

# nature

## TWO STRIKES

CRISPR screen identifies drug duo that improves treatment of liver cancer

### Coronavirus

Infectious secrets of the SARS-CoV-2 life cycle

### People power

Use locally managed areas to expand conservation toolkit

### Light touch

X-rays seen being bent around supermassive black hole

# Nature.2021.07.31

[Sat, 31 Jul 2021]

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

---

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

---

# Donation

---

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

# This Week

- **[Artificial intelligence in structural biology is here to stay](#)** [ 27 July 2021]  
Editorial • Machine learning will transform our understanding of protein folding. And it's essential that all data be open.
- **[The lack of people of colour in science images must be fixed](#)** [ 28 July 2021]  
Editorial • Archives, libraries, photo agencies and publishers need to do better to reflect science's true past and present.
- **[Remember Ebola: stop mass COVID deaths in Africa](#)** [ 27 July 2021]  
World View • The wealthy world must rally or nations will collapse across Africa.
- **[The surprising place where pot farming first blossomed](#)** [ 21 July 2021]  
Research Highlight • Humans first began cultivating cannabis, a source of both fibres and drugs, some 12,000 years ago.
- **[Rubbish-raiding parrots take lessons from co-conspirators](#)** [ 22 July 2021]  
Research Highlight • Sulphur-crested cockatoos are the first parrots known to have complex culture centred on food-gathering.
- **[A telescope watches as a black hole spews stuff into space](#)** [ 22 July 2021]  
Research Highlight • The trailblazing collaboration that first imaged a black hole turns its sights on a nearby black-hole jet.
- **[How to win big for the climate: rein in the ‘super polluters’](#)** [ 21 July 2021]  
Research Highlight • Just 5% of the world's power plants account for almost three-quarters of carbon emissions from electricity generation.
- **[China wastes almost 30% of its food](#)** [ 15 July 2021]  
Research Highlight • Out-of-home dining accounts for some of the nation's wasted food, but much more is lost during food storage and processing.
- **[The guts of a ‘bog body’ reveal sacrificed man’s final meal](#)** [ 20 July 2021]

Research Highlight • Tollund Man, who lived more than 2,000 years ago, ate well before he was hanged.

- **Pikas in high places have a winter-time treat: yak poo** [ 19 July 2021]

Research Highlight • Snacks of faeces help the pocket-sized mammals survive the cold and wind atop a vast plateau that abuts the Himalayas.

- EDITORIAL
- 27 July 2021

# Artificial intelligence in structural biology is here to stay

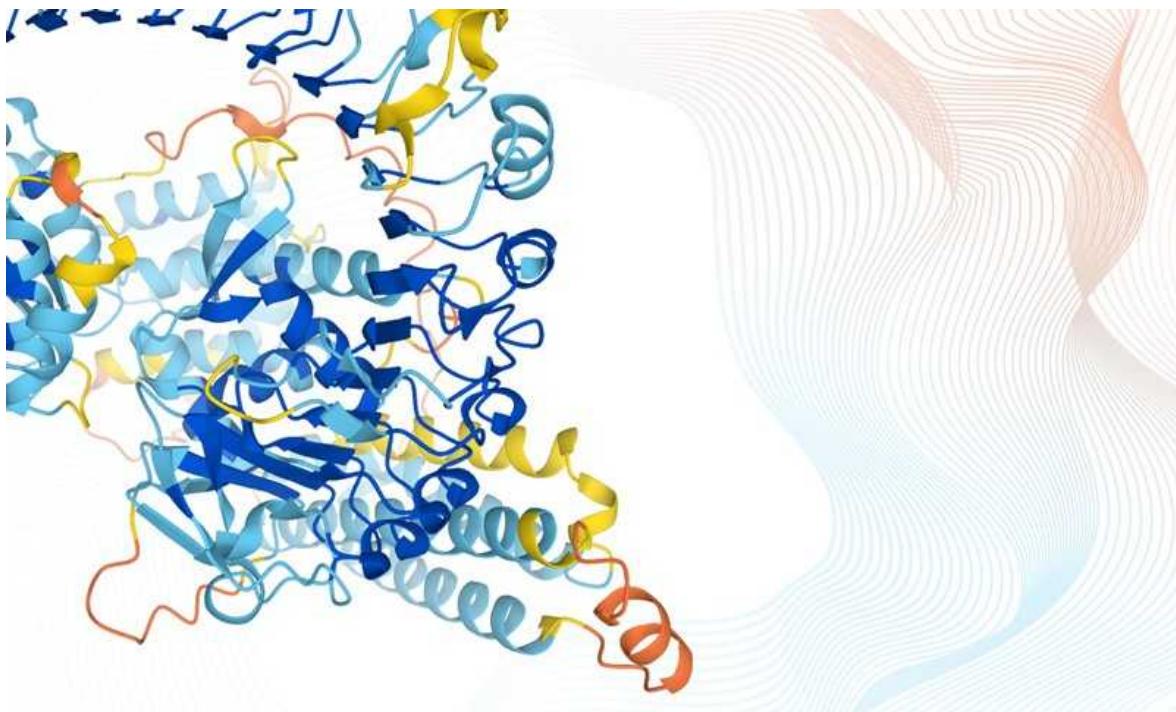
Machine learning will transform our understanding of protein folding. And it's essential that all data be open.





•

[Download PDF](#)



The AlphaFold machine-learning tool can predict 3D structures of full protein chains for 98% of human proteins

Credit: Karen Arnott/EMBL-EBI

“I didn’t think we would get to this point in my lifetime.” That’s how one research leader in structural biology responded to last week’s publication of research in which artificial intelligence (AI) was used to predict the structure of more than 20,000 human proteins, as well as that of nearly all the known proteins produced by 20 model organisms such as *Escherichia coli*, fruit flies and yeast, but also soya bean and Asian rice. That is a combined total of around 365,000 predictions<sup>1</sup>.

The data, publicly accessible for the first time (see <https://alphafold.ebi.ac.uk>), were released online on 22 July by researchers at DeepMind, a London-based AI company owned by Google’s parent company, Alphabet, and the European Bioinformatics Institute, based at the European Molecular Biology Laboratory (EBI-EMBL) near Cambridge, UK.



### DeepMind's AI predicts structures for a vast trove of proteins

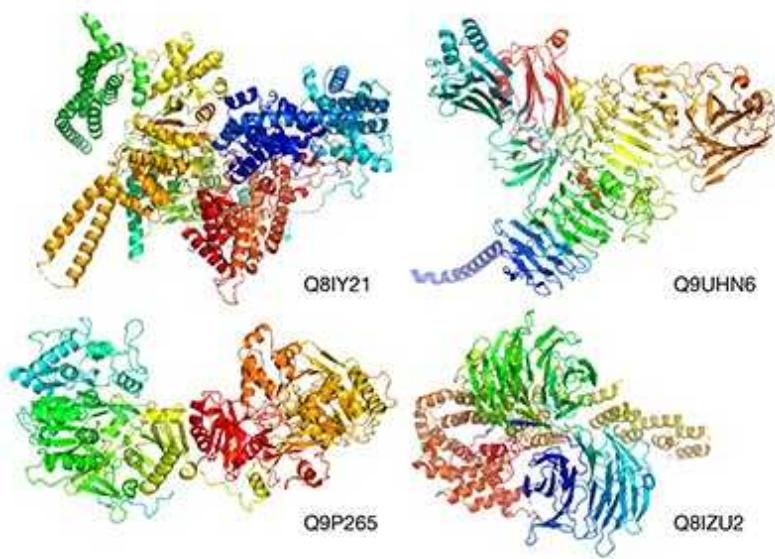
The DeepMind team developed a [machine-learning tool called AlphaFold](#). The team trained this program on DNA sequences, including their evolutionary history, and the already-known shapes of tens of the thousands of proteins contained in a public-access database of proteins hosted by the EBI-EMBL researchers. A week earlier, DeepMind also released the source code for AlphaFold and detailed how it was constructed<sup>2</sup>, at the same time that researchers from the University of Washington, Seattle, published details of another protein-structure prediction program — inspired by AlphaFold — called RoseTTAFold<sup>3</sup>.

The unveiling of this catalogue of predicted structures would not be nearly such good news were the data and the methodology not open and freely available. Structural biologists and other researchers are already starting to use AlphaFold to obtain more-accurate models for proteins that have been difficult or impossible to characterize by current experimental methods.

## **Speeding up structure prediction**

Predicting the 3D shape that proteins fold into has been one of biology's unsolved 'grand challenges' since the discovery in 1953 of the structure of DNA itself. Before AI, structure prediction from sequence was an intensely

time-consuming, not to say labour-intensive, process with little guarantee of getting an accurate result. The new data will still need to be validated and experimentally verified. But the AI tools can accurately predict protein structures in minutes to hours — compared with the months, or years, that it used to take to determine the structure of just one or two proteins. And that opens up possibilities for applications, for example in the engineering of enzymes to break down environmental pollutants such as microplastics.



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

Last week's breakthrough depended not just on the sharing of open data, but on advances in fundamental science and technology. Since the 1960s, structural biologists have worked on parallel approaches to understanding the science of protein folding. One involves piecing together the structures of proteins by understanding the underlying physical forces. Another attempts to predict the shapes by making comparisons with closely related proteins, using an organism's evolutionary history. And then there's been the all-important role of imaging technologies, starting with X-ray crystallography and now cryo-electron microscopy.

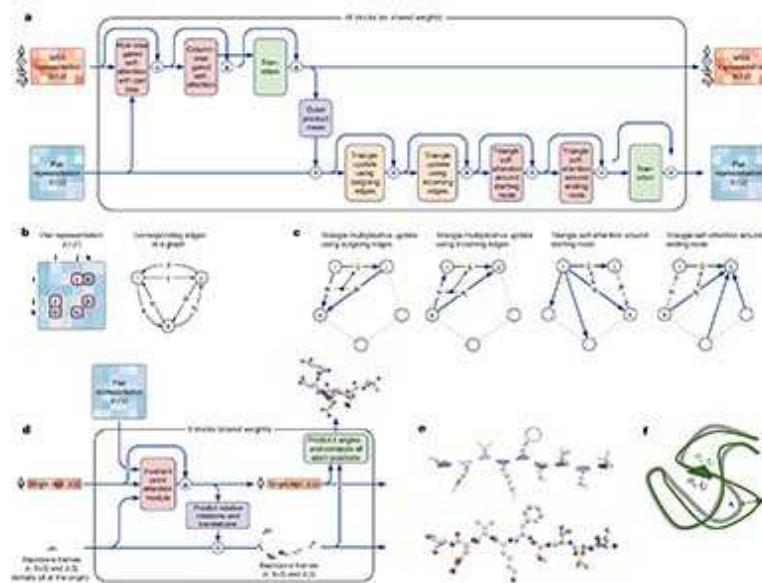
In the basic science of structural biology, key problems remain to be solved. Although AI in science and technology is good at producing accurate results, it doesn't (at least for now) explain how, or why, those results happened. The

teams at DeepMind, EBI-EMBL, the University of Washington and elsewhere should be congratulated for crucial breakthroughs. But there is still work to be done to unlock the science — the essential biology, chemistry and physics — of how and why proteins fold.

## Public and private

In terms of significance, some are comparing the latest advances to the first draft human genome sequence 20 years ago. And it's true that there are comparisons to be made. Both the Human Genome Project and DeepMind's catalogue of human protein-structure predictions equip their fields with a tool that is set to markedly accelerate discovery.

The human genome's first draft was the result of a race. Solving protein folding has also benefited from a kind of competition — an annual event called the Critical Assessment of Protein Structure Prediction (or CASP), which has been essential to getting a result.



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

Today's research teams — just like those involved in early genome sequencing — needed open access to data. In making the data and the methodology openly available to all, DeepMind now sets a benchmark that

will make it harder for other corporations in this space, such as Facebook and Microsoft, to continue arguing for proprietary data.

And so, what of the future? Over the past week, *Nature* interviewed nearly a dozen researchers in the field. The consensus is that it's too early to predict exactly what impact the application of AI in the life sciences will have, except that any impact will be transformative.

Accurately predicting how AI will change biology needs good training data, which we don't yet have. But in AI, the structural-biology research community — and its collaborators in other fields — have a vast trove of fresh data. In addition to its research and data, AI provides a window into models for research organization and management that universities should study. For today's researchers, and those in future generations, there is much work to follow up on.

*Nature* **595**, 625-626 (2021)

doi: <https://doi.org/10.1038/d41586-021-02037-0>

## References

1. 1.

Tunyasuvunakool, K. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-03828-1> (2021).

[Article](#) [Google Scholar](#)

2. 2.

Jumper, J. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-03819-2> (2021).

[Article](#) [Google Scholar](#)

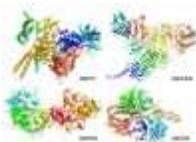
3. 3.

Baek, M. *et al.* *Science* <https://doi.org/10.1126/science.abj8754> (2021).

[Article](#) [Google Scholar](#)

[Download references](#)

## Related Articles



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



- [Read the paper: Highly accurate protein structure prediction with AlphaFold](#)
- [Read the paper: Accurate prediction of protein structures and interactions using a three-track neural network](#)



- [DeepMind's AI predicts structures for a vast trove of proteins](#)



- [DeepMind's AI for protein structure is coming to the masses](#)

## Subjects

- [Machine learning](#)

- [Proteomics](#)
- [Structural biology](#)

## nature careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



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



- [Read the paper: Highly accurate protein structure prediction with AlphaFold](#)
- [Read the paper: Accurate prediction of protein structures and interactions using a three-track neural network](#)



- [DeepMind's AI predicts structures for a vast trove of proteins](#)



- [DeepMind's AI for protein structure is coming to the masses](#)

## Subjects

- [Machine learning](#)
- [Proteomics](#)
- [Structural biology](#)

- EDITORIAL
- 28 July 2021

# The lack of people of colour in science images must be fixed

Archives, libraries, photo agencies and publishers need to do better to reflect science's true past and present.





•

[Download PDF](#)



Unlike many other Black scientists from history, high-quality images of agricultural scientist George Washington Carver — shown here with students at the Tuskegee Institution in Alabama — are readily available. Credit: Bettmann/Getty

Last month, *Nature* published a Comment article on how researchers and communities [helped each other during a water crisis in Flint, Michigan](#). While sourcing pictures for the article, *Nature*'s photo editor discovered that there are few images available of the people involved, many of whom are Black.

Recently, we also needed an image of the physicist Elmer Imes, who, in 1918, became only the [second African American to be awarded a PhD in physics in the United States](#). His doctoral work provided early evidence of the quantum behaviour of molecules. But university archives that *Nature* contacted did not have a copy of his photograph. Commercial photography

agencies also had nothing. Low-resolution, grainy images do exist, but, shockingly, even the US Library of Congress in Washington DC — which holds images of many important scientists from the nation’s history — does not have a photograph. However, such images are available for a number of notable white scientists from Imes’s time.



### Community-academic partnerships helped Flint through its water crisis

This is far from an isolated case. *Nature* often illustrates articles reporting on communities and countries that are under-represented in science using generic images, in part because universities, national libraries and commercial photo agencies hold relatively few images of people from such communities.

Although we do our best to work with generic images in such situations, they tend to be less compelling than pictures showing real scientists doing real research. When we do use photographs of the researchers themselves, this can boost the impact of the article — attracting greater social media attention, for example — which, in turn, can benefit those individuals and their work.

Systemic racism and science’s diversity deficit extend to images, creating a distorted and exclusionary picture of science’s past and present. This is an issue that needs attention, and there are several potential ways to rectify it.

When it comes to photographs of living scientists from under-represented communities, it can be surprisingly difficult to obtain high-resolution images of a standard that international publishers generally require. The resolution of images on institutional websites is often insufficient. But there's a relatively straightforward fix, at least for some institutions: where resources allow it, universities could ensure that they can provide appropriate access to high-resolution images of their researchers, if individuals have consented to this.



### [Katherine Johnson's memoir charts her bold trajectory to NASA and beyond](#)

A second and related problem — the lack of high-quality historical images, particularly of people of colour — is also not insurmountable. For example, such images might be available in university records or archives, and, if not, these institutions will often know how to find such images or will have access to ways of improving the quality of the images they do have. National libraries need to work with universities to identify and publish images of notable researchers.

Arguably the most difficult, although no less important, task will be to bring about change in the commercial photography agencies. These agencies are a crucial source of images for media organizations. At *Nature*, we use them all the time, and credit them next to the images. But, more often than not, our searches for photos of particular Black scientists and scientists of other

marginalized ethnicities yield negative results, and we are compelled to fall back on generic images of people modelling a generic scene, instead of photos of the scientists themselves. In some cases, photos do exist, but are incorrectly captioned or are not tagged with appropriate keywords, meaning they cannot be found.

*Nature* approached six large agencies and asked whether they have a dedicated staff member — or an organized process — for improving diversity in their science-related images. Representatives of three agencies responded. None has such a person. One photo repository acknowledged that Black people are not represented in its images of clinical medicine, and that it is actively working to correct this. Tracking diversity needs to be a priority for these agencies.

## Community action

Science publishers and media outlets — including *Nature* — also have a responsibility to do more to ensure we are publishing images of the people we feature. And we need to commission more photographers from the communities that we're writing about, something *Nature* has particularly tried to address in our weekly article, [Where I Work](#).

Universities, libraries, publishers and photo agencies — the organizations that hold the keys to so much of the world's photography — must all take steps to diversify our imagery. Science's historical record will remain incomplete while it is missing pictures of people who have contributed to discovery and invention. Such efforts are also essential to make research more welcoming for people from under-represented communities, and to ensure that future generations of researchers reflect those that science has often failed to attract in the past.

*Nature* 595, 626 (2021)

doi: <https://doi.org/10.1038/d41586-021-02036-1>

## Related Articles



- [Community-academic partnerships helped Flint through its water crisis](#)



- [Katherine Johnson's memoir charts her bold trajectory to NASA and beyond](#)



- [The time tax put on scientists of colour](#)



- [Racism is magnifying the deadly impact of rising city heat](#)



- [Science diversified: Black researchers' perspectives](#)

## Subjects

- [History](#)
- [Publishing](#)
- [Society](#)

**nature** careers

[Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- [Community–academic partnerships helped Flint through its water crisis](#)



- Katherine Johnson's memoir charts her bold trajectory to NASA and beyond



- The time tax put on scientists of colour



- Racism is magnifying the deadly impact of rising city heat



- Science diversified: Black researchers' perspectives

## Subjects

- History
- Publishing
- Society

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02036-1>

- WORLD VIEW
- 27 July 2021

## Remember Ebola: stop mass COVID deaths in Africa



The wealthy world must rally or nations will collapse across Africa.

- [Mosoka Fallah](#) 0

1. [Mosoka Fallah](#)

1. Mosoka Fallah is the former director-general of the National Public Health Institute of Liberia and is a contractor for the US National Institute of Allergy and Infectious Disease.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

## [Download PDF](#)

As I write, mass fatalities from COVID-19 have begun in Africa. Overall, only 1.1% of people in low-income countries have received at least one vaccine dose. In my own Liberia, the proportion of people who have had both doses is just 0.18%. Recorded cases in the country jumped by 144% between 1 June and 21 July. This is probably an undercount.

The disease is more deadly here than elsewhere: the current case fatality rate in Africa is 18% higher than the global average. Rates are worse among the critically ill. About half of those who are admitted to intensive care die

within 30 days ([African COVID-19 Critical Care Outcomes Study](#). *Investigators Lancet* **397**, 1885–1894; 2021).

