORCID: [orcid.org/0000-0002-7567-1891<sup>1,2</sup>](https://orcid.org/0000-0002-7567-1891),
- [Shiqi Wu<sup>1</sup>](#),
- [Li Zhuang<sup>1</sup>](#),
- [Bingliang Fang<sup>3</sup>](#),
- [Masha V. Poyurovsky<sup>4</sup>](#),
- [Kellen Olszewski](#) ORCID: [orcid.org/0000-0001-9691-0831<sup>4</sup>](https://orcid.org/0000-0001-9691-0831) &
- [Boyi Gan](#) ORCID: [orcid.org/0000-0001-8884-6040<sup>1,2</sup>](https://orcid.org/0000-0001-8884-6040)

*Nature* volume **596**, page E13 (2021)

- 727 Accesses
- 2 Altmetric
- [Metrics details](#)

The [Original Article](#) was published on 12 May 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03539-7> Published online 12 May 2021

In the Acknowledgements section of this Article, the text “PDX generation and annotation were supported by the University of Texas MD Anderson Cancer Center Moon Shots Program, and Specialized Program of Research Excellence (SPORE) grant CA070907.” has been changed to “PDX generation and annotation were supported by the University of Texas MD Anderson Cancer Center Moon Shots Program, Specialized Program of Research Excellence (SPORE) grant CA070907 **and University of Texas PDX Development and Trial Center grant U54CA224065.**” The original Article has been corrected online.

## Author information

### Affiliations

1. Department of Experimental Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Chao Mao, Xiaoguang Liu, Yilei Zhang, Guang Lei, Yuelong Yan, Hyemin Lee, Pranavi Koppula, Shiqi Wu, Li Zhuang & Boyi Gan

2. The University of Texas MD Anderson UTHealth Graduate School of Biomedical Sciences, Houston, TX, USA

Pranavi Koppula & Boyi Gan

3. Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Bingliang Fang

4. Kadmon Corporation, LLC, New York, NY, USA

Masha V. Poyurovsky & Kellen Olszewski

## Corresponding authors

Correspondence to [Kellen Olszewski](#) or [Boyi Gan](#).

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Mao, C., Liu, X., Zhang, Y. *et al.* Author Correction: DHODH-mediated ferroptosis defence is a targetable vulnerability in cancer. *Nature* **596**, E13 (2021). <https://doi.org/10.1038/s41586-021-03820-9>

- Published: 02 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03820-9>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03820-9>

- Publisher Correction
- [Published: 02 August 2021](#)

# Publisher Correction: Observation of first and second sound in a BKT superfluid

- [Panagiotis Christodoulou](#) ORCID: [orcid.org/0000-0001-7800-9484<sup>1</sup>](https://orcid.org/0000-0001-7800-9484),
- [Maciej Gałka<sup>1</sup>](#),
- [Nishant Dogra<sup>1</sup>](#),
- [Raphael Lopes](#) ORCID: [orcid.org/0000-0003-3877-8478<sup>2</sup>](https://orcid.org/0000-0003-3877-8478),
- [Julian Schmitt](#) ORCID: [orcid.org/0000-0002-0002-3777<sup>1</sup>](https://orcid.org/0000-0002-0002-3777) nAff3 &
- [Zoran Hadzibabic](#) ORCID: [orcid.org/0000-0002-0118-9285<sup>1</sup>](https://orcid.org/0000-0002-0118-9285)

[Nature](#) volume 596, page E14 (2021)

- 312 Accesses
- 2 Altmetric
- [Metrics details](#)

## Subjects

- [Phase transitions and critical phenomena](#)
- [Quantum fluids and solids](#)
- [Ultracold gases](#)

The [Original Article](#) was published on 09 June 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03537-9>

Published online 9 June 2021

In this Article, in the sentence beginning ‘In absolute terms...’, there is an ‘equal to’ sign missing from the second equation. The equation should read:

$$\langle T \rangle_{\{c\}} = 2\pi n \hbar^2 / [m k_B l n] = 37(6) \times n K.$$

The Article has been corrected online.

## Author information

### Author notes

1. Julian Schmitt

Present address: Institut für Angewandte Physik, Universität Bonn, Bonn, Germany

## Affiliations

1. Cavendish Laboratory, University of Cambridge, Cambridge, UK

Panagiotis Christodoulou, Maciej Gałka, Nishant Dogra, Julian Schmitt & Zoran Hadzibabic

2. Laboratoire Kastler Brossel, Collège de France, CNRS, ENS-PSL Research University, Sorbonne Université, Paris, France

Raphael Lopes

## Corresponding author

Correspondence to [Panagiotis Christodoulou](#).