Enough statistics. The number of graves being dug each day in Liberia recalls the 2014 Ebola catastrophe. Like Ebola, COVID-19 infects and kills a disproportionate number of health-care workers, and Liberia already has far too few.

My social media used to be alive with wishes for happy birthdays and anniversaries. Now it is inundated with “rest in peace”. The moment for high-income nations to come to the aid of Africa, a continent they depend on for an abundance of human and natural resources, has almost passed.

I was head of case detection in Liberia during the Ebola crisis in 2014, when the world took eight months to act. I remember weeping in the homes of entire families swept away by the disease. Now I spend my days advocating that global leaders avoid repeating the mistakes of that outbreak. Lately my focus has been on increasing West Africa’s access to equipment needed to safely administer oxygen to patients. Oxygen supplies are also a concern: Sierra Leone, for instance, has only 2 oxygen-production plants for 7.8 million people. Uganda has the capacity to produce about 3,000 oxygen cylinders a day, but the likelihood is that in the coming months it will need many, many more.



## How COVID spurred Africa to plot a vaccines revolution

Let me tell you about two brothers who attended church with my aunt. The younger, a family man and valued member of his community, fell ill on a Monday. He was incorrectly diagnosed with and treated for malaria and typhoid fever, which are common here. Despite government warnings that cases of COVID-19 were rapidly increasing, neither the man nor his family or caregivers realized this risk until Wednesday, when he had difficulty breathing and was rushed to a private hospital. With all of the oxygen cylinders in use by the dozen people with COVID-19 already there, his family watched, helpless, as he died, unable to breathe. They approached funeral home after funeral home to find one with space for his body. The man's eldest brother, who had retired to Liberia after a career in the United States, agreed to cover all the expenses, but also died from COVID-19 even before the younger man's funeral.

Africa was largely spared by the pandemic in 2020, but not this year. We lack vaccines and we are gasping for air.

Had the global north delivered on its promises to provide vaccines, many deaths could have been averted. By the beginning of July, fewer than 50 million of the 700 million doses that the COVAX initiative promised to deliver to Africa this year had arrived. They might yet arrive by December, but, by then, it will be too late for many.

During the worst period of the Ebola outbreak for Liberia, in July, August and September 2014, I saw people die in the streets. The world largely left us to combat a global health threat alone. One Ebola treatment unit built for 34 people had to serve 74 patients. The sick would wait for someone to die to free up space. Yet when a handful of Ebola infections reached developed countries, some US\$3.5 billion poured in to fight the outbreak. This enormous sum came too late to save many of the 11,300 who died in West Africa. Less than a decade later, here we are again.



## What it will take to vaccinate the world against COVID-19

Last October, before vaccines were available, the head of the Africa Centres for Disease Control and Prevention issued a call in these pages to ensure that provision was made to get vaccines to the poorest countries ([J. N. Nkengasong et al. \*Nature\* 586, 197–199; 2020](#)). In January, as rich countries were rolling out vaccine distribution, I helped to organize a call for action with 30 Ebola veterans and 81 global health professionals who, recognizing the threat COVID-19 poses to Africa, petitioned the World Health Assembly to act to secure vaccines. Why are we still waiting?

As I and my peers across African health ministries see it, rich countries are hoarding vaccines, allowing doses to expire while unvaccinated people who want to be immunized die.

Let me say this as an African: our world as we know it is on the brink; we face massive death tolls, and the collapse of economies and nations. What is the real meaning of humanity? For all lives to be given the same value, irrespective of geography or economy.

In the aftermath of the Second World War, the United States developed the Marshall Plan to enable devastated countries in Western Europe to recover. It was an inspiring moment in human history. The current human calamity must be stopped with a new Marshall Plan, whereby prosperous nations

freely share vaccines, manufacturing capacity and resources — if not for the sake of their consciences, then for health security.

Regions where COVID-19 cases are allowed to soar are the places where the next variant will emerge. That could undo all the advances made with the vaccine roll-out in developed countries.

*Nature* **595**, 627 (2021)

*doi:* <https://doi.org/10.1038/d41586-021-01964-2>

## Competing Interests

The author declares no competing interests.

## Related Articles

-  [How poorer countries are scrambling to prevent a coronavirus disaster](#)
-  [How COVID spurred Africa to plot a vaccines revolution](#)
-  [What it will take to vaccinate the world against COVID-19](#)

## Subjects

- [Vaccines](#)
- [SARS-CoV-2](#)
- [Society](#)
- [Developing world](#)

## nature careers

### [Jobs](#) >

- [\*\*Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics\*\*](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [\*\*FACULTY POSITION IN MICROBIOME BIOCHEMISTRY\*\*](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [\*\*Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology\*\*](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

-  [How poorer countries are scrambling to prevent a coronavirus disaster](#)
-  [How COVID spurred Africa to plot a vaccines revolution](#)
-  [What it will take to vaccinate the world against COVID-19](#)

## Subjects

- [Vaccines](#)
- [SARS-CoV-2](#)
- [Society](#)
- [Developing world](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01964-2>



Domesticated cannabis in Qinghai province, China. Credit: Guangpeng Ren

Genomics

21 July 2021

# **The surprising place where pot farming first blossomed**

Humans first began cultivating cannabis, a source of both fibres and drugs, some 12,000 years ago.

•



•

Early humans domesticated cannabis about 12,000 years ago in what is now China — not in Central Asia, as had been widely thought.

Guangpeng Ren at Lanzhou University in China, Luca Fumagalli at the University of Lausanne in Switzerland and their colleagues sequenced the genomes of 82 cannabis plants (*Cannabis sativa*, the source of marijuana) from around the world. The team's analysis of those genomes and 28 others that had been sequenced previously revealed that they fell into four genetically distinct categories: a primordial group, two groups bred for drugs, and a hemp group, which comprises plants grown for their fibre.

The hemp and drug types diverged from ancestral members of the primordial group about 12,000 years ago, suggesting that humans have long cultivated the plant for use in textiles and for medicinal purposes. In the hemp type, domestication led to changes in genes responsible for branch formation and the timing of blooming; in the drug types, genes responsible for potency underwent change.

Legal restrictions have hindered studies of the plant's history, and the authors hope that these findings will inspire future research.

*[Sci. Adv. \(2021\)](#)*

- [Genomics](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01982-0>

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



A sulphur-crested cockatoo grabs a snack by walking along the edge of a rubbish bin to open the lid. Credit: Barbara Klump/Max Planck Institute of Animal Behavior

Animal behaviour

22 July 2021

## Rubbish-raiding parrots take lessons from co-conspirators

Sulphur-crested cockatoos are the first parrots known to have complex culture centred on food-gathering.





•

Parrots in eastern Australia have picked up a variety of bin-diving techniques by watching their neighbours — the first evidence that these large birds have distinct food-gathering cultures.

Over the last three years, sulphur-crested cockatoos (*Cacatua galerita*) have been seen flipping open household bins in the suburbs around Sydney to hunt for food in the litter. Barbara Klump and Lucy Aplin at the Max Planck Institute of Animal Behavior in Radolfzell, Germany, and their colleagues surveyed residents in 44 suburbs in the area and compiled reports of cockatoos prying open rubbish bins. Most sightings involved multiple birds.

The authors marked and watched almost 500 birds in 3 neighbourhoods. They observed how the parrots shuffled over the bins, and the precise beak and foot movements that the birds used to lift the lids.

Birds in one region tended to have similar sequences of perching and lid-flipping, suggesting that they learned the routine from each other, the authors say. The behaviour probably originated in 3 suburbs and spread to 41 others.

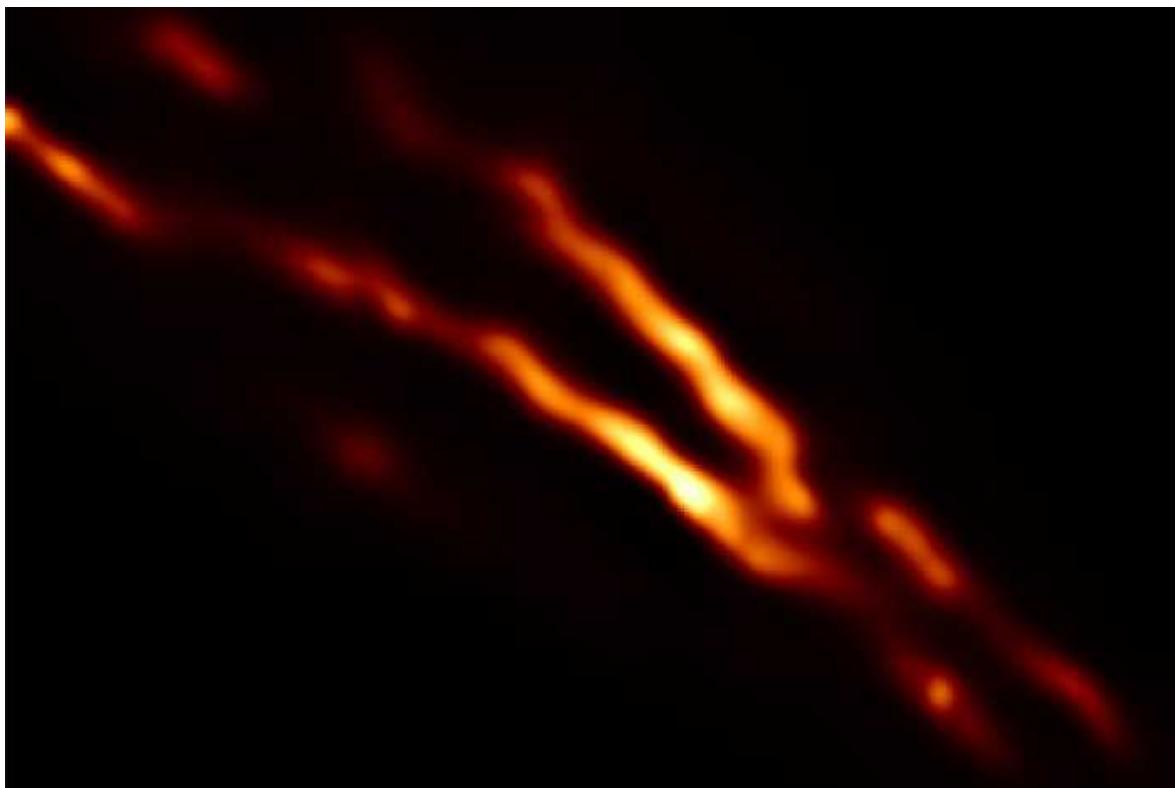
[Science \(2021\)](#)

- [Animal behaviour](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01999-5>

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



A plume issuing from the gigantic black hole at the centre of the Centaurus A galaxy is fed by this cone-shaped structure. Credit: M. Janssen *et al.*/*Nat. Astron.* ([CC BY 4.0](#))

Astronomy and astrophysics

22 July 2021

## A telescope watches as a black hole spews stuff into space

The trailblazing collaboration that first imaged a black hole turns its sights on a nearby black-hole jet.





•

The galaxy called Centaurus A is one of the Milky Way's noisiest neighbours. Now, astronomers have gotten a close look at a prime reason for the racket — a radiation-generating jet of material kicked out by the supermassive black hole at the galaxy's heart.

Michael Janssen at the Max-Planck Institute for Radio Astronomy in Bonn, Germany, and his colleagues captured a detailed picture of the jet close to its source. To do so, they used the Event Horizon Telescope array, the global network of radio telescopes behind the first image of a black hole, which was published in 2019.

The resulting picture of Centaurus A's jet has 16 times the resolution of the previous highest-resolution image. The new image revealed that the base of the jet is a cone of ejected matter with bright edges, similar to what has been seen in a range of other black-hole systems. The findings support the idea that the black hole in Centaurus A — a mid-size model by cosmic standards — behaves similarly to much smaller and much bigger ones.

[Nature Astron. \(2021\)](#)

- [Astronomy and astrophysics](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01985-x>

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



The Niederaussem facility near Bergheim, Germany, ranked among the top ten power plants for carbon emissions in 2018. Credit: Oliver Berg/DPA/AFP/Getty

Climate change

21 July 2021

## How to win big for the climate: rein in the ‘super polluters’

Just 5% of the world’s power plants account for almost three-quarters of carbon emissions from electricity generation.





•

A crackdown on a limited number of ‘hyper-emitting’ power plants could yield outsize cuts in the carbon emissions resulting from global electricity generation.

Don Grant and his colleagues at the University of Colorado Boulder, combed an inventory of more than 29,000 fossil-fuel power plants across 221 countries to identify the world’s largest polluters as of 2018. They then calculated the potential emissions reductions that could be attained if the worst offenders boosted their efficiency, switched to lower-carbon fuels or implemented carbon-capture technologies.

The team found that extreme emitters — power plants that ranked in the top 5% in terms of climate pollution — were responsible for 73% of global emissions from electricity generation, and tended to be less efficient than average power plants in their home countries. The team's calculations suggest that emissions would fall by roughly 25% if those polluters boosted their efficiency to match the global average. Switching their fuel source from coal or oil to natural gas could cut global emissions by almost 30%, and implementing carbon-capture technologies could cut emissions nearly in half.

Targeting the world's super polluters could yield disproportionate national and global benefits, the researchers argue.

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

- [Climate change](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01983-z>

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



A worker sweeps up watermelons at a wholesale market in Chongqing, China. Waste from households accounts for only a small fraction of China's discarded food. Credit: China Photos/Getty

Sustainability

15 July 2021

## China wastes almost 30% of its food

Out-of-home dining accounts for some of the nation's wasted food, but much more is lost during food storage and processing.





•

More than a quarter of food produced for human consumption in China gets lost along the supply chain or lands on rubbish heaps.

Liu Gang at the University of Southern Denmark in Odense and his colleagues looked at data from field surveys and reviewed the published literature to assess the ultimate fate of food produced in the nation of 1.4 billion people. They estimate that around 350 million tonnes of China's annual farm product, or about 27%, is discarded by retailers, restaurants or consumers — or is ruined and disposed of before reaching retail level.

Although almost half of the loss occurs during food storage and processing, out-of-home eating, including at food stalls, restaurants and canteens,

produces some 45 million tonnes of food waste each year, the researchers found.

Food waste on such a large scale threatens environmental and sustainability goals, the scientists say. To reduce waste, the authors suggest, among other steps, that rural Chinese households use more-efficient storage systems and that urban restaurants reduce portions and encourage patrons to take their leftovers.

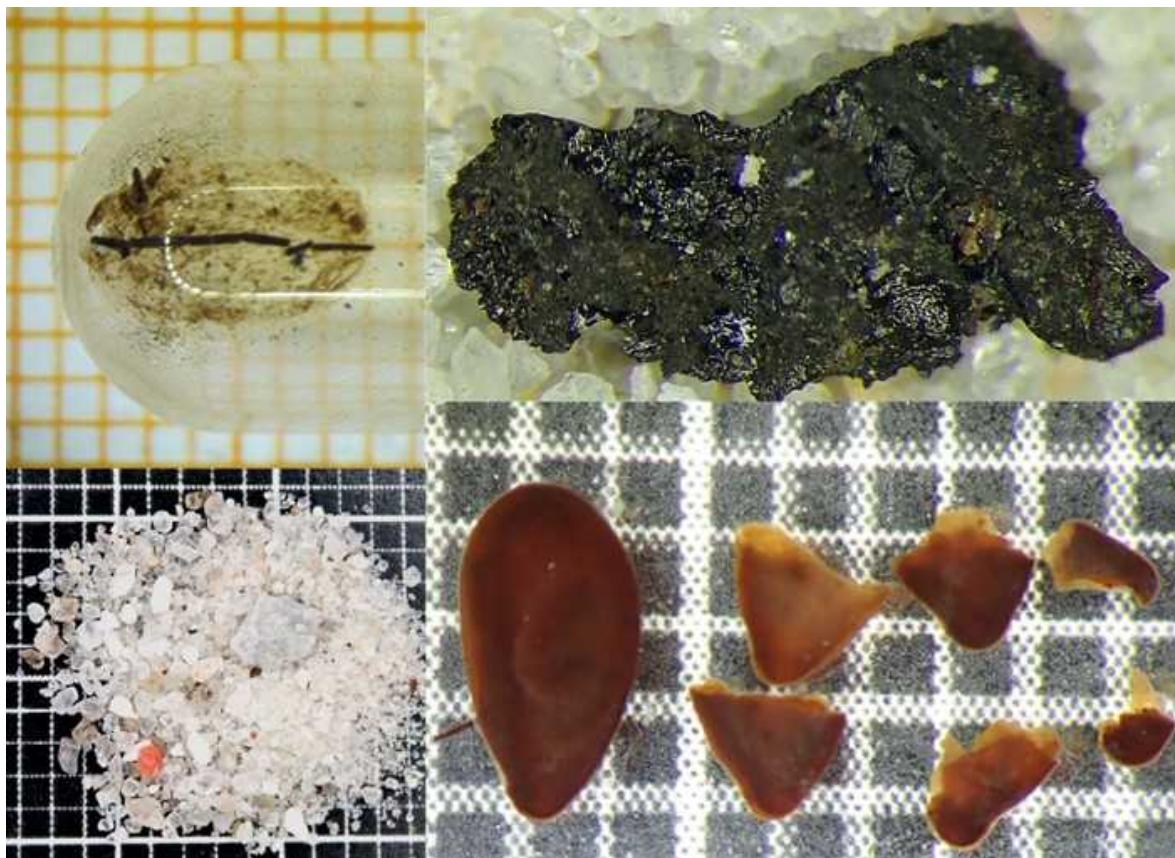
[Nature Food \(2021\)](#)

- [Sustainability](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01963-3>

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



The intestinal contents of a man killed in a prehistoric ritual (clockwise from upper left): barley, charred food that had been encrusted in a clay pot, flax seeds and sand. Credit: Peter Steen Henriksen, the Danish National Museum

Archaeology

20 July 2021

## The guts of a ‘bog body’ reveal sacrificed man’s final meal

Tollund Man, who lived more than 2,000 years ago, ate well before he was hanged.





•

Some 2,400 years ago, a man was killed and placed in a bog in modern Denmark — probably as some kind of ritual sacrifice. Now, researchers have reconstructed the man's last meal, down to its recipe.

The naturally preserved remains of Tollund Man, as he has come to be known, were found in peatlands in northern Denmark in 1950. Nina Helt Nielsen at the Silkeborg Museum in Denmark and her colleagues re-examined Tollund Man's gut contents and found that, some 12 to 24 hours before his death, he ate a hearty meal.

The meal was cooked in a clay pot and consisted of a porridge containing barley and flax, and probably fish. It also included seeds from pale persicaria

(*Persicaria lapathifolia*), a weed that grew in barley fields and whose seeds were typically removed during threshing and cleaning the grain. These seeds seem to have been added back after the threshing process, which might signal that the porridge was eaten during a human sacrifice ritual.

The presence of proteins and eggs from intestinal worms indicates that Tollund Man was infected with parasites, most likely as a result of consuming contaminated food or water.

[Antiquity \(2021\)](#)

- [Archaeology](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01984-y>.

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



Plateau pikas ‘kiss’ during the growing season, when their metabolic rates are much higher than during the winter. Credit: Zhou Jinshuai/Xinhua/Alamy

Zoology

19 July 2021

## Pikas in high places have a winter-time treat: yak poo

Snacks of faeces help the pocket-sized mammals survive the cold and wind atop a vast plateau that abuts the Himalayas.





•

To survive the harsh winters of their high-altitude home, a fluffball-like animal called the pika becomes strategically lazy — and gorges on yak dung.

The plateau pika (*Ochotona curzoniae*), a round-eared relative of rabbits, lives on China's Qinghai–Tibet plateau, where the air is thin and winter temperatures often drop below  $-30^{\circ}\text{C}$ . To understand how the pika, which does not hibernate, survives the cold season, Yanming Zhang at the Northwest Institute of Plateau Biology, Chinese Academy of Sciences in Qinghai and colleagues intensively studied the mammals over 20 field trips between 2007 and 2020.

The researchers filmed pikas and implanted them with temperature sensors. The team also injected the animals with water bearing a distinctive isotopic signature to assess their metabolic rate.

The results show that on average, plateau pikas can lower their daily energy expenditure by almost 30% in the winter. They also rely on an unexpected but nutritious and easy-to-digest food that they can access without expending too much precious energy: domestic yak faeces, which local people also use as fuel.

[Proc. Natl. Acad. Sci. USA \(2021\)](#)

- [Zoology](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01978-w>

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

# News in Focus

- **Delta's success, satellite swarms and a 'super' COVID antibody** [ 28 July 2021]  
News Round-Up • The latest science news, in brief.
- **The US is boosting funding for research monkeys in the wake of COVID** [ 15 July 2021]  
News • The investments are promising but won't fix the primate shortage, experts say.
- **DeepMind's AI predicts structures for a vast trove of proteins** [ 22 July 2021]  
News • AlphaFold neural network produced a 'totally transformative' database of more than 350,000 structures from Homo sapiens and 20 model organisms.
- **Massive DNA 'Borg' structures perplex scientists** [ 16 July 2021]  
News • Researchers say they have discovered unique and exciting DNA strands in the mud — others aren't sure of their novelty.
- **China launches world's largest carbon market: but is it ambitious enough?** [ 20 July 2021]  
News • Experts welcome the trading scheme, but question whether it is up to the task of helping China achieve its climate goals.
- **Should children get COVID vaccines? What the science says** [ 20 July 2021]  
News Explainer • With vaccination campaigns under way in some countries while others weigh the options, Nature looks at the evidence for vaccinating younger people.
- **Deaths from COVID 'incredibly rare' among children** [ 15 July 2021]  
News • Studies find that overall risk of death or severe disease from COVID-19 is very low in kids.
- **How the coronavirus infects cells — and why Delta is so dangerous** [ 28 July 2021]  
News Feature • Scientists are unpicking the life cycle of SARS-CoV-2 and how the virus uses tricks to evade detection.

---

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

- NEWS ROUND-UP
- 28 July 2021

# Delta's success, satellite swarms and a 'super' COVID antibody

The latest science news, in brief.





•

[Download PDF](#)



A queue at a hospital in Surabaya, Indonesia, that treats people with COVID-19. Indonesia has been hard hit by the Delta variant. Credit: Juni Kriswanto/AFP/Getty

## How the Delta coronavirus variant achieves its ultrafast spread

Since it first appeared in late 2020, the Delta variant of SARS-CoV-2 has become the predominant strain in much of the world. Researchers might now know why Delta has been so successful: people infected with it produce many more virus particles than do those infected with the original version of SARS-CoV-2, making it very easy to spread.

According to current estimates, the Delta variant could be more than twice as transmissible as the original strain. To find out why, epidemiologist Jing Lu at the Guangdong Provincial Center for Disease Control and Prevention

in Guangzhou, China, and his colleagues tracked 62 people who were quarantined after exposure to COVID-19 and who were some of the first people in mainland China to become infected with the Delta strain.

The team tested study participants' 'viral load' — a measure of the density of viral particles in the body — every day throughout the course of infection to see how it changed over time. They then compared participants' infection patterns with those of 63 people who contracted the original SARS-CoV-2 strain in 2020.

The virus was first detectable in people with the Delta variant four days after exposure, compared with an average of six days among people with the original strain, suggesting that Delta replicates much faster ([B. Li et al. Preprint at medRxiv https://doi.org/gk78tn; 2021](#)). People infected with Delta also had viral loads up to 1,260 times higher than did people infected with the original strain.

This combination of a high number of virus particles and a short incubation period could [explain Delta's heightened transmissibility](#), scientists say.

Other questions about the variant remain unanswered. It's still unclear, for instance, whether Delta is more likely to cause severe disease than the original strain, and how good it is at evading the immune system.



The firm SpaceX launched a batch of Starlink satellites into orbit around Earth on 6 January 2020.Credit: Paul Hennessy/NurPhoto/Getty

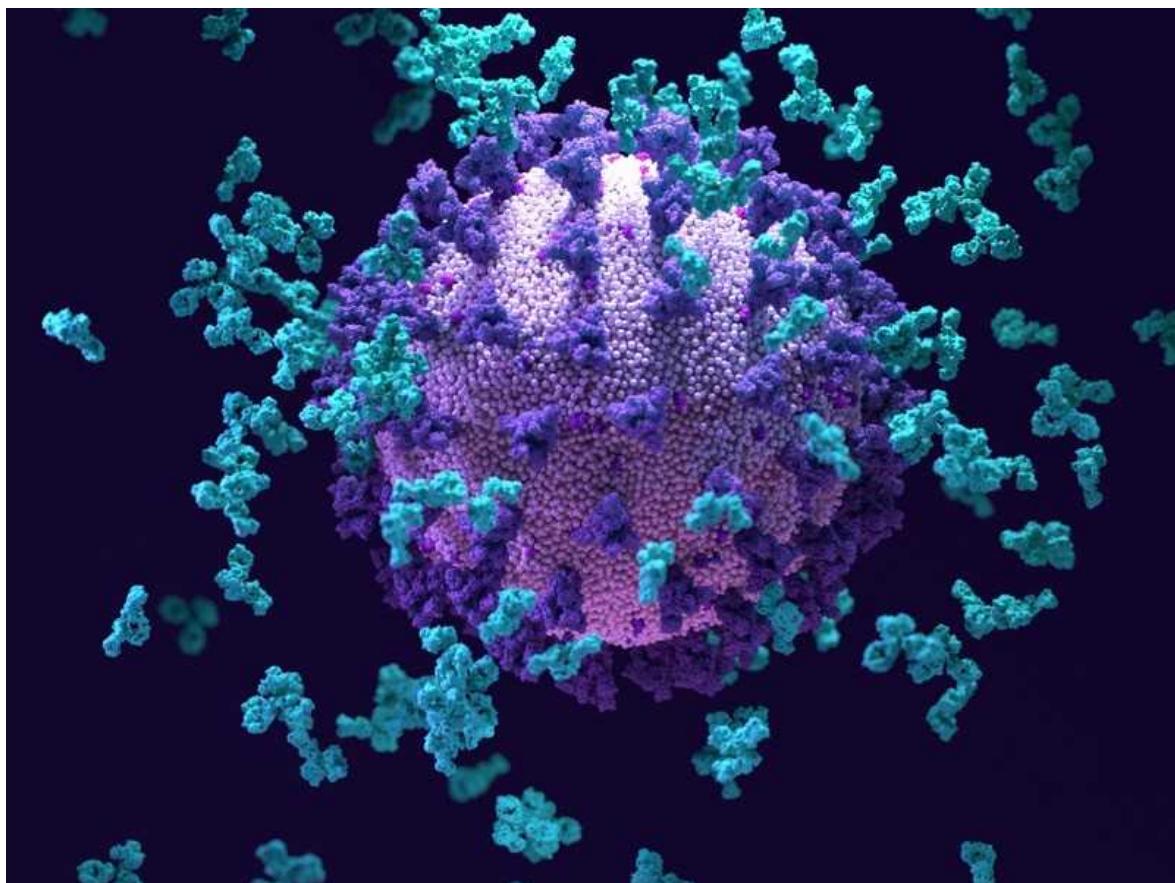
## Astronomers want global debate on satellite swarms

Aerospace companies have launched about 2,000 Internet satellites into orbit around Earth over the past 2 years, sparking [concerns among astronomers and other skygazers](#) over interference with observations of the night sky.

In what would be the biggest international step yet towards addressing these worries, diplomats at a United Nations forum next month might discuss whether humanity has a right to ‘dark and quiet skies’.

Astronomers have been working through the International Astronomical Union to raise awareness of how these satellite networks, or ‘megaconstellations’, are affecting scientists and members of the public. They are pushing for the UN Committee on the Peaceful Uses of Outer Space to take up the topic at its next meeting, which begins on 25 August.

They say the goal is not to pit astronomers against satellite companies, but to develop a vision of how to use the shared realm of outer space fairly. “The consensus has to come from all the countries,” says Connie Walker, an astronomer at NOIRLab based in Tucson, Arizona, an umbrella organization for several US-funded observatories.



Antibodies (light blue; artist's impression) swarm around a SARS-CoV-2 particle. Credit: Design Cells/SPL

## ‘Super antibody’ fights off many coronaviruses

Scientists have [uncovered an antibody](#) that can fight off not only a wide range of SARS-CoV-2 variants, but also closely related coronaviruses ([T. N. Starr et al. \*Nature\* https://doi.org/gpqn; 2021](#)).

Researchers examined antibodies isolated from people who had been infected with either SARS-CoV-2 or its close relative SARS-CoV. One

antibody, S2H97, stood out for its capacity to adhere to a protein fragment known as the receptor binding domain on multiple SARS-CoV-2 variants and dozens of coronaviruses belonging to a group called the sarbecoviruses.

S2H97 was able to prevent a range of SARS-CoV-2 variants and other sarbecoviruses from spreading among cultured cells. It was also powerful enough to protect hamsters against SARS-CoV-2 infection. A closer examination of its molecular structure revealed that it targets a previously unseen and well-hidden region of the binding domain (pictured, artist's impression of antibodies swarming a viral particle).

Molecules targeting this binding-domain region could provide protection against multiple viruses, and might one day be used to produce broad-ranging vaccines and treatments, the researchers say.

*Nature* **595**, 631 (2021)

doi: <https://doi.org/10.1038/d41586-021-02038-z>

## Subjects

- [Astronomical instrumentation](#)
- [SARS-CoV-2](#)

**nature** careers

## [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [\*\*Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology\*\*](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Subjects

- [\*\*Astronomical instrumentation\*\*](#)
- [\*\*SARS-CoV-2\*\*](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02038-z>

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

- NEWS
- 15 July 2021

# The US is boosting funding for research monkeys in the wake of COVID

The investments are promising but won't fix the primate shortage, experts say.

- [Nidhi Subbaraman](#)

1. Nidhi Subbaraman  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



Rhesus macaques are the monkey most often used in US biomedical research. Credit: Bernard Castelein/NaturePL

The US government is investing heavily to breed more monkeys at the national facilities that house primates for biomedical research, *Nature* has learnt. The goal is to offset an ongoing shortage of these animals, which grew worse in 2020 as scientists tested scores of COVID-19 vaccines and treatments on primates before trials began in people.

To make room for more monkeys, the US National Institutes of Health (NIH) has invested about US\$29 million over the past two years in refurbishing housing, building outdoor enclosures and making other infrastructure improvements at the US National Primate Research Centers (NPRCs), which it funds. The agency is expecting to spend another \$7.5 million or so by October. And US President Joe Biden's administration has proposed investing even more: citing the pandemic, it suggests a 27% increase in funding for the NPRCs in its budget request for fiscal year 2022. If approved by Congress, that would add \$30 million for the centres.



## [US lawmakers propose plan to reduce primate research at National Institutes of Health](#)

“We have been making investments to bring the levels up and to plan for the future,” says James Anderson, director of the NIH Division of Program Coordination, Planning, and Strategic Initiatives in Bethesda, Maryland. “What happens if [a pandemic] happens again, with another virus in three years? We want to be ready for that.”

US scientists use non-human primates, most commonly rhesus macaques (*Macaca mulatta*), to study a range of medical conditions, including infectious diseases. Genetically and physiologically similar to people, primate models offer a way to run tests and experiments before human trials or when human trials are not possible. In 2019, US scientists used 68,257 non-human primates in research, [according to the US government](#).

“A couple of years ago, we were feeling the pinch,” says Nancy Haigwood, director of the Oregon NRPC in Beaverton, which houses about 5,000 non-human primates. But owing to the pandemic, “we are truly out of animals”, she says. “We’re turning away everyone.”

Experts say the new financing is a step in the right direction, but it will take a bigger investment to fully address the shortage.

“It’s very encouraging to see the Biden administration make an investment in the future of primate research in the US,” says Matthew Bailey, president of the National Association for Biomedical Research (NABR) in Washington DC, a group that advocates support for animal research. But he adds that it takes time to establish monkey colonies, and yields could be years away. “It’s a smart decision, but it isn’t like flipping a switch — it’s not going to change overnight.”

## A shortage exacerbated

Research demand for non-human primates in the United States has soared in the past five years. A multitude of NIH grants awarded in 2016 to study HIV/AIDS sparked a rise in the use of rhesus macaques, [according to a 2018 report](#). Non-human primates are expensive to house and feed, and budget caps meant the NRPCs didn’t have the funding to build up infrastructure to expand their capacity, the report said. The analysis warned that demand for rhesus macaques, marmosets and possibly baboons could overwhelm capacity at the centres in the coming years.

The issue has drawn public interest — and opposition. Animal-rights groups seeking to stop the use of animals in research have collided with scientists and funders who insist that experiments in model species are necessary to treat and understand scores of conditions, from neurodegenerative diseases to cancer.



Rhesus macaques perch in an outdoor enclosure at the Tulane National Primate Research Center in Covington, Louisiana. Credit: Kathleen Flynn/Reuters/Alamy

Transporting animals has [arisen as a particular challenge](#). In the past decade, after pressure from animal-rights groups, many airlines have stopped carrying primates for research. The NABR filed a complaint with the US Department of Transportation in 2018, asking it to order airlines to carry the animals. Last month, 90 universities, science societies and companies [petitioned the department to take up the issue](#).

The pandemic brought the need for research monkeys into sharp relief. “As expected, non-human primates, largely rhesus, were absolutely critical in the early testing of vaccines and therapeutics,” says Anderson.



### [Animal-research data show effects of EU's tough regulations](#)

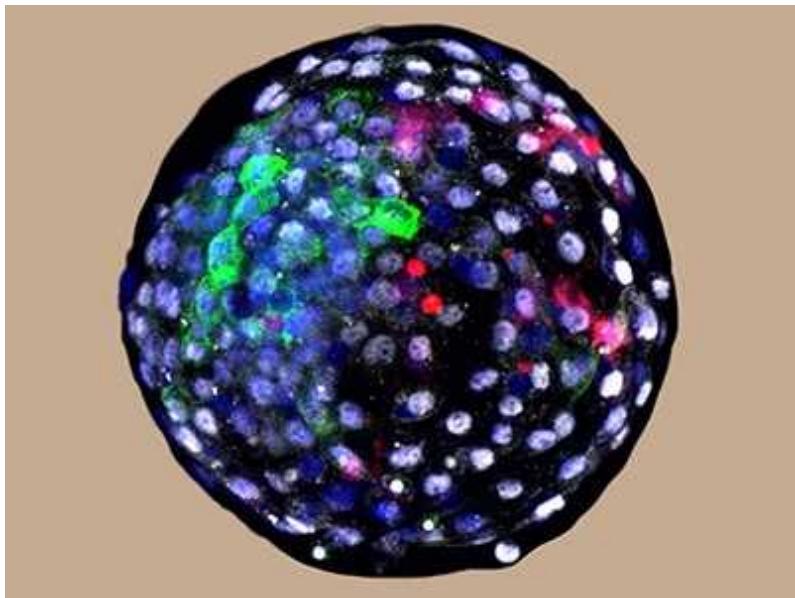
Facing a rush of applications for monkey use last year, the NIH convened an internal committee to review and prioritize the projects that required non-human primates most urgently. The goal was to funnel resources to COVID-19 work without cutting off studies of other conditions, says Anderson.

China has become an important supplier of cynomolgus macaques (*Macaca fascicularis*), but stopped shipping the animals once the pandemic began. The change was hardest on pharmaceutical companies, which prefer that species for drug trials. Anderson says the NIH's focus is on rhesus macaques, which are most in demand at academic labs. Rhesus monkeys tend to thrive in captive research environments, and decades of research on the species means their biology and genetics are well understood.

### **Room to grow**

About \$8 million of the recent funding boost for the NRPCs came from last year's CARES (Coronavirus Aid, Relief, and Economic Security) Act, through which the US government authorized an emergency bundle of money for COVID-19-related work. Some of these funds helped increase capacity at biosafety labs that securely house monkeys infected with the coronavirus SARS-CoV-2.

Centres with room to grow received the first grants when the NIH began its spending increase two years ago, says Sheri Hild, health-science administrator in the Division of Comparative Medicine at the NIH Office of Research Infrastructure Programs in Bethesda. In particular those that could build outdoor housing for monkeys were prioritized — for example, the Oregon NRPC has received about \$3.5 million to allow it to house up to 20% more animals in the short term. An outdoor arrangement is cheaper and potentially better for the animals than an indoor one, says Hild.



### [First monkey–human embryos reignite debate over hybrid animals](#)

At the Tulane NRPC in Covington, Louisiana, associate director Skip Bohm is aiming to add 1,000 monkeys to the 4,500 in the breeding colony that the centre currently houses. The centre has received \$5 million from the NIH, says Bohm. But he cautions that the impact of the funds will be some time coming. “The investment they’re making now is for the future — [the monkeys] are not going to be here next year.”

The Southwest NRPC at Texas Biomedical Research Institute in San Antonio has received about \$3 million from the NIH to expand capacity by 10–20%. The centre currently holds 2,500 animals, mostly baboons, rhesus macaques and marmosets. Separately, Texas Biomedical is planning a new building to house 1,000 more monkeys, and funding most of the \$13.5-million cost.

The centre felt the effects of the COVID-19 crunch first-hand. During the early months of the pandemic, the pharmaceutical firm Pfizer, based in New York City, collaborated with the Southwest NRPC to test the company's COVID-19 vaccine, based on messenger RNA. The centre asked Pfizer to supply its own non-human primates for the work, says its director, Deepak Kaushal.

The expansion funding from the NIH is "both unusual and novel", he says, but still limited. "Talking about the broader context, it's like a drop in the ocean."

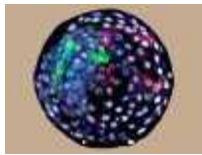
To fully reset and revamp the current set-up at the NRPCs, Kaushal estimates the NIH would need to invest a one-time sum of \$50 million — far beyond current levels, and even more than the ambitious request in Biden's 2022 budget proposal. "I could be wrong, but I think that would be a good start."

*Nature* **595**, 633-634 (2021)

doi: <https://doi.org/10.1038/d41586-021-01894-z>

## Related Articles

-  [US lawmakers propose plan to reduce primate research at National Institutes of Health](#)
-  [Animal-research data show effects of EU's tough regulations](#)



- [First monkey–human embryos reignite debate over hybrid animals](#)

## Subjects

- [SARS-CoV-2](#)
- [Policy](#)
- [Medical research](#)
- [Funding](#)

**nature** careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- [US lawmakers propose plan to reduce primate research at National Institutes of Health](#)



- [Animal-research data show effects of EU's tough regulations](#)



- [First monkey-human embryos reignite debate over hybrid animals](#)

## Subjects

- [SARS-CoV-2](#)
- [Policy](#)
- [Medical research](#)
- [Funding](#)

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

- NEWS
- 22 July 2021

# DeepMind's AI predicts structures for a vast trove of proteins

AlphaFold neural network produced a ‘totally transformative’ database of more than 350,000 structures from *Homo sapiens* and 20 model organisms.

- [Ewen Callaway](#)
  1. Ewen Callaway  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



The human mediator complex has long been one of the most challenging multi-protein systems for structural biologists to understand. Credit: Yuan He

The human genome holds the instructions for more than 20,000 proteins. But only about one-third of those have had their 3D structures determined experimentally. And in many cases, those structures are only partially known.