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Christodoulou, P., Gałka, M., Dogra, N. *et al.* Publisher Correction: Observation of first and second sound in a BKT superfluid. *Nature* **596**, E14 (2021). <https://doi.org/10.1038/s41586-021-03746-2>

- Published: 02 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03746-2>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03746-2>

- Publisher Correction
- [Published: 02 August 2021](#)

# Publisher Correction: A highly magnetized and rapidly rotating white dwarf as small as the Moon

- [Ilaria Caiazzo](#) ORCID: [orcid.org/0000-0002-4770-5388<sup>1</sup>](https://orcid.org/0000-0002-4770-5388),
- [Kevin B. Burdge](#) ORCID: [orcid.org/0000-0002-7226-836X<sup>1</sup>](https://orcid.org/0000-0002-7226-836X),
- [James Fuller](#) ORCID: [orcid.org/0000-0002-4544-0750<sup>1</sup>](https://orcid.org/0000-0002-4544-0750),
- [Jeremy Heyl](#) ORCID: [orcid.org/0000-0001-9739-367X<sup>2</sup>](https://orcid.org/0000-0001-9739-367X),
- [S. R. Kulkarni<sup>1</sup>](#),
- [Thomas A. Prince](#) ORCID: [orcid.org/0000-0002-8850-3627<sup>1</sup>](https://orcid.org/0000-0002-8850-3627),
- [Harvey B. Richer<sup>2</sup>](#),
- [Josiah Schwab<sup>3</sup>](#),
- [Igor Andreoni<sup>1</sup>](#),
- [Eric C. Bellm](#) ORCID: [orcid.org/0000-0001-8018-5348<sup>4</sup>](https://orcid.org/0000-0001-8018-5348),
- [Andrew Drake<sup>1</sup>](#),
- [Dmitry A. Duev](#) ORCID: [orcid.org/0000-0001-5060-8733<sup>1</sup>](https://orcid.org/0000-0001-5060-8733),
- [Matthew J. Graham](#) ORCID: [orcid.org/0000-0002-3168-0139<sup>1</sup>](https://orcid.org/0000-0002-3168-0139),
- [George Helou](#) ORCID: [orcid.org/0000-0003-3367-3415<sup>5</sup>](https://orcid.org/0000-0003-3367-3415),
- [Ashish A. Mahabal](#) ORCID: [orcid.org/0000-0003-2242-0244<sup>1,6</sup>](https://orcid.org/0000-0003-2242-0244),
- [Frank J. Masci](#) ORCID: [orcid.org/0000-0002-8532-9395<sup>5</sup>](https://orcid.org/0000-0002-8532-9395),
- [Roger Smith<sup>7</sup>](#) &
- [Maayane T. Soumagnac<sup>8,9</sup>](#)

*Nature* volume **596**, page E15 (2021)

- 366 Accesses

- 2 Altmetric
- [Metrics details](#)

## Subjects

- [Astrophysical magnetic fields](#)
- [Compact astrophysical objects](#)
- [Stellar evolution](#)

The [Original Article](#) was published on 30 June 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03615-y>.  
Published online 30 June 2021

In Table 1 of this Article, owing to a typesetting error, the value of mass ( $M$  ( $M_{\odot}$ )) was incorrectly shown as ‘ $1.327 \pm 1.365$ ’ instead of as a range as ‘ $1.327–1.365$ ’. The original Article has been corrected online

## Author information

### Affiliations

1. Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, CA, USA

Ilaria Caiazzo, Kevin B. Burge, James Fuller, S. R. Kulkarni, Thomas A. Prince, Igor Andreoni, Andrew Drake, Dmitry A. Duev, Matthew J. Graham & Ashish A. Mahabal

2. Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia, Canada

Jeremy Heyl & Harvey B. Richer

3. Department of Astronomy and Astrophysics, University of California,  
Santa Cruz, CA, USA

Josiah Schwab

4. Department of Astronomy, University of Washington, Seattle, WA,  
USA

Eric C. Bellm

5. IPAC, California Institute of Technology, Pasadena, CA, USA

George Helou & Frank J. Masci

6. Center for Data Driven Discovery, California Institute of Technology,  
Pasadena, CA, USA

Ashish A. Mahabal

7. Caltech Optical Observatories, California Institute of Technology,  
Pasadena, CA, USA

Roger Smith

8. Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Maayane T. Soumagnac

9. Department of Particle Physics and Astrophysics, Weizmann Institute  
of Science, Rehovot, Israel

Maayane T. Soumagnac

## **Corresponding author**

Correspondence to [Ilaria Caiazzo](#).

## **Rights and permissions**

[Reprints and Permissions](#)

## About this article

### Cite this article

Caiazzo, I., Burdge, K.B., Fuller, J. *et al.* Publisher Correction: A highly magnetized and rapidly rotating white dwarf as small as the Moon. *Nature* **596**, E15 (2021). <https://doi.org/10.1038/s41586-021-03799-3>

- Published: 02 August 2021
- Issue Date: 26 August 2021
- DOI: <https://doi.org/10.1038/s41586-021-03799-3>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03799-3>

# Nature Outlook