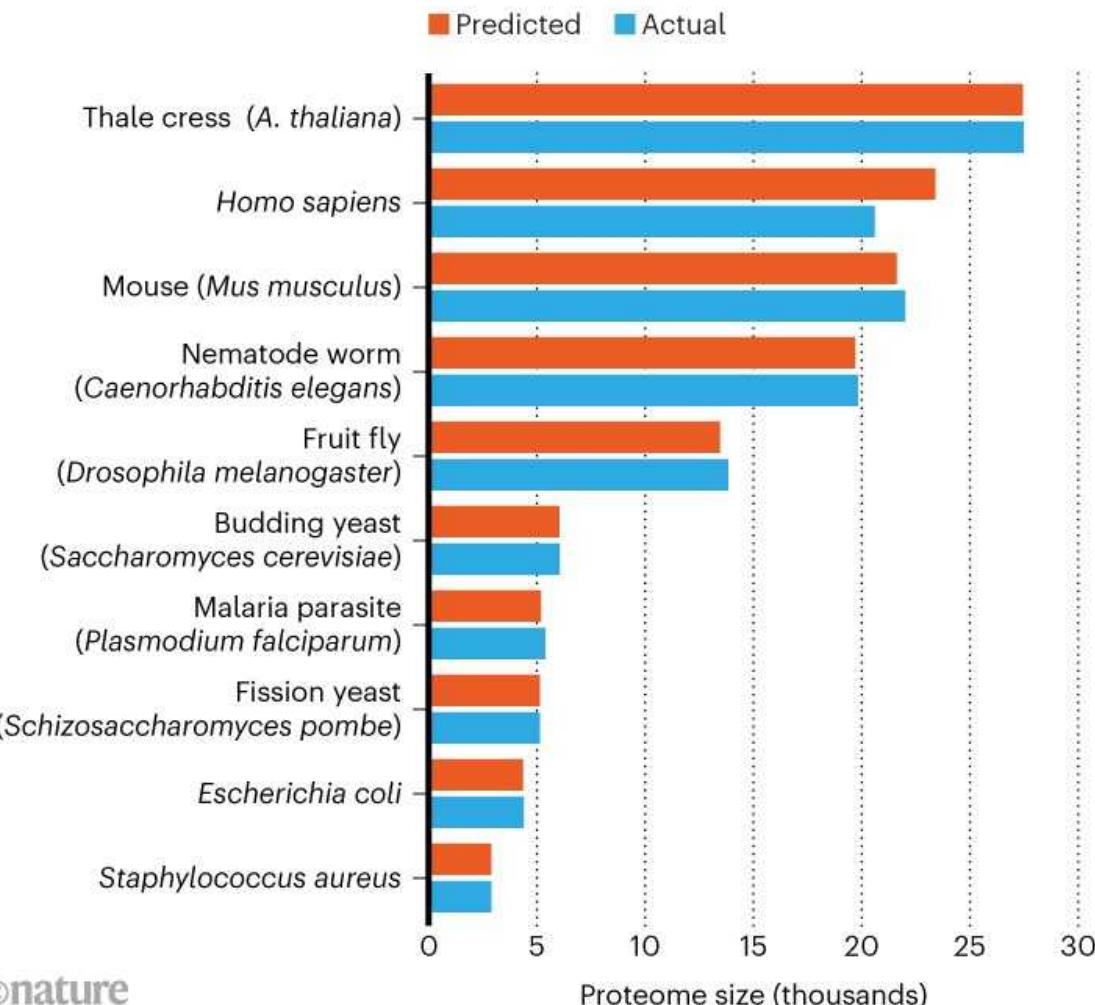
Now, a transformative artificial intelligence (AI) tool called AlphaFold, which has been developed by Google's sister company DeepMind in London, has predicted the structure of nearly the entire human proteome (the full complement of proteins expressed by an organism). In addition, the tool has predicted almost complete proteomes for various other organisms, ranging from mice and maize (corn) to the malaria parasite (see 'Folding options').

The more than 350,000 protein structures, which are available through a public database, vary in their accuracy. But researchers say the resource —

which is set to grow to 130 million structures by the end of the year — has the potential to revolutionize the life sciences.

## FOLDING OPTIONS

AlphaFold has aimed to predict the structure of every protein in humans as well as in 20 model organisms, including those listed here. For some of the proteins, it has provided multiple predictions, which explains why the numbers can be higher than the size of the proteome. In the case of *Homo sapiens*, the predictions include 98.5% of known proteins.



Source: EMBL–EBI and <https://swissmodel.expasy.org/repository>

“It’s totally transformative from my perspective. Having the shapes of all these proteins really gives you insight into their mechanisms,” says Christine Orengo, a computational biologist at University College London (UCL).

“This is the biggest contribution an AI system has made so far to advancing scientific knowledge. I don’t think it’s a stretch to say that,” says Demis Hassabis, co-founder and chief executive of DeepMind.

But researchers emphasize that the data dump is a beginning, not an end. They will want to validate the predictions and, more importantly, apply them to experiments that were hitherto impossible. “It’s an amazing first step, that we have all this data on that scale,” says David Jones, a UCL computational biologist who advised DeepMind on an earlier iteration of AlphaFold.

## Prizewinning predictions

DeepMind stunned the life-sciences community last year, when an updated version of AlphaFold swept a biennial protein-prediction exercise called CASP (Critical Assessment of Protein Structure Prediction). In this long-running competition, which has traditionally been the domain of academics, researchers predict the structures of proteins whose structures have been experimentally solved, but not yet made public.

Some of AlphaFold’s predictions were on par with very good experimental models, and some scientists said the network’s influence would be epochal. Last week, DeepMind released the source code behind the latest version of AlphaFold, and a detailed description of how it was developed<sup>1</sup> (academic teams have already begun using these resources to make useful predictions). In the process of preparing AlphaFold’s code for public release, DeepMind refined it to make the code run more efficiently. Some of the CASP predictions took days, but the updated version of AlphaFold could now compute them in minutes to hours.

With this added efficiency, the DeepMind team set out to predict the structures of nearly every known protein encoded by the human genome, as well as those of 20 model organisms. The structures are [available in a database](#) maintained by EMBL-EBI (the European Molecular Biology Laboratory European Bioinformatics Institute) in Hinxton, UK.

In addition to the predicted structures, which cover 98.5% of known human proteins and a similar percentage for other organisms, AlphaFold generated

a measurement of the confidence of its predictions. “We want to give experimentalists and biologists a really clear signal of which parts of the predictions they should rely on,” says Kathryn Tunyasuvunakool, a science engineer at DeepMind and first author of a *Nature* paper describing the human proteome predictions<sup>2</sup>. For the human proteome, 58% of its predictions for the locations of individual amino acids were good enough to be confident in the shape of the protein’s folds, Tunyasuvunakool says. A subset of those predictions — 36% of the total — are potentially precise enough to detail atomic features useful for drug design, such as the active site of an enzyme.

Even the less-accurate predictions might offer insights. Biologists think that a large proportion of human proteins and those of other eukaryotes — organisms with cells that have nuclei — contain regions that are inherently disordered and take on a defined structure only in concert with other molecules. “Many proteins are just wiggly in solution, they don’t have a fixed structure,” says AlphaFold lead researcher John Jumper. Some of the regions that AlphaFold predicted with low confidence match up with those that biologists suspect are disordered, says Pushmeet Kohli, head of AI for science at DeepMind.

Determining how individual proteins interact with other cellular players is one of the greatest challenges to the AlphaFold predictions, say researchers. For the CASP competition, most of its predictions were of independently folding units of a protein, called domains. But the human proteome, and those of other organisms, contains proteins with multiple domains that fold semi-independently. Human cells also contain molecules made of multiple chains of interacting proteins, such as receptors on cell membranes.

## Deluge of data

The approximately 365,000 structure predictions deposited this week should swell to 130 million — nearly half of all known proteins — by the year’s end, says Sameer Velankar, a structural bioinformatician at EMBL-EBI. The database will be updated as new proteins are identified and predictions improved. “This is not a resource you expect to have access to,” says Tunyasuvunakool, and she is eager to see what scientists come up with.

Researchers are already using AlphaFold and related tools to help make sense out of experimental data generated using X-ray crystallography and cryo-electron microscopy. Marcelo Sousa, a biochemist at the University of Colorado Boulder, used AlphaFold to make models from X-ray data of proteins that bacteria use to evade an antibiotic called colistin. The parts of the experimental model that differed from the AlphaFold prediction were typically regions that the software had assigned with low confidence, Sousa notes, a sign that AlphaFold is accurately predicting its limits.

Still, biologists will want to continue benchmarking these predictions to experimental data to get a better handle on their reliability, says Venki Ramakrishnan, a structural biologist at the MRC Laboratory of Molecular Biology in Cambridge, UK. “We need to be able to trust these data,” adds Orengo.

Jones is impressed with what the network has achieved. But he says that many of the models predicted by AlphaFold could have been generated with earlier software developed by academics. “For most proteins, those results are probably good enough for quite a lot of the things you want to do.” Scientists dead-set on obtaining the structure of any particular protein could probably succeed using experimental approaches.

But the availability of so many protein structures is likely to mark a “paradigm shift” in biology, says Mohammed AlQuraishi, a computational biologist at Columbia University in New York City who works on protein-structure prediction. His field has spent so much time and energy on predicting accurate protein structures on this scale that it hasn’t yet worked out what to do with such resources. “Everything we do today that relies on a protein sequence, we can now do with protein structure.”

Orengo hopes that the database will help her to better understand the structural constraints of proteins. She has mapped a database of known proteins into about 5,000 ‘structural families’, but about half of the proteins in the database are excluded because there is nothing else like them for which a structure has been determined. AlphaFold’s predictions could help uncover new shapes, she says. “We’ll really see what folding space looks like.”

Jones expects AlphaFold will lead to a lot of soul-searching among biologists about what to do with so many structures — and the ease of creating many more. “There will be conferences. Now we’ve got 130 million models, how does this change our view of biology? It may be it doesn’t change it,” he says. “I suspect it will.”

*Nature* **595**, 635 (2021)

doi: <https://doi.org/10.1038/d41586-021-02025-4>

## References

1. 1.

Jumper, J. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-03819-2> (2021).

[Article](#) [Google Scholar](#)

2. 2.

Tunyasuvunakool, K. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-03828-1> (2021).

[Article](#) [Google Scholar](#)

[Download references](#)

## Related Articles

- [DeepMind’s AI for protein structure is coming to the masses](#)

-  [‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures](#)
-  [AI protein-folding algorithms solve structures faster than ever](#)

## Subjects

- [Proteomics](#)
- [Structural biology](#)
- [Machine learning](#)

**nature** careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

- [DeepMind's AI for protein structure is coming to the masses](#)



- ['It will change everything': DeepMind's AI makes gigantic leap in solving protein structures](#)



- [AI protein-folding algorithms solve structures faster than ever](#)

## Subjects

- [Proteomics](#)
- [Structural biology](#)
- [Machine learning](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02025-4>

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

- NEWS
- 16 July 2021

# Massive DNA ‘Borg’ structures perplex scientists

Researchers say they have discovered unique and exciting DNA strands in the mud — others aren’t sure of their novelty.

- [Amber Dance](#) <sup>0</sup>
  1. Amber Dance is a science journalist in Los Angeles, California.  
[View author publications](#)  
You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



Borgs seem to be associated with single-celled microorganisms known as archaea, shown in this scanning-electron microscopy image. Credit: Eye of Science/SPL

The Borg have landed — or, at least, researchers have discovered their counterparts here on Earth. Scientists analysing samples from muddy sites in the western United States have found novel DNA structures that seem to scavenge and ‘assimilate’ genes from microorganisms in their environment, much like the fictional *Star Trek* ‘Borg’ aliens who assimilate the knowledge and technology of other species.

These extra-long DNA strands, which the scientists named in honour of the aliens, join a diverse collection of genetic structures — circular plasmids, for example — known as extrachromosomal elements (ECEs). Most microbes have one or two chromosomes that encode their primary genetic blueprint. But they can host, and often share between them, many distinct ECEs. These carry non-essential but useful genes, such as those for antibiotic resistance.

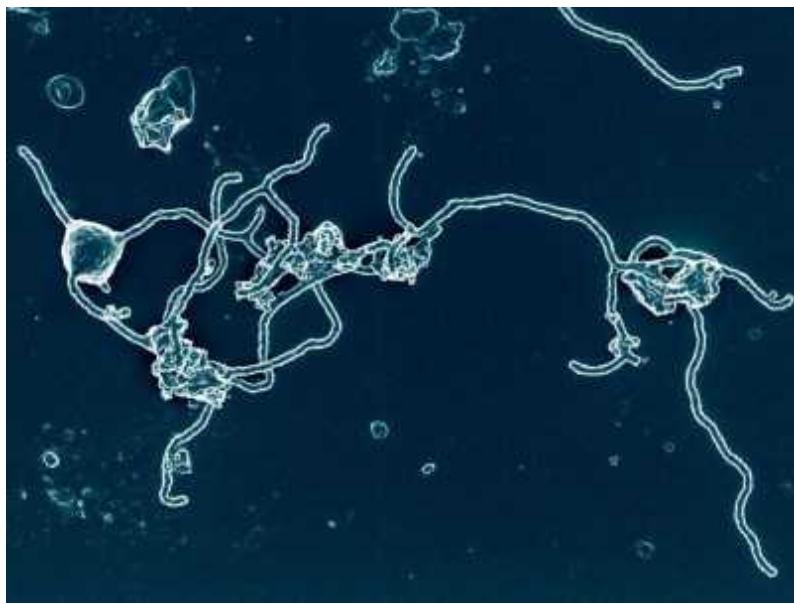
Borgs are a previously unknown, unique and “absolutely fascinating” type of ECE, says Jill Banfield, a geomicrobiologist at the University of California,

Berkeley. She and her colleagues describe their discovery of the structures in a preprint posted to the server bioRxiv<sup>1</sup>. The work is yet to be peer-reviewed.

## Unlike anything seen before

Borgs are DNA structures “not like any that’s been seen before”, says Brett Baker, a microbiologist at the University of Texas at Austin. Other scientists agree that the find is exciting, but have questioned whether Borgs really are unique, noting similarities between them and other large ECEs.

In recent years “people have become used to surprises in the field of ECEs”, says Huang Li, a microbiologist at the Chinese Academy of Sciences in Beijing. “However, the discovery of Borgs, which undoubtedly enriches the concept of ECEs, has fascinated many in the field.”



### The life of archaea

Their vast size, ranging between more than 600,000 and about 1 million DNA base pairs in length, is one feature that distinguishes Borgs from many other ECEs. In fact, Borgs are so huge that they are up to one-third of the length of the main chromosome in their host microbes, Banfield says.

Banfield studies how microbes influence the carbon cycle — including the production and degradation of methane, a potent greenhouse gas — and, in October 2019, she and her colleagues went hunting for ECEs containing genes involved in the carbon cycle in Californian wetlands. There, they found the first Borgs and later identified 19 different types from this and similar sites in Colorado and California.

Borgs seem to be associated with archaea, which are single-celled microorganisms distinct from bacteria. Specifically, those Banfield and her team have discovered are linked to the *Methanoperedens* variety, which digest and destroy methane. And Borg genes seem to be involved in this process, says Banfield.



Researchers found Borgs at this site on the East River, Colorado.Credit: Jordan Hoff

Scientists can't yet culture *Methanoperedens* in the laboratory — an [ongoing challenge](#) for many microbes — so the team's conclusions that Borgs might be used by the archaea for methane processing are based on sequence data alone.

“They’ve made an interesting observation,” says systems biologist Nitin Baliga, at the Institute for Systems Biology in Seattle, Washington. But he cautions that when researchers sift through fragments of many genomes and piece them together, as Banfield’s team has done, it’s possible to make errors. Finding Borgs in cultured *Methanoperedens* will be necessary for the finding to be considered definitive, he adds.

## Costs and benefits

Assuming Borgs are real, maintaining such a massive ECE would be costly for *Methanoperedens*, Banfield and colleagues say, so the DNA structures must provide some benefit. To learn what that might be, the researchers analysed the sequences of hundreds of Borg genes and compared them with known genes.

Borgs seem to house many genes needed for entire metabolic processes, including digesting methane, says Banfield. She describes these collections as “a toolbox” that might super-charge the abilities of *Methanoperedens*.

So what makes a Borg a Borg? In addition to their remarkable size, Borgs share several structural features: they’re linear, not circular as many ECEs are; they have mirrored repetitive sequences at each end of the strand; and they have many other repetitive sequences both within and between the presumptive genes.



## [Scientists glimpse oddball microbe that could help explain rise of complex life](#)

Individually, these features of Borgs can overlap with those seen in other large ECEs, such as elements in certain salt-loving archaea, so Baliga says the novelty of Borgs is still debatable at this stage. Borgs also resemble [giant linear plasmids](#) found in soil-dwelling [Actinobacteria](#), says Julián Rafael Dib, a microbiologist at the Pilot Plant for Microbiological Industrial Processes in Tucumán, Argentina.

Banfield counters that although the individual features of Borgs have been seen before, “the size, combination and metabolic gene load” is what makes them different. She speculates that they were once entire microbes, and were assimilated by *Methanoperedens* in much the same way that eukaryotic cells gained energy-generating mitochondria by assimilating free-living bacteria.

Now that scientists know what to look for, they might find more Borgs by sifting through old data, says Baker, who used to work in Banfield's lab. He thinks he might already have discovered some candidates in his own genetic database since the preprint was posted.

## **Resistance is futile**

When analysing the Borg genome, Banfield and colleagues also saw features suggesting that Borgs have assimilated genes from diverse sources, including the main *Methanoperedens* chromosome, Banfield says. This potential to ‘assimilate’ genes led her son to propose the name ‘Borg’ over Thanksgiving dinner in 2020.

Banfield’s team is now investigating the function of Borgs and the role of their DNA repeats. Repeats are important to microbes: differently-structured repeats called CRISPR are snippets of genetic code from viruses that microbes incorporate into their own DNA to ‘remember’ the pathogens so they can defend against them in the future.



### Pioneers of revolutionary CRISPR gene editing win chemistry Nobel

CRISPR and its associated proteins have been a boon for biotechnology because they have been adapted into a powerful gene-editing technique — hinting that Borg genomes might also yield useful tools. “It could be as important and interesting as CRISPR, but I think it’s going to be a new thing,” says Banfield, who is collaborating on future investigations with her preprint co-author, Jennifer Doudna, a pioneer of CRISPR-based gene editing at the University of California.

One potential application that the researchers see for Borgs could be as an aid in the fight against climate change. Fostering the growth of microbes

containing them could, perhaps, cut down the methane emissions generated by soil-dwelling archaea, which add up to about 1 gigatonne globally each year. It would be risky to do this in natural wetlands, Banfield says, but it might be appropriate at agricultural sites. So, as a first step, her group is now hunting Borgs in Californian rice paddies.

*Nature* **595**, 636 (2021)

doi: <https://doi.org/10.1038/d41586-021-01947-3>

## References

1. 1.

Al-Shayeb, B. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.07.10.451761> (2021).

[Download references](#)

## Related Articles

-  [The life of archaea](#)
-  [Scientists glimpse oddball microbe that could help explain rise of complex life](#)
-  [Pioneers of revolutionary CRISPR gene editing win chemistry Nobel](#)

# Subjects

- [Climate change](#)
- [Biotechnology](#)
- [Microbiology](#)

**nature** careers

## [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- [The life of archaea](#)



- [Scientists glimpse oddball microbe that could help explain rise of complex life](#)



- [Pioneers of revolutionary CRISPR gene editing win chemistry Nobel](#)

## Subjects

- [Climate change](#)
- [Biotechnology](#)
- [Microbiology](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01947-3>

- NEWS
- 20 July 2021

# China launches world's largest carbon market: but is it ambitious enough?

Experts welcome the trading scheme, but question whether it is up to the task of helping China achieve its climate goals.

- [Bianca Nogrady](#)
  1. Bianca Nogrady  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



A coal-fired power plant in Hangzhou in China's Zhejiang province. Credit: Feature China/Barcroft Media/Getty

China, the world's largest emitter of greenhouse gases, has launched its first national emissions-trading scheme. Such carbon-pricing mechanisms exist in around 45 countries already, but China's scheme, which began trading last week, is the world's biggest.

It has been plagued by delays, and researchers argue it might not be ambitious enough to enable China to meet its emissions-reduction goals, including a 2030 deadline for peak emissions and a 2060 goal of net-zero emissions.



## How China could be carbon neutral by mid-century

“We can’t put all the eggs in one basket,” says Hongbo Duan, an economist at the School of Economics and Management at the Chinese Academy of Sciences in Beijing. “We need to do more, like develop renewables and also CCS — carbon capture and storage.”

But Duan is hopeful that the scheme will have a far-reaching impact over time. Unlike other national mechanisms, China is using intensity of emissions (the amount of emissions per unit of energy generated) rather than absolute emissions to help reduce its impact on climate. “In the future, I think it may play a formidable role in curbing carbon emissions,” he says.

## **Trading has begun**

China began testing the waters in 2013, when it launched seven pilot schemes in cities including Beijing, Shanghai and Shenzhen. Reports that some companies were falsifying emissions data have driven a stronger focus in the national scheme on robust monitoring and reporting, says Yan Qin, an economist and lead carbon analyst in Oslo at Refinitiv, a global company that provides data on financial markets.

China's rules and regulations around the scheme came into force in February but online trading did not begin until 16 July.

China's scheme is based on a cap-and-trade model, in which emitters — initially just coal- and gas-fired energy plants — are allocated a certain number of emissions allowances up to a set limit, or cap, and then either trade or buy allowances if they remain below or exceed this. The aim is to expand the plan to industries including construction, oil and chemicals in coming years.

What makes China's scheme different from those operating in other countries and regions, such as the European Union, Canada and Argentina, is that China has chosen to focus on reducing the intensity of emissions generation, rather than absolute emissions.

## **Incentivized efficiency**

Power companies are incentivized to reduce the intensity of emissions, which means producing the same or greater amount of energy while reducing their emissions or keeping them at the same level. That means the absolute emissions can still increase as energy output increases, as long as the companies are reducing the volume of emissions per unit of energy output.

A company's initial emissions cap is a function of both its current energy output and the emissions intensity of its current operations, which is based on factors such as the type of coal and equipment it uses, says Brad Kerin, general manager at the Carbon Market Institute in Melbourne, Australia.

This then helps the authorities "look at how many allowances they offer initially, and then restrict that over time", Kerin says. Each year, the cap is recalculated and reduced, which drives greater efficiency by requiring companies to reduce the amount of emissions they generate for the energy they produce.



### Net-zero emissions targets are vague: three ways to fix

“That’s the core of the emission-trading scheme: it provides incentive to more efficient generation or less carbon-intensive generation” of energy, says Qin, who explains that power companies can upgrade equipment and facilities to be more efficient and trade the emissions allowances they save, or otherwise buy allowances to cover excess emissions.

The issue for China is that its economy is expected to grow by 4–5% per year, which means a significant increase in power consumption, and therefore emissions, Qin says.

The intensity target allows for “future economic development” while also meaning China can still “reduce the carbon emission per unit of economic output”, says Jianlei Mo, an economist at the Center for Energy and Environmental Policy Research at the Chinese Academy of Sciences in Beijing.

## **Doesn’t go far enough**

However, researchers are concerned that the initial allowances are too generous, the prices for these allowances too low, and the penalties for failing to comply are not severe enough to be a deterrent.

“The current design, this intensity-based target that you allow emissions to increase, that is not very helpful,” Qin adds. She suggests that the relatively soft opening is likely to be a concession to power producers and the fossil-fuel industry. “You need to have the thermal producers on board, but after a few years that scheme has to be tightened.”

Frank Jotzo, an environmental economist and director of the Centre for Climate and Energy Policy at the Australian National University in Canberra, says it is a positive sign that China’s emissions-trading scheme has started. It’s unlikely to have much effect on emissions in its current form, he says, “however, it establishes the infrastructure that could be used in future to effectively and quite efficiently reduce emissions in China’s power sector”.

Another challenge for China will be to ensure the integrity of reporting and monitoring of emissions, says Qin. To address issues seen under the pilot schemes, the national mechanism has a tighter standard for companies reporting their emissions, which requires them to provide detailed technical information, such as data on coal type and consumption.

China is also relying on independent agencies to verify the data — similarly to the EU’s scheme, where operators submit annual emissions reports that are then independently confirmed by accredited verifiers. China is also encouraging its citizens and media to police this and report violations.

In 2019, China was responsible for 27% of global carbon emissions — more than 10 billion tonnes of carbon dioxide. However, its per capita emissions — at about 6.8 tonnes of CO<sub>2</sub> per person — are less than half those of nations including the United States, Australia and Canada.

*Nature* **595**, 637 (2021)

doi: <https://doi.org/10.1038/d41586-021-01989-7>

## Related Articles

-  [Net-zero emissions targets are vague: three ways to fix](#)
-  [COVID curbed carbon emissions in 2020 — but not by much](#)
-  [How China could be carbon neutral by mid-century](#)

## Subjects

- [Industry](#)
- [Energy](#)
- [Climate change](#)

**nature** careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [\*\*Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology\*\*](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- [\*\*Net-zero emissions targets are vague: three ways to fix\*\*](#)



- [\*\*COVID curbed carbon emissions in 2020 — but not by much\*\*](#)



- [\*\*How China could be carbon neutral by mid-century\*\*](#)

## Subjects

- [Industry](#)
  - [Energy](#)
  - [Climate change](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01989-7>

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

- NEWS EXPLAINER
- 20 July 2021

# Should children get COVID vaccines? What the science says

With vaccination campaigns under way in some countries while others weigh the options, *Nature* looks at the evidence for vaccinating younger people.

- [Heidi Ledford](#)
  1. Heidi Ledford  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



A student in Bogor, Indonesia, receives a Sinovac jab in a COVID-19 vaccination campaign for school kids aged 12–18. Credit: Adriana Adie/NurPhoto/Getty

At a time when much of the world is still [struggling to access COVID-19 vaccines](#), the question of whether to vaccinate children can feel like a privilege. On 19 July, vaccine advisers in the United Kingdom recommended delaying vaccines for most young people under 16, citing the very low rates of serious disease in this age group. But several countries, including the United States and Israel, have forged ahead, and others are hoping to follow suit when supplies allow.

*Nature* looks at where the evidence stands on children and COVID vaccines.

## Is it necessary?

Since the early days of the pandemic, parents have been taking some comfort from the fact that SARS-CoV-2 is [much less likely to cause serious illness in children than it is in adults.](#)

But some children do still become very ill, and the spectre of [long COVID](#) — a constellation of sometimes debilitating symptoms that can linger for months after even a mild bout of COVID-19 — is enough for many paediatricians to urge vaccination as quickly as possible. “I spent the pandemic taking care of kids in a children’s hospital,” says Adam Ratner, a paediatric infectious-disease specialist at New York University. “We saw not as many as in the adult side, but plenty of children who were quite ill.”

Vaccine advisers in the United Kingdom, however, have recommended that only adolescents who are clinically vulnerable, or who live with vulnerable adults, be vaccinated for the time being. Severe illness, deaths and even long COVID are rare among healthy adolescents and children, and soon, nearly all vulnerable adults will have received two vaccine doses, University of Bristol paediatrician Adam Finn told reporters at a media briefing.

But in some countries, little is yet known about how COVID-19 affects children. Some official tallies of hospitalizations and deaths due to COVID-19 in sub-Saharan Africa, for example, do not break down the cases by age. As a result, paediatricians don’t know which deaths were in children and young people, and how outcomes of COVID-19 might be affected by conditions such as malnutrition, or concurrent tuberculosis or HIV infection. “We are feeling in the dark,” says Nadia Sam-Agudu, a paediatrician with the University of Maryland School of Medicine in Baltimore who works in Nigeria.

In addition, some paediatricians are concerned about what will happen to children who are co-infected with SARS-CoV-2 and other common viruses, such as respiratory syncytial virus, which is one of the causes of the common cold but can sometimes cause more severe breathing illness in young children. Strict lockdowns have kept this problem at bay in some regions, but as social-distancing measures are eased, there are already signs that respiratory syncytial virus infections in children are rising, says Danilo Buonsenso, a paediatrician at the Gemelli University Hospital in Rome. “We

don't know yet what will be the burden of co-infections in children when we have a massive circulation of routine viruses and COVID," he says.

## Is vaccinating children safe?

A handful of vaccines have been tested in young people over the age of 12, including mRNA vaccines made by Moderna and Pfizer–BioNTech, and two Chinese vaccines made by Sinovac and Sinopharm. And several countries, including the United States, Israel and China, are now offering vaccines to this age group. Other studies are expected to report results in young people over the age of 12 soon, including studies on the Zydus Cadila vaccine and the Covaxin inactivated coronavirus vaccine, both made in India.

Thus far, the vaccines seem to be safe in adolescents<sup>1</sup>, and some companies have moved on to carrying out clinical trials in children as young as six months old. In the United States, vaccines for those under 12 might be available later this year, says paediatrician Andrea Shane at Emory University in Atlanta, Georgia.

A potential link between the Pfizer vaccine and heart inflammation — conditions called myocarditis and pericarditis — has emerged since Israel and the United States began vaccinating young people. However, researchers have yet to establish that the vaccine caused the inflammation. Most of those affected have recovered, and the data suggest that the risk of these conditions is “extremely low”, says paediatrician David Pace at the University of Malta in Msida — about 67 cases per million second doses in adolescent males aged 12–17, and 9 per million in adolescent females in the same age group.

## How will vaccinating children and young people affect the pandemic?

Malta has fully vaccinated 80% of its population — one of the highest vaccination rates in the world — and is now also vaccinating adolescents over the age of 12. There, the decision to vaccinate young people was shaped, among other factors, by the close-knit family structures in a country

where adolescents often have frequent contact with their grandparents, says Pace. “On a population level, vaccinated adolescents may result in a reduction in transmission to vulnerable older people,” he says. Young people in Malta also often travel abroad for school, potentially importing coronavirus infections and variants from abroad, he adds.

Data show that children and particularly adolescents can play a significant part in coronavirus transmission, says Catherine Bennett, an epidemiologist at Deakin University in Melbourne, Australia. And concerns about transmission by children and adolescents are growing as new coronavirus variants emerge. It’s possible that more-transmissible variants will develop a way to push through whatever it is in a young person’s immune response that makes them more resistant to infection, says Bennett, making it all the more important that they are vaccinated.

Hopes of achieving herd immunity through immunization have waned, so countries need to do the best that they can to keep transmission low, she adds: “You only need one poorly vaccinated population to generate global variants.”

## Is vaccinating children fair?

Chile, another country with one of the highest COVID-19 vaccination rates in the world, is also rolling out vaccines to those aged 12 and older.

But Miguel O’Ryan, a former member of two advisory committees to the government there who has pushed for aggressive vaccination campaigns, now finds himself wondering whether it’s time to slow down. “Probably countries should not move forward with paediatric vaccinations so fast,” says O’Ryan, who is a paediatric infectious-disease specialist at the University of Chile in Santiago. “Other countries, even our neighbours, are struggling very hard to get enough vaccines for their high-risk groups.”

O’Ryan is not the only one concerned about using valuable vaccines to inoculate children, when more vulnerable populations around the world are still struggling to secure supplies. In May, World Health Organization chief Tedros Adhanom Ghebreyesus said that wealthier countries that are

vaccinating children are doing so at the expense of health-care workers and high-risk groups in other countries. But advocates for vaccinating children and young adults argue that it need not be a case of one or the other. “This is sort of a false dichotomy,” says Ratner. Sam-Agudu agrees, pointing out that some wealthy countries bought more than enough doses to fully vaccinate their populations. “The argument for sending vaccines outside the country should not preclude vaccinating children in higher-income countries,” she says.

And there are other steps that could be taken to improve the supply of vaccines to needy countries, says Bennett. More could be done to better target donations, she notes. For example, rather than allocating donated vaccine doses to countries solely on the basis of how many people live there, they could be distributed according to other factors, such as the need to preserve health-care services in the face of an oncoming malaria season, or ongoing measles outbreak. “We probably still haven’t had the deep epidemiological war room that we need to map out the problem and the best way to address it,” she says. “There’s a whole range of ways you could look at this.”

*Nature* **595**, 638-639 (2021)

doi: <https://doi.org/10.1038/d41586-021-01898-9>

## References

1. 1.

Han, B. *et al. Lancet Infect. Dis.* [https://doi.org/10.1016/S1473-3099\(21\)00319-4](https://doi.org/10.1016/S1473-3099(21)00319-4) (2021).

[Article](#) [Google Scholar](#)

[Download references](#)

## Related Articles

-  [Deaths from COVID ‘incredibly rare’ among children](#)
-  [Long COVID and kids: scientists race to find answers](#)
-  [Will COVID become a disease of the young?](#)
-  [COVID and schools: the evidence for reopening safely](#)
-  [Does vaccinating adults stop kids from spreading COVID too?](#)

## Subjects

- [Infection](#)
- [Medical research](#)
- [Immunology](#)
- [Vaccines](#)
- [SARS-CoV-2](#)

**nature** careers

## Jobs >

- **Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics**

City University of Hong Kong (CityU)

Hong Kong, China

- **FACULTY POSITION IN MICROBIOME BIOCHEMISTRY**

Duke University

Durham, NC, United States

- **Bioinformatician (m/f/div)**

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- **Student Research Assistant / HIWI (m/f/x) for Literature Research in the Field of Soil Microbiology**

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## **Related Articles**

-  [Deaths from COVID ‘incredibly rare’ among children](#)
-  [Long COVID and kids: scientists race to find answers](#)
-  [Will COVID become a disease of the young?](#)
-  [COVID and schools: the evidence for reopening safely](#)
-  [Does vaccinating adults stop kids from spreading COVID too?](#)

## Subjects

- [Infection](#)
- [Medical research](#)
- [Immunology](#)
- [Vaccines](#)
- [SARS-CoV-2](#)

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

- NEWS
- 15 July 2021

# Deaths from COVID ‘incredibly rare’ among children

Studies find that overall risk of death or severe disease from COVID-19 is very low in kids.

- [Heidi Ledford](#)
  1. Heidi Ledford  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



A child performing a lateral-flow test for COVID-19 in Chertsey, UK. Credit: Dan Kitwood/Getty

A comprehensive analysis of hospital admissions and reported deaths across England suggests that COVID-19 carries a lower risk of dying or requiring intensive care among children and young people than was previously thought.

In a series of preprints published on medRxiv<sup>1–3</sup>, a team of researchers picked through all hospital admissions and deaths reported for people younger than 18 in England. The studies found that COVID-19 caused 25 deaths in that age group between March 2020 and February 2021.

About half of those deaths were in individuals with an underlying complex disability with high health-care needs, such as tube feeding or assistance with breathing.

The studies did not evaluate rates of less-severe illness or debilitating ['long COVID' symptoms](#) that can linger months after the acute phase of the infection has past. “The low rate of severe acute disease is important news,

but this does not have to mean that COVID does not matter to children,” says paediatrician Danilo Buonsenso at the Gemelli University Hospital in Rome. “Please, let’s keep attention — as much as is feasible — on immunization.”

In one of the preprints, the researchers trawled for published accounts of COVID-19 among children and young people, and ultimately analysed data from 57 studies and 19 countries<sup>3</sup>. They then picked apart risk factors for severe disease and death from the data.

## Study findings

Some conditions — including obesity and cardiac or neurological conditions — were associated with a higher risk of death or intensive-care treatment, the researchers found. But the absolute increase in risk was very small, study author Rachel Harwood, a paediatric surgical registrar at Alder Hey Children’s Hospital in Liverpool, UK, told reporters at a media briefing.

For the other two preprints, the researchers focused on England, drawing on nationwide health-care data on intensive-care admissions and deaths among those under 18 years old. The team found that, of 6,338 hospital admissions for COVID-19, 259 children and young people required treatment in paediatric intensive-care units.

Black children were more likely than their white counterparts to require intensive care, both for COVID-19 and for paediatric multisystem inflammatory syndrome, a rare syndrome associated with coronavirus infection. But overall, the need for intensive care was “incredibly rare” among these patients, says study author Joseph Ward of the University College London Great Ormond Street Institute of Child Health.

Of 3,105 deaths from all causes among the 12 million or so people under 18 in England between March 2020 and February 2021, 25 were attributable to COVID-19 — a rate of about 2 for every million people in this age range. None had asthma or type-1 diabetes, the authors note, and about half had conditions that put them at a higher risk than healthy children of dying from any cause.

Taken together, the unusually comprehensive studies could provide some comfort to parents who have been shielding children who they thought might be vulnerable to severe complications from COVID-19. “There’s a general feeling among paediatricians that probably too many children were shielded during the first wave of the pandemic,” Russell Viner, who studies adolescent health at the University College London, told reporters.

In some cases, those efforts might have done more harm than good, added Elizabeth Whittaker, an infectious-disease specialist at Imperial College London. “Shields are very leaky,” she said. “The shields have not been perfect, and have probably caused more stress and anxiety for families than benefit.”

The work does not tackle the spectre of long COVID, but other studies suggest that [it does occur in children](#) — including in those who had mild initial symptoms or were asymptomatic — but less frequently than in adults.

Despite the very low death and intensive-care rates, Buonsenso hopes that schools will embrace measures such as masks and improved ventilation, and that parents will focus on immunization — for either their children, where possible, or themselves.

“When adults are immunized, fewer children are infected,” he says. “We need to do as much as possible to reduce COVID-19 infection in children.”

*Nature* **595**, 639 (2021)

*doi:* <https://doi.org/10.1038/d41586-021-01897-w>

## References

1. 1.

Ward, J. L. *et al.* Preprint at medRxiv  
<https://doi.org/10.1101/2021.07.01.21259785> (2021).

2. 2.

Smith, C. *et al.* Preprint at medRxiv  
<https://doi.org/10.1101/2021.07.07.21259779> (2021).

3. 3.

Harwood, R. *et al.* Preprint at medRxiv  
<https://doi.org/10.1101/2021.06.30.21259763> (2021).

[Download references](#)

## Related Articles

-  [\*\*Will COVID become a disease of the young?\*\*](#)
-  [\*\*COVID and schools: the evidence for reopening safely\*\*](#)
-  [\*\*Does vaccinating adults stop kids from spreading COVID too?\*\*](#)
-  [\*\*How kids' immune systems can evade COVID\*\*](#)

**nature** careers

[Jobs >](#)

- **Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics**

City University of Hong Kong (CityU)

Hong Kong, China

- **FACULTY POSITION IN MICROBIOME BIOCHEMISTRY**

Duke University

Durham, NC, United States

- **Bioinformatician (m/f/div)**

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- **Student Research Assistant / HIWI (m/f/x) for Literature Research in the Field of Soil Microbiology**

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- **Will COVID become a disease of the young?**



- [COVID and schools: the evidence for reopening safely](#)



- [Does vaccinating adults stop kids from spreading COVID too?](#)



- [How kids' immune systems can evade COVID](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01897-w>

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

- NEWS FEATURE
- 28 July 2021

# How the coronavirus infects cells — and why Delta is so dangerous

Scientists are unpicking the life cycle of SARS-CoV-2 and how the virus uses tricks to evade detection.

- [Megan Scudellari](#) 0

1. Megan Scudellari

1. Megan Scudellari is a science journalist in Boston, Massachusetts.

[View author publications](#)

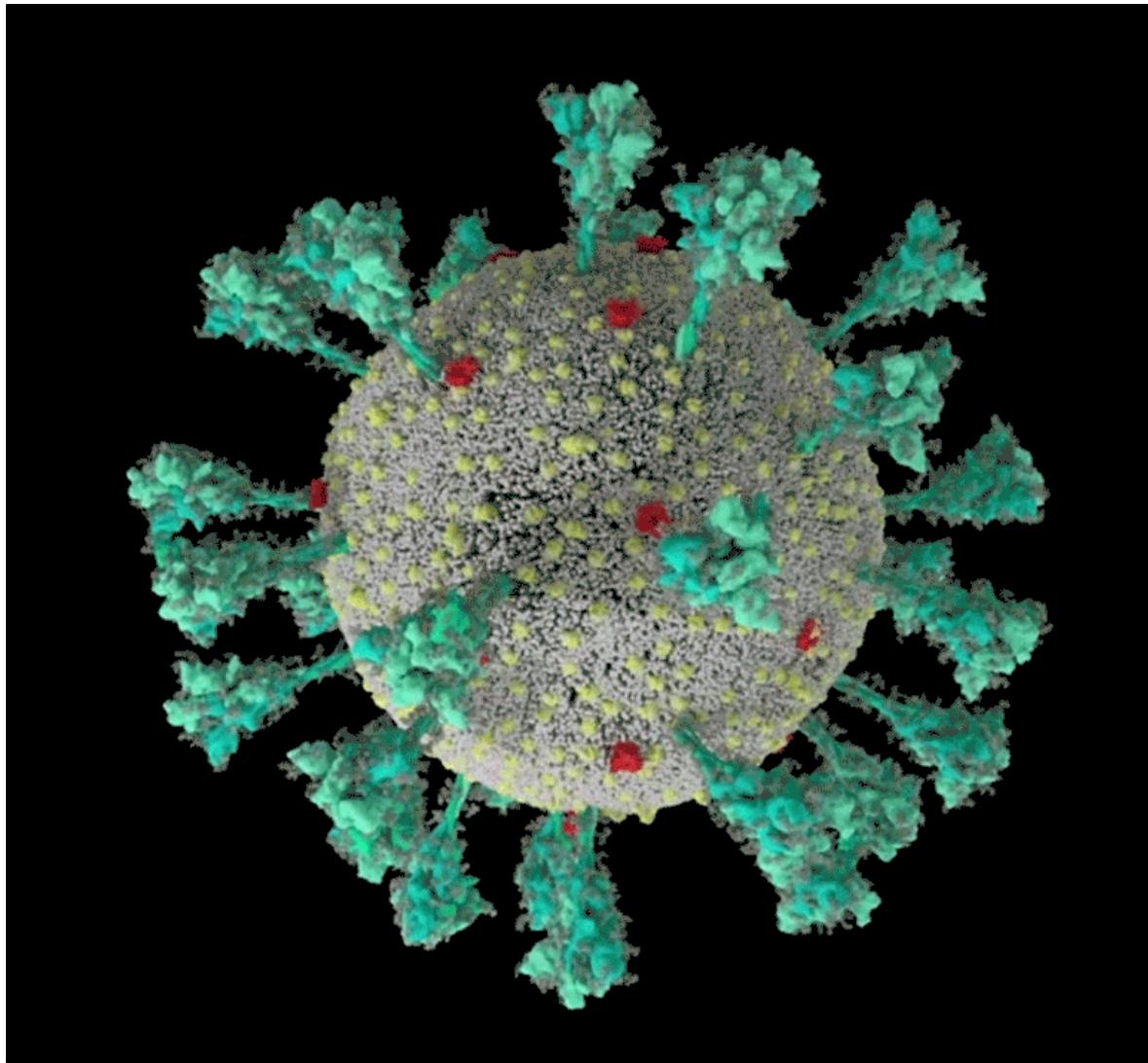
You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



A computer simulation of the structure of the coronavirus SARS-CoV-2. Credit: [Janet Iwasa, University of Utah](#)

The coronavirus sports a luxurious sugar coat. “It’s striking,” thought Rommie Amaro, staring at her computer simulation of one of the trademark spike proteins of SARS-CoV-2, which stick out from the virus’s surface. It was swathed in sugar molecules, known as glycans.

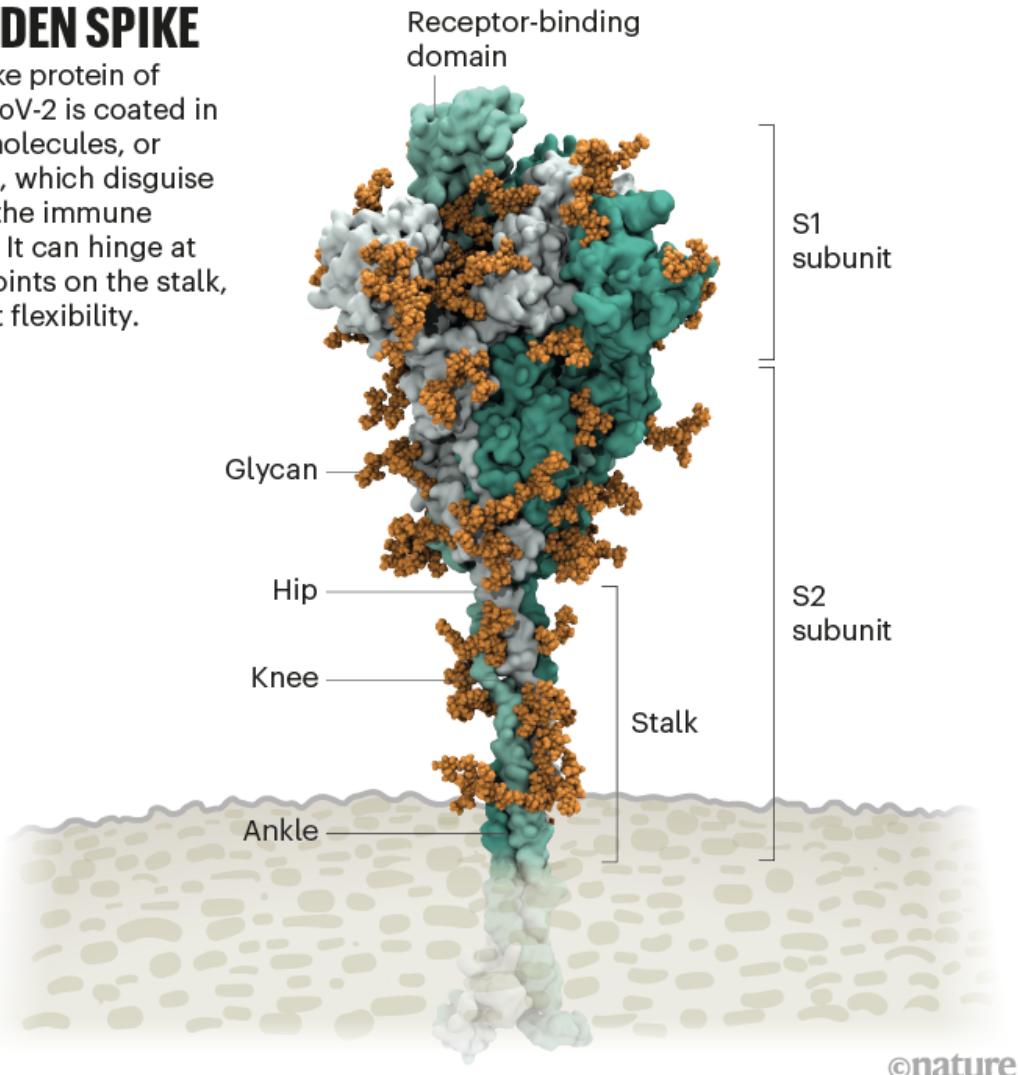
“When you see it with all the glycans, it’s almost unrecognizable,” says Amaro, a computational biophysical chemist at the University of California, San Diego.

Many viruses have glycans covering their outer proteins, camouflaging them from the human immune system like a wolf in sheep's clothing. But last year, Amaro's laboratory group and collaborators created the most detailed visualization yet of this coat, based on structural and genetic data and rendered atom-by-atom by a supercomputer. On 22 March 2020, [she posted the simulation to Twitter](#). Within an hour, one researcher asked in a comment: what was the naked, uncoated loop sticking out of the top of the protein?

Amaro had no idea. But ten minutes later, structural biologist Jason McLellan at the University of Texas at Austin chimed in: the uncoated loop was a receptor binding domain (RBD), one of three sections of the spike that bind to receptors on human cells (see 'A hidden spike').

## A HIDDEN SPIKE

The spike protein of SARS-CoV-2 is coated in sugar molecules, or glycans, which disguise it from the immune system. It can hinge at three points on the stalk, giving it flexibility.



©nature

Source: Structural image from Lorenzo Casalino, Univ. California, San Diego (Ref. 1); Graphic: Nik Spencer/*Nature*

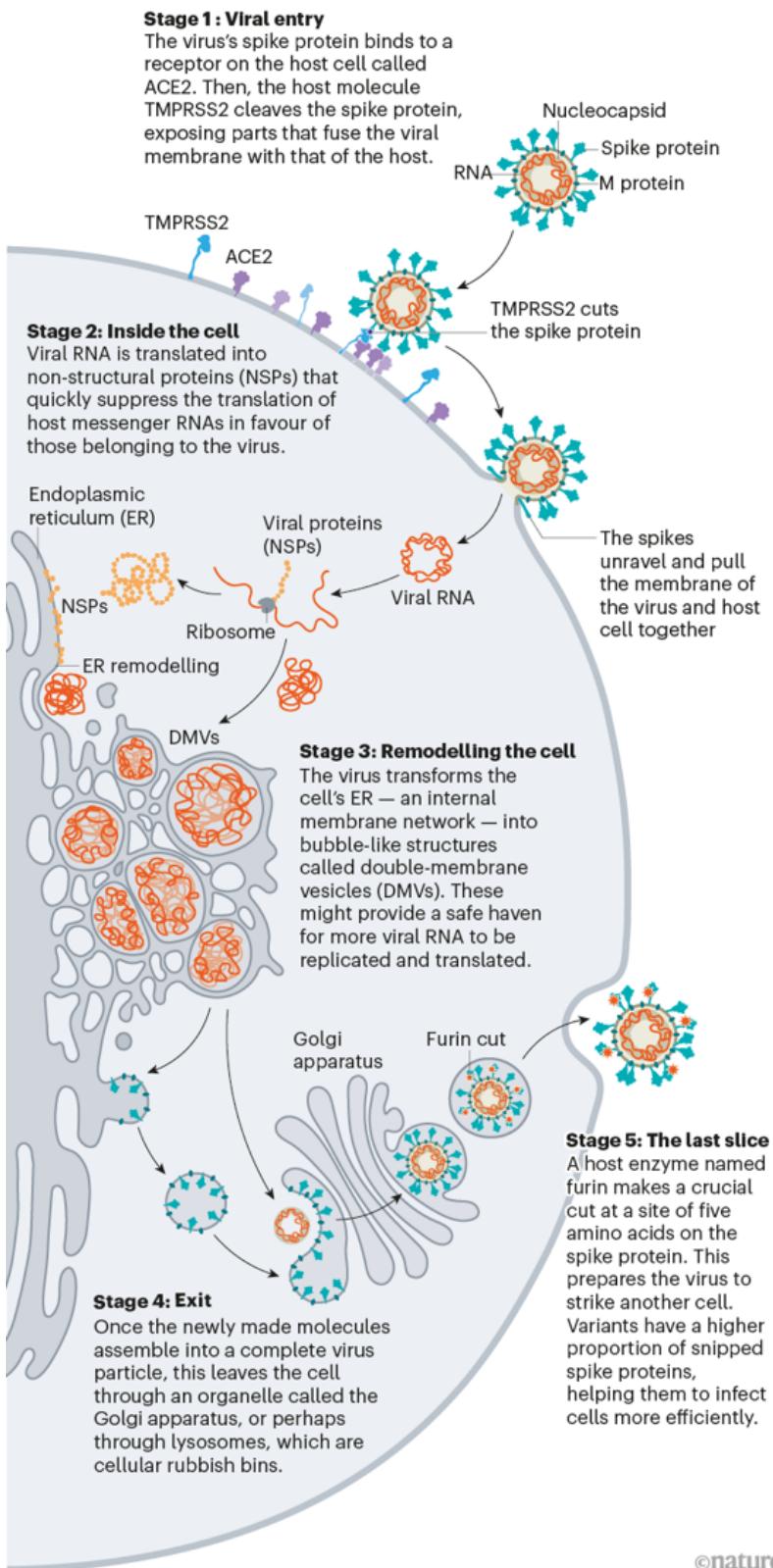
In Amaro's simulation, when the RBD lifted up above the glycan cloud, [two glycans swooped in to lock it into place](#), like a kickstand on a bicycle. When Amaro mutated the glycans in the computer model, the RBD collapsed. McLellan's team built a way to try the same experiment in the lab, and by June 2020, the collaborators had reported that mutating the two glycans reduced the ability of the spike protein to bind to a human cell receptor<sup>1</sup> — a role that no one has previously recognized in coronaviruses, McLellan says. It's possible that snipping out those two sugars could reduce the virus's infectivity, says Amaro, although researchers don't yet have a way to do this.

Since the start of the COVID-19 pandemic, scientists have been developing a detailed understanding of how SARS-CoV-2 infects cells. By picking apart the infection process, they hope to find better ways to interrupt it through improved treatments and vaccines, and learn why the latest strains, such as the Delta variant, [are more transmissible](#).

What has emerged from 19 months of work, backed by decades of coronavirus research, is a blow-by-blow account of how SARS-CoV-2 invades human cells (see ‘Life cycle of the pandemic coronavirus’). Scientists have discovered key adaptations that help the virus to grab on to human cells with surprising strength and then hide itself once inside. Later, as it leaves cells, SARS-CoV-2 executes a crucial processing step to prepare its particles for infecting even more human cells. These are some of the tools that have enabled the virus to spread so quickly and claim millions of lives. “That’s why it’s so difficult to control,” says Wendy Barclay, a virologist at Imperial College London.

# LIFE CYCLE OF THE PANDEMIC CORONAVIRUS

A simplified account of how SARS-CoV-2 enters and exits cells.

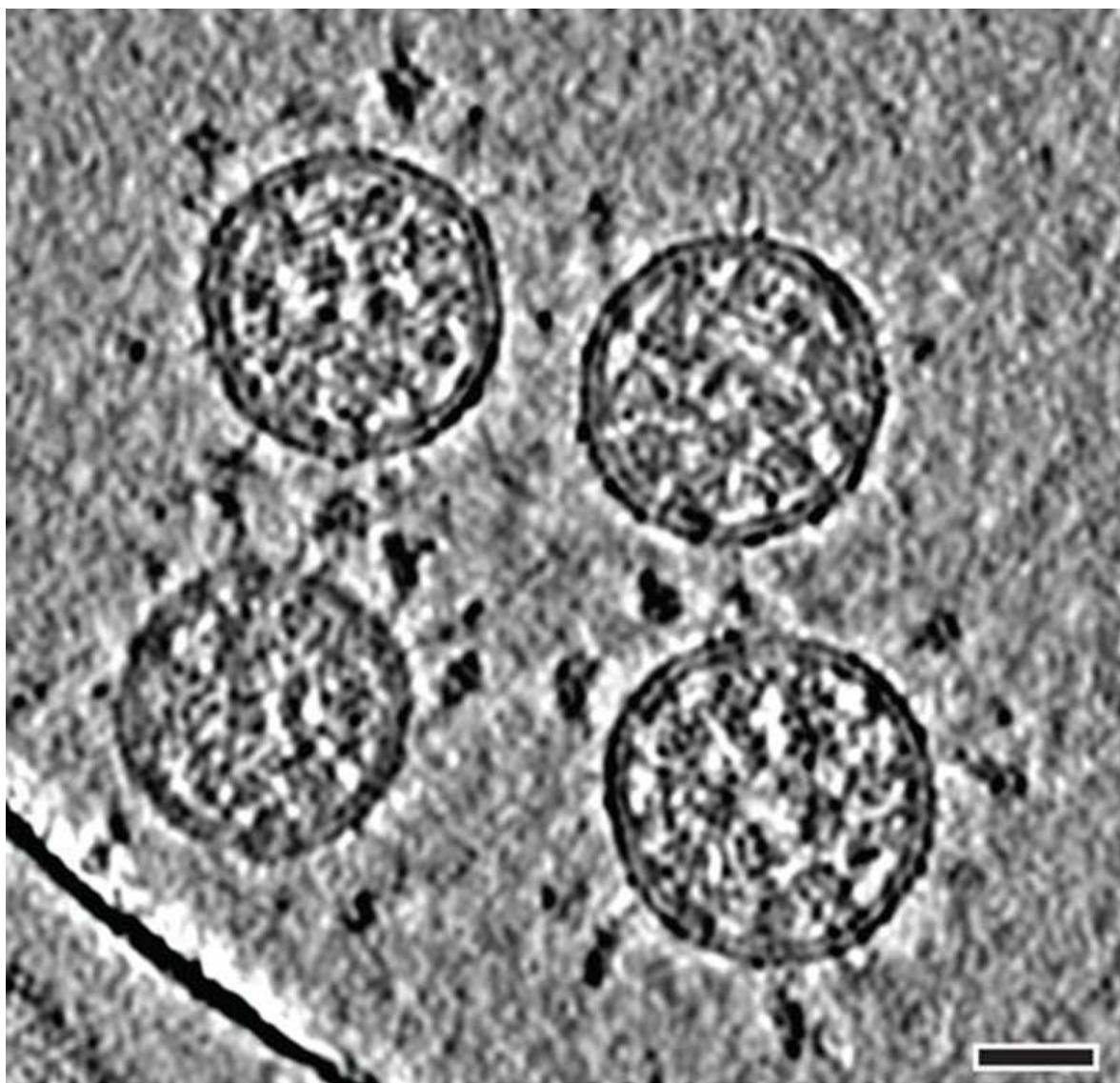


Source: [Hui \(Ann\) Liu, Univ. Utah](#); Graphic: Nik Spencer/*Nature*

## Barbed and ready

It starts with the spikes. Each SARS-CoV-2 virion (virus particle) has an outer surface peppered with 24–40 haphazardly arranged spike proteins that are its key to fusing with human cells<sup>2</sup>. For other types of virus, such as influenza, external fusion proteins are relatively rigid. SARS-CoV-2 spikes, however, are wildly flexible and hinge at three points, according to work published in August 2020 by biochemist Martin Beck at the Max Planck Institute of Biophysics in Frankfurt, Germany, and his colleagues<sup>3</sup>.

That allows the spikes to flop around, sway and rotate, which could make it easier for them to scan the cell surface and for multiple spikes to bind to a human cell. There are no similar experimental data for other coronaviruses, but because spike-protein sequences are highly evolutionarily conserved, it is fair to assume the trait is shared, says Beck.



Cryo-electron tomography images of SARS-CoV-2 virions. (Scale bar: 30 nanometres.) Credit: B. Turoňová *et al./Science*

Early in the pandemic, researchers confirmed that the RBDs of SARS-CoV-2 spike proteins attach to a familiar protein called the ACE2 receptor, which adorns the outside of most human throat and lung cells. This receptor is also the docking point for SARS-CoV, the virus that causes severe acute respiratory syndrome (SARS). But compared with SARS-CoV, SARS-CoV-2 binds to ACE2 an estimated 2–4 times more strongly<sup>4</sup>, because several changes in the RBD stabilize its virus-binding hotspots<sup>5</sup>.

Worrying variants of SARS-CoV-2 tend to have mutations in the S1 subunit of the spike protein, which hosts the RBDs and is responsible for binding to the ACE2 receptor. (A second spike subunit, S2, prompts viral fusion with the host cell's membrane.)

The Alpha variant, for example, includes ten changes in the spike-protein sequence, which result in RBDs being more likely to stay in the ‘up’ position<sup>6</sup>. “It is helping the virus along by making it easier to enter into cells,” says Priyamvada Acharya, a structural biologist at the Duke Human Vaccine Institute in Durham, North Carolina, who is studying the spike mutations.

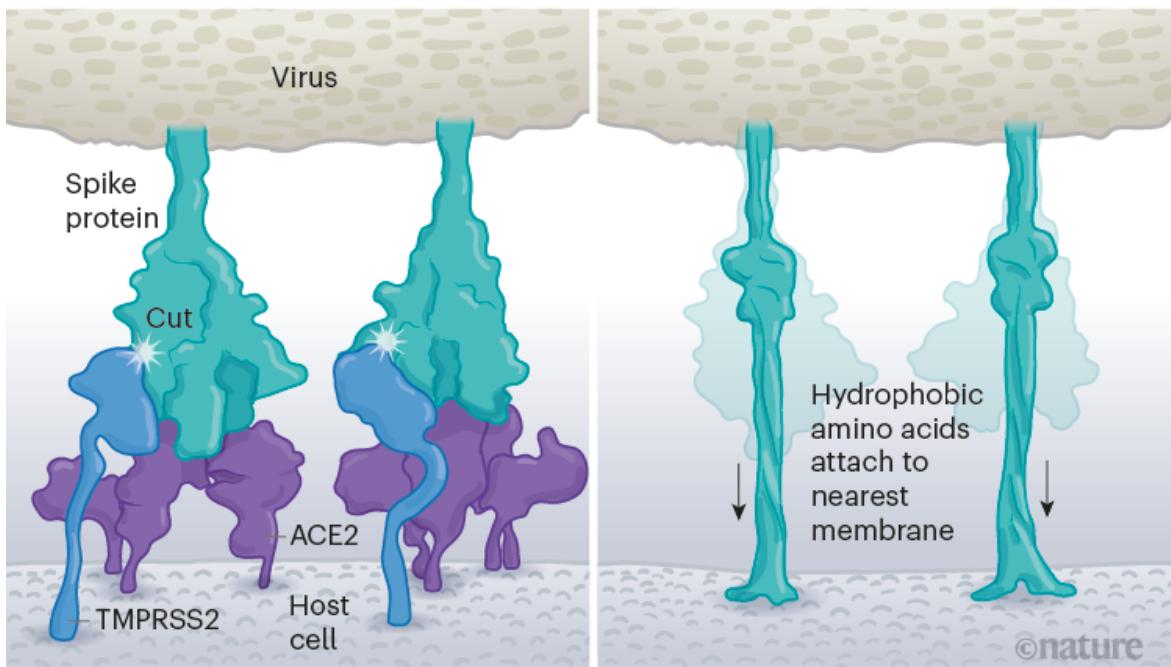
The Delta variant, which is now spreading around the world, hosts multiple mutations in the S1 subunit, including three in the RBD that seem to improve the RBD’s ability to bind to ACE2 and evade the immune system<sup>7</sup>.

## Restricted entry

Once the viral spikes bind to ACE2, other proteins on the host cell’s surface initiate a process that leads to the merging of viral and cell membranes (see ‘Viral entry up close’).

## VIRAL ENTRY UP CLOSE

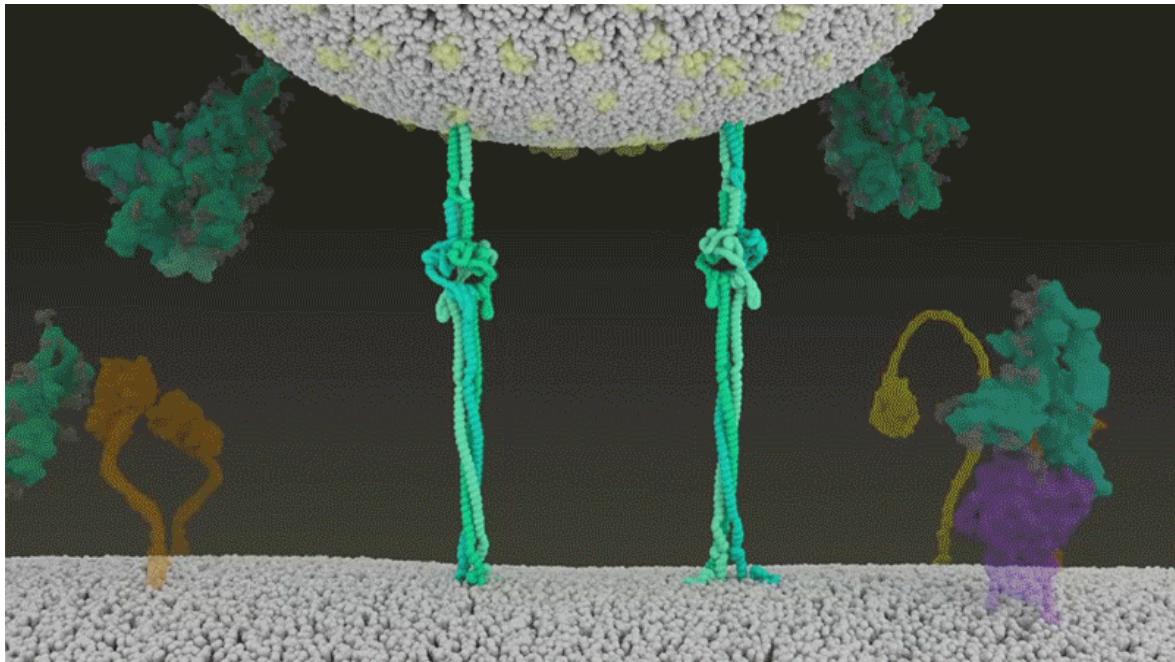
Virus and host-cell membranes fuse after the TMPRSS2 enzyme cuts a SARS-CoV-2 spike protein. This exposes hydrophobic amino acids in the spike that rapidly embed themselves into the nearest membrane — that of the host cell.



Source: [Janet Iwasa, Univ. Utah](#); Graphic: Nik Spencer/*Nature*

The virus that causes SARS, SARS-CoV, uses either of two host protease enzymes to break in: TMPRSS2 (pronounced ‘tempress two’) or cathepsin L. TMPRSS2 is the faster route in, but SARS-CoV often enters instead through an endosome — a lipid-surrounded bubble — which relies on cathepsin L. When virions enter cells by this route, however, antiviral proteins can trap them.

SARS-CoV-2 differs from SARS-CoV because it efficiently uses TMPRSS2, an enzyme found in high amounts on the outside of respiratory cells. First, TMPRSS2 cuts a site on the spike’s S2 subunit<sup>8</sup>. That cut exposes a run of hydrophobic amino acids that rapidly buries itself in the closest membrane — that of the host cell. Next, the extended spike folds back onto itself, like a zipper, forcing the viral and cell membranes to fuse.



An animation of the way SARS-CoV-2 fuses with cells.Credit: [Janet Iwasa, University of Utah](#)

The virus then ejects its genome directly into the cell. By invading in this spring-loaded manner, SARS-CoV-2 infects faster than SARS-CoV and avoids being trapped in endosomes, according to work published in April by Barclay and her colleagues at Imperial College London<sup>9</sup>.

The virus's speedy entry using TMPRSS2 explains why the malaria drug chloroquine didn't work in clinical trials as a COVID-19 treatment, despite early promising studies in the lab<sup>10</sup>. Those turned out to have used cells that rely exclusively on cathepsins for endosomal entry. "When the virus transmits and replicates in the human airway, it doesn't use endosomes, so chloroquine, which is an endosomal disrupting drug, is not effective in real life," says Barclay.

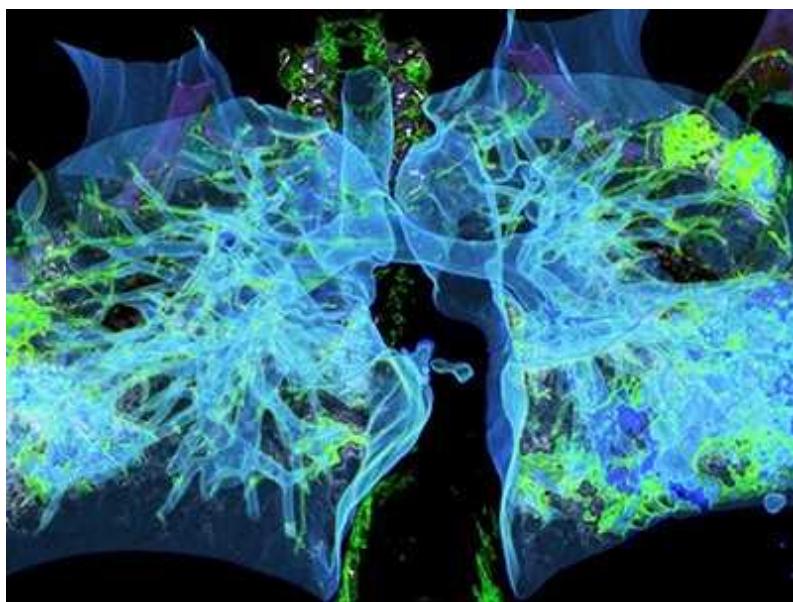
The discovery also points to protease inhibitors as a promising therapeutic option to prevent a virus from using TMPRSS2, cathepsin L or other proteases to enter host cells. One TMPRSS2 inhibitor, camostat mesylate, which is approved in Japan to treat pancreatitis, blocked viral entry into lung cells<sup>8</sup>, but the drug did not improve patients' outcomes in an initial clinical trial<sup>11</sup>.

“From my perspective, we should have such protease inhibitors as broad antivirals available to fight new disease outbreaks and prevent future pandemics at the very beginning,” says Stefan Pöhlmann, director of the Infection Biology Unit at the German Primate Center in Göttingen, who has led research on ACE2 binding and the TMPRSS2 pathway.

## Deadly competition

The next steps of infection are murkier. “There are a lot more black boxes once you are inside the cell,” says chemist Janet Iwasa at the University of Utah in Salt Lake City, who is developing an annotated animation of the viral life cycle. “There’s more uncertainty, and competing hypotheses.”

After the virus shoots its RNA genome into the cell, ribosomes in the cytoplasm translate two sections of viral RNA into long strings of amino acids, which are then snipped into 16 proteins, including many involved in RNA synthesis. Later, more RNAs are generated that code for a total of 26 known viral proteins, including structural ones used to make new virus particles, such as the spike, and other accessory proteins. In this way, the virus begins churning out copies of its own messenger RNA. But it needs the cell’s machinery to translate those mRNAs into proteins.



[How a rampant coronavirus variant blunts our immune defences](#)

Coronaviruses take over that machinery in many ways. Virologist Noam Stern-Ginossar and her team at the Weizmann Institute of Science in Rehovot, Israel, zoomed in on three mechanisms by which SARS-CoV-2 suppresses the translation of host mRNA in favour of its own. None are exclusive to this virus, but the combination, speed and magnitude of the effects seem unique, says Stern-Ginossar.

First, the virus eliminates the competition: viral protein Nsp1, one of the first proteins translated when the virus arrives, recruits host proteins to systematically chop up all cellular mRNAs that don't have a viral tag. When Stern-Ginossar's team put that same tag on the end of a host mRNA, the mRNA was not chopped up<sup>[12](#)</sup>.

Second, infection reduces overall protein translation in the cell by 70%. Nsp1 is again the main culprit, this time physically blocking the entry channel of ribosomes so mRNA can't get inside, according to work from two research teams<sup>[13,14](#)</sup>. The little translation capacity that remains is dedicated to viral RNAs, says Stern-Ginossar.

Finally, the virus shuts down the cell's alarm system. This happens in numerous ways, but Stern-Ginossar's team identified one clear mechanism for SARS-CoV-2: the virus prevents cellular mRNA from getting out of the nucleus, including instructions for proteins meant to alert the immune system to infection. A second team confirmed this finding, and again pointed to Nsp1: the protein seems to jam up exit channels in the nucleus so nothing can escape<sup>[15](#)</sup>.

Because gene transcripts can't get out of the nucleus, the infected cells don't release many interferons — these are signalling proteins that alert the immune system to the presence of a virus. SARS-CoV-2 is particularly efficient at shutting down this alarm system: compared with other respiratory viruses, including SARS-CoV and respiratory syncytial virus, SARS-CoV-2 infection induces significantly lower levels of interferons<sup>[16](#)</sup>. And this June, researchers [reported mutations in the Alpha variant](#) that seem to enable it to subdue interferon production even more efficiently<sup>[17](#)</sup>.

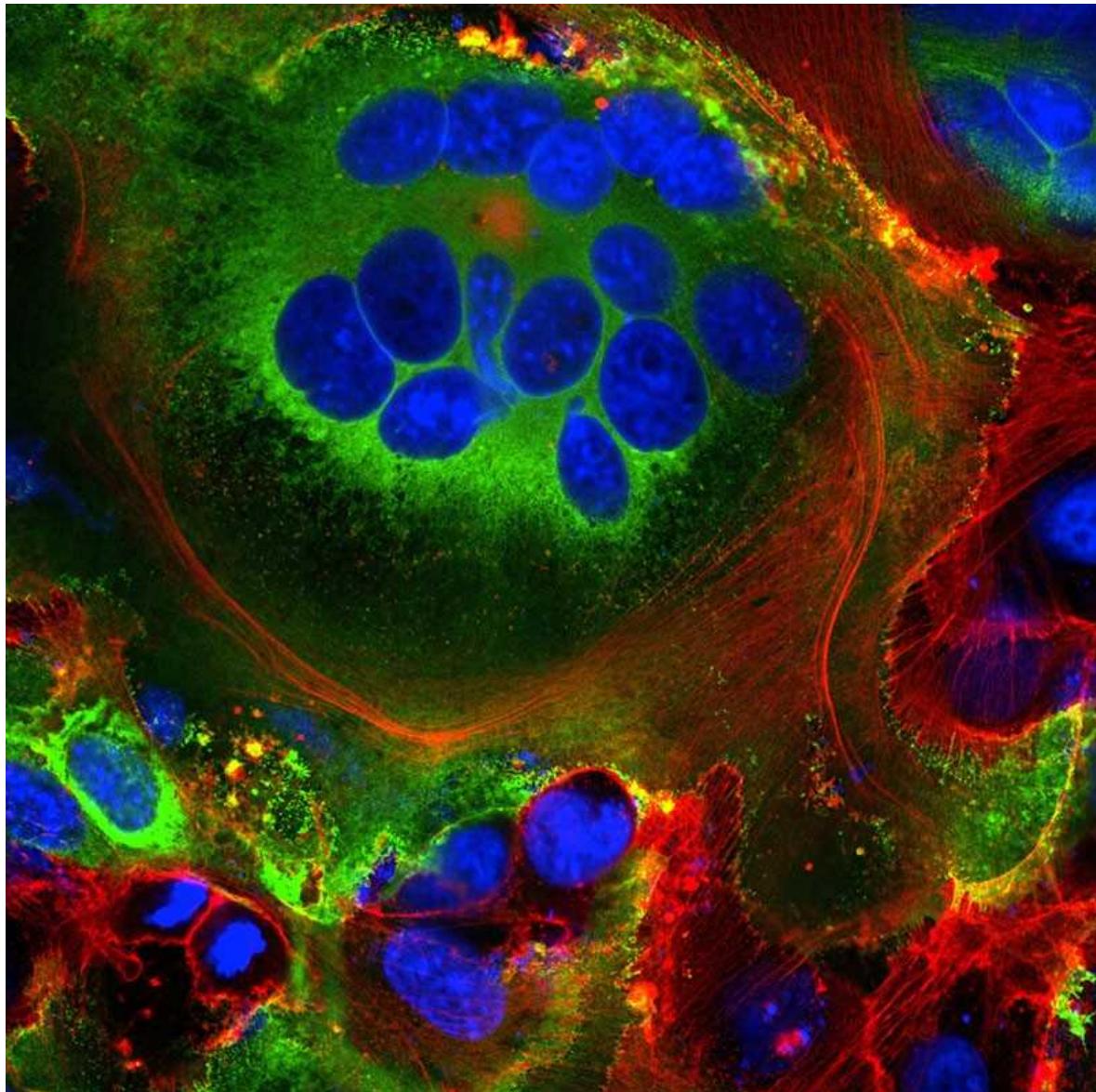
“It’s clear that SARS-CoV-2 is a very fast virus that has a unique ability to prevent our immune system from recognizing and combating infection in the

first stages,” says Stern-Ginossar. By the time the immune system does realize there is a virus, there is so much of it that immune-response proteins sometimes flood the bloodstream at a faster rate than normal — which can cause damage. Doctors saw early in the pandemic that some people with COVID-19 who become very ill are harmed by an overactive immune response to SARS-CoV-2, as well as by the virus itself. Some proven treatments [work by dampening down this immune response](#).

## Renovation station

Once the virus has taken over host translation, it starts a home makeover, extensively remodelling the interior and exterior of the cell to its needs.

First, some of the newly made viral spike proteins travel to the surface of the cell and poke out of the host-cell membrane. There, they activate a host calcium-ion channel, which expels a fatty coating onto the outside of the cell — the same coating found on cells that naturally fuse together, such as muscle cells. At this point, the infected cell fuses to neighbouring cells expressing ACE2, developing into massive individual respiratory cells filled with up to 20 nuclei.



Fused cell structures (syncytia) seen in cells expressing the SARS-CoV-2 spike protein (green). Nuclei are in blue and the cell skeleton is in red. Credit: Mauro Giacca

These fused structures, called syncytia, are induced by viral infections such as HIV and herpes simplex virus, but not by the SARS virus, says molecular biologist Mauro Giacca at King's College London, who led the team that published the finding in April<sup>18</sup>. He hypothesizes that forming syncytia allows infected cells to thrive for long periods of time, churning out more and more virions. "This is not a hit-and-run virus," he says. "It persists." A second team, led by researcher Qiang Sun at the Chinese Academy of

Medical Sciences in Beijing, found that some COVID-19-infected cells even form syncytia with lymphocytes — one of the body's own immune cells<sup>[19](#)</sup>. This is a known mechanism of immune evasion by tumour cells, but not by viruses. It suggests that infected cells avoid immune detection by simply grabbing on to and merging with nearby immune scouts.

On the inside of the cell, even more change is occurring. Like other coronaviruses, SARS-CoV-2 transforms the long, thin endoplasmic reticulum (ER), a network of flat membranes involved in protein synthesis and transport, into double-membrane spheres, as if the ER were blowing bubbles. These double-membrane vesicles (DMVs) might provide a safe place for viral RNA to be replicated and translated, shielding it from innate immune sensors in the cell, but that hypothesis is still being investigated.

Proteins involved in making DMVs could be good drug targets, because they seem to be necessary for viral replication. For instance, a host protein, TMEM41B, is needed to mobilize cholesterol and other lipids to expand the ER membranes so that all the virus parts will fit inside<sup>[20](#)</sup>. “When you take TMEM41B out, it has a major impact on infection,” says Vineet Menachery, a coronavirus researcher at the University of Texas Medical Branch in Galveston, who was involved in the research. The coronavirus transmembrane protein Nsp3 could also be a target: it creates a crown-like pore in the walls of the DMVs to shuttle out newly made viral RNA<sup>[21](#)</sup>.

Most viruses that have an outer wrapping, known as an envelope, form this feature by assembly directly at the edge of the cell, co-opting some of the cell's own plasma membrane on their way out. But newly made coronavirus proteins take a different path.

For years, evidence has suggested that coronaviruses are transported out of the cell through the Golgi complex, an organelle that works like a post office, packaging molecules in membranes and sending them off to other parts of the cell. There, the virus forms a lipid envelope from the Golgi complex's membrane; newly formed virions are then carried inside Golgi vesicles to the cell surface, where they are spat out of the cell, says virologist and cell biologist Carolyn Machamer at Johns Hopkins University in Baltimore, Maryland, who has studied coronaviruses for 30 years.

But in December, cell biologist Nihal Altan-Bonnet at the US National Heart, Lung, and Blood Institute in Bethesda, Maryland, and her colleagues reported that they had detected coronaviruses leaving the cell through lysosomes — cellular rubbish bins full of enzymes that break down cell parts<sup>22</sup>. Blocking the Golgi-based secretory pathway didn't seem to affect the amount of infectious virus being released, says Altan-Bonnet. Her team's evidence<sup>22</sup> suggests that viral proteins form an envelope by budding into the ER, then take over lysosomes to get out of the cell. The researchers are currently testing inhibitors that block the lysosomal exit process as potential antiviral candidates.

Leaving a cell through either the Golgi or lysosomes is slow and inefficient compared with budding out of a plasma membrane, so scientists don't know why SARS-CoV-2 does it. Machamer suspects that the lipid composition of a Golgi- or lysosome-derived envelope is somehow more beneficial to the virus than one from the plasma membrane. "If we understood this part a little bit better, there would be great opportunities for novel antiviral therapeutics," she says.

## Last slice

On the way out of the cell, one more event makes this virus into an infectious juggernaut: a quick snip at a site of five amino acids prepares the virus to strike its next target.

Where other coronaviruses have a single arginine amino acid at the junction of the S1 and S2 subunits of the spike, SARS-CoV-2 has a line of five amino acids: proline, arginine, arginine, alanine and arginine. "Because the site was unusual, we focused on it, and it turned out that, yes, the site is essential for invasion of lung cells," says Pöhlmann. In May 2020, he and his colleagues reported that a host-cell protein called furin recognizes and clips that string of amino acids — and the cut is "essential" for the virus to enter human lung cells efficiently<sup>23</sup>.

It's not the first time that researchers have identified a furin cleavage site on a virus; highly pathogenic avian influenza viruses also have it, says Barclay. When a colleague sent Barclay a strain of SARS-CoV-2 in culture that had

spontaneously lost the furin cleavage site, her team found that ferrets infected with this strain shed viral particles in lower amounts than did those infected with the pandemic strain, and did not transmit the infection to nearby animals<sup>9</sup>. At the same time as Barclay's team reported its results in a September 2020 preprint, a study in the Netherlands also found that coronavirus with an intact furin cleavage site enters human airway cells faster than do those without it<sup>24</sup>.

Furin is suspected to cut the site at some point during virion assembly, or just before release. The timing might explain why the virus exits through the Golgi or lysosomes, says Tom Gallagher, a virologist at Loyola University Chicago in Illinois. “The virus, once assembled, moves into an organelle where it can be bathed in the presence of the furin protease.”

By snipping the bond between the S1 and S2 subunits, the furin cut loosens up virion spike proteins so that during cell entry they respond to a second cut by TMPRSS2, which exposes the hydrophobic area that rapidly buries itself in a host-cell membrane, says Gallagher. If spikes are not pre-clipped by furin —and they aren’t always — they bypass TMPRSS2, and enter through the slower endosomal pathway, if at all.



[The race for antiviral drugs to beat COVID — and the next pandemic](#)

Two coronavirus variants, Alpha and Delta, have altered furin cleavage sites. In the Alpha variant, the initial proline amino acid is changed to a histidine (P681H); in the Delta variant, it is changed to an arginine (P681R). Both changes make the sequence less acidic, and the more basic the string of amino acids, the more effectively furin recognizes and cuts it, says Barclay. “We would hypothesize that this is the virus getting even better at transmitting.”

More furin cuts mean more spike proteins primed to enter human cells. In SARS-CoV, less than 10% of spike proteins are primed, says Menachery, whose lab group has been quantifying the primed spike proteins but is yet to publish this work. In SARS-CoV-2, that percentage rises to 50%. In the Alpha variant, it’s more than 50%. In the highly transmissible Delta variant, the group has found, greater than 75% of spikes are primed to infect a human cell.

## Known unknowns

The scientific community is still scratching the surface of its understanding of SARS-CoV-2. Key unknowns include the number of ACE2 receptors needed to bind to each spike protein; when exactly the S2 site is cleaved by TMPRSS2; and the number of spikes needed for virus–cell membrane fusion, says McLellan — and that’s just for entry. In April 2020, a team at the University of California, San Francisco, identified at least 332 interactions between SARS-CoV-2 and human proteins<sup>25</sup>.

It is not easy to keep pace with the quickly mutating virus. Most mutations so far are associated with how effectively the virus spreads, not with how much the virus damages the host, experts agree. This month, a study reported that [the Delta variant grew more rapidly and at higher levels](#) inside people’s lungs and throats than did earlier versions of the virus<sup>26</sup>.

But it is not yet certain how Delta’s mutations have supercharged the variant in this way, says Stern-Ginossar. “This is something many labs are trying to figure out.”

*doi: <https://doi.org/10.1038/d41586-021-02039-y>*

## References

1. 1.

Casalino, L. *et al.* *ACS Cent. Sci.* **6**, 1722–1734 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

2. 2.

Ke, Z. *et al.* *Nature* **588**, 498–502 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

3. 3.

Turoňová, B. *et al.* *Science* **370**, 203–208 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

4. 4.

Nguyen, H. L. *et al.* *J. Phys. Chem. B* **124**, 7336–7347 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

5. 5.

Shang, J. *et al.* *Nature* **581**, 221–224 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

6. 6.

Gobeil, S. M.-C. *et al.* *Science* <https://doi.org/10.1126/science.abi6226> (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

7. 7.

Khateeb, J., Li, Y. & Zhang, H. *Crit. Care* **25**, 244 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

8. 8.

Hoffmann, M. *et al.* *Cell* **181**, 271–280 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

9. 9.

Peacock, T. P. *et al.* *Nature Microbiol.* **6**, 899–909 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

10. 10.

Wang, M. *et al.* *Cell Res.* **30**, 269–271 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

11. 11.

Gunst, J. D. *et al.* *EClinicalMedicine* **35**, 100894 (2021).

[Article](#) [Google Scholar](#)

12. 12.

Finkel, Y. *et al.* *Nature* **594**, 240–245 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

13. 13.

Schubert, K. *et al.* *Nature Struct. Mol. Biol.* **27**, 959–966 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

14. 14.

Thoms, M. *et al.* *Science* **369**, 1249–1255 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

15. 15.

Zhang, K. *et al.* *Sci. Adv.* **7**, eabe7386 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

16. 16.

Blanco-Melo, D. *et al.* *Cell* **181**, 1036–1045 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

17. 17.

Thorne, L. G. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.06.06.446826> (2021).

18. 18.

Braga, L. *et al.* *Nature* **594**, 88–93 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

19. 19.

Zhang, Z. *et al.* *Cell Death Differ.* <https://doi.org/10.1038/s41418-021-00782-3> (2021).

[Article](#) [Google Scholar](#)

20. 20.

Trimarco, J. D. *et al.* *PLoS Pathog.* **17**, e1009599 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

21. 21.

Wolff, G. *et al.* *Science* **369**, 1395–1398 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

22. 22.

Ghosh, S. *et al.* *Cell* **183**, 1520–1535 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

23. 23.

Hoffmann, M., Kleine-Weber, H. & Pöhlmann, S. *Mol. Cell* **78**, 779–784 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

24. 24.

Mykytyn, A. Z. *et al.* *eLife* **10**, e64508 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

25. 25.

Gordon, D. E. *et al.* *Nature* **583**, 459–468 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

26. 26.

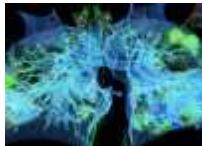
Li, B. *et al.* Preprint at medRxiv  
<https://doi.org/10.1101/2021.07.07.21260122> (2021).

[Download references](#)

## Related Articles



- [The sprint to solve coronavirus protein structures — and disarm them with drugs](#)



- [How a rampant coronavirus variant blunts our immune defences](#)



- [Profile of a killer: the complex biology powering the coronavirus pandemic](#)

## Subjects

- [SARS-CoV-2](#)
- [Virology](#)

**nature** careers

[Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- **FACULTY POSITION IN MICROBIOME BIOCHEMISTRY**

Duke University

Durham, NC, United States

- **Bioinformatician (m/f/div)**

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- **Student Research Assistant / HIWI (m/f/x) for Literature Research in the Field of Soil Microbiology**

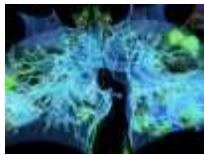
Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

-  **The sprint to solve coronavirus protein structures — and disarm them with drugs**



- [How a rampant coronavirus variant blunts our immune defences](#)



- [Profile of a killer: the complex biology powering the coronavirus pandemic](#)

## Subjects

- [SARS-CoV-2](#)
- [Virology](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02039-y>.

# Opinion

- **[Richard R. Ernst \(1933—2021\)](#)** [ 26 July 2021]  
Obituary • Nobel prizewinner who revolutionized nuclear magnetic resonance.
- **[Biodiversity needs every tool in the box: use OECMs](#)** [ 26 July 2021]  
Comment • To conserve global biodiversity, countries must forge equitable alliances that support sustainability in traditional pastoral lands, fisheries-management areas, Indigenous territories and more.
- **[Scientists — reach across the aisle, Republicans could be back soon](#)** [ 27 July 2021]  
Correspondence •
- **[Global climate models do not need more behavioural science](#)** [ 27 July 2021]  
Correspondence •
- **[Peru to punish bending of clinical-trial rules](#)** [ 27 July 2021]  
Correspondence •
- **[Europe's reformed agricultural policy disappoints](#)** [ 27 July 2021]  
Correspondence •

- OBITUARY
- 26 July 2021

# Richard R. Ernst (1933—2021)

Nobel prizewinner who revolutionized nuclear magnetic resonance.

- [Kurt Wüthrich](#) <sup>9</sup>

## 1. [Kurt Wüthrich](#)

1. Kurt Wüthrich is a professor of biophysics at ETH Zurich in Switzerland and Cecil H. and Ida M. Green Professor of Structural Biology at Scripps Research in La Jolla, California. He won a share of the Nobel Prize in Chemistry in 2002 for developing NMR techniques for structural biology.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)

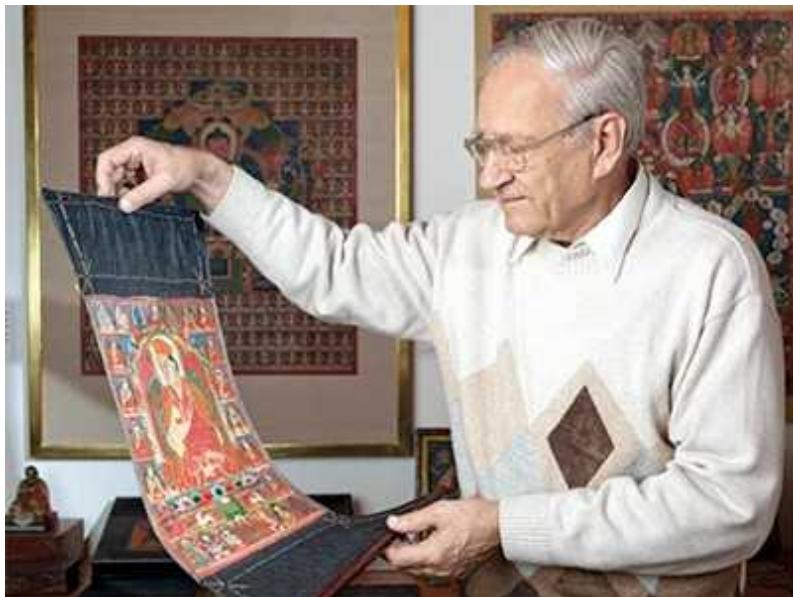


Credit: Tina Ruisinger/Visum/eyevine

Richard R. Ernst turned applications of nuclear magnetic resonance (NMR) into methods for daily use in chemistry, structural biology and medical diagnosis. His introduction of Fourier transform-nuclear magnetic resonance (FT-NMR) in the 1960s paved the way for magnetic resonance imaging (MRI) to be widely used in medical diagnosis and in materials research.

As well as gaining him the Nobel Prize in Chemistry in 1991, Ernst's work underpinned the achievements rewarded by the two NMR Nobel prizes of the twenty-first century so far — mine included. He has died, aged 87.

Richard Robert Ernst was born in Winterthur, Switzerland. His family was part of the city's establishment, with a history stretching back at least 500 years. His father taught architecture at the Winterthur 'Technikum' (today the University of Applied Sciences) and was a high-ranking officer in the Swiss army. In this distinguished milieu, musical gatherings had an important place; Richard played the cello. His plans to become a composer were set aside at 13, when he found a box of chemicals in the attic, left by a late uncle, and began doing experiments. After studying chemistry at the Swiss Federal Institute of Technology (ETH) in Zurich, he pursued doctoral studies in the new field of NMR.



### A man of many dimensions

Shortly after the Second World War, Felix Bloch and Edward Mills Purcell had described NMR spectroscopy, sharing the 1952 Nobel Prize in Physics. The principle is this: in a magnetic field, the moments of certain atomic nuclei of a substance come into alignment. Applying radio-frequency irradiation perturbs this alignment, causing the atoms to give off electromagnetic signals that are unique to the chemical composition and molecular structure of the substance.

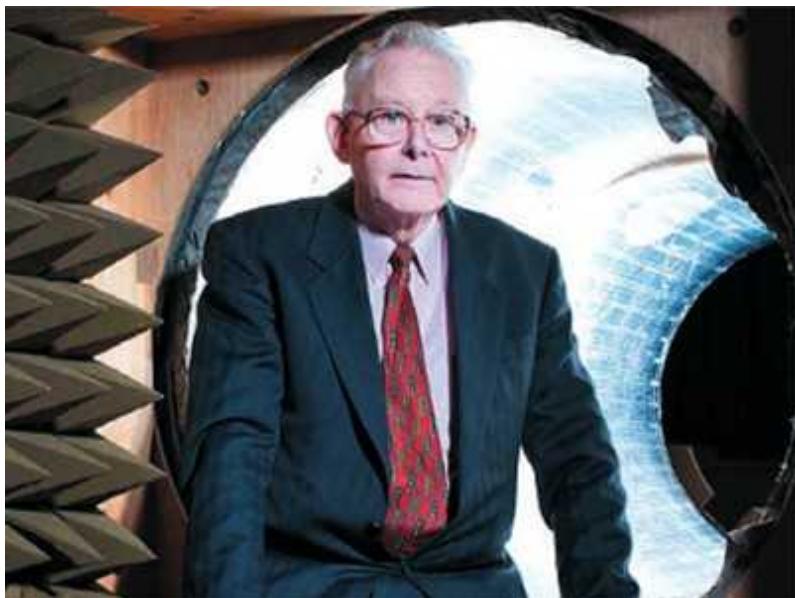
Soon after its discovery, NMR became widely used in physics, for example in particle physics, solid-state physics and in studies of phase transitions. Applications as an analytical tool in chemistry started in the 1950s with

continuous-wave NMR experiments, which had limited impact because their sensitivity was too low.

Working in the mid-1960s at the then leading manufacturer of NMR instrumentation, the Varian Company in Palo Alto, California, Ernst experimented with applying a single, hard pulse that excites the entire spectrum, instead of sweeping the NMR spectra with weak continuous irradiation. He then subjected the resulting signal to Fourier transformation to generate a high-resolution spectrum.

Neither the materials used in NMR instruments nor the computer facilities of the 1960s were up to the demands of the new technique. So it took another decade before FT-NMR found applications in chemistry and structural biology. FT-NMR also provided the foundation for the more-informative 2D NMR in the late 1970s; related mathematical procedures made MRI efficient for medical diagnosis.

In the late 1960s, Ernst returned to ETH Zurich, where he spent the rest of his career. Over four decades, he inspired generations of scientists as a professor, and as a lecturer around the world.



[Peter Mansfield \(1933–2017\)](#)

His graduate students and postdoctoral fellows witnessed him diving again and again into the depths of a theory or a physical-chemical effect. In great mathematical detail and with unique didactic skills, using his arms to illustrate how the magnetization in an NMR sample behaves under the influence of radio-frequency fields, say, he would generate rigorous formalisms to capture all nuances of a phenomenon. Thanks to the Ernst magic, the NMR community evolved from a few hundred scientists in the early 1970s to gatherings of thousands in a range of specialized areas of NMR today.

Ernst described himself as a “work-addict”. A toolmaker, he liked to provide others with new ways to solve problems. Yet he also believed that scientists should have a wide range of interests outside their field. He played the cello throughout his life, and had an encyclopedic knowledge of classical music. In his later years, he wrote a frank and insightful autobiography (translated as *Searching and Researching*, published this year).

On their return from California to Switzerland in 1968, Ernst and his wife travelled to Tibet, where they became fascinated by the local scroll paintings, and in the following decades they assembled a rich collection of Tibetan art. Ernst’s joy in this culture is touchingly apparent in his autobiography, which features beautiful illustrations of the works. After he retired in 1998, Ernst devoted his time and energy, as well as the Nobel prize money, to expanding this collection. In his home, he used cutting-edge spectroscopic techniques (not NMR!) to examine the paint on his precious Tibetan scrolls to trace their provenance.

As a colleague on the faculty of ETH Zurich, I had the privilege of a close scientific collaboration with Ernst from 1976 to 1986, which resulted in the introduction of multi-dimensional NMR to chemistry laboratories and to structural biology. I knew Richard as a strikingly intelligent, conscientious and gentle colleague and friend. Together with so many who are indebted to him for having enriched their scientific lives, I keep fond memories of the good times spent with this great scientist.

*Nature* **595**, 645 (2021)

doi: <https://doi.org/10.1038/d41586-021-02040-5>

## Competing Interests

The author declares no competing interests.

## Related Articles



- [A man of many dimensions](#)



- [Nicolaas Bloembergen \(1920–2017\)](#)



- [Peter Mansfield \(1933–2017\)](#)

## Subjects

- [History](#)
- [Chemistry](#)
- [Structural biology](#)

**nature** careers

## [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- **FACULTY POSITION IN MICROBIOME BIOCHEMISTRY**

Duke University

Durham, NC, United States

- **Bioinformatician (m/f/div)**

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- **Student Research Assistant / HIWI (m/f/x) for Literature Research in the Field of Soil Microbiology**

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- [A man of many dimensions](#)



- [Nicolaas Bloembergen \(1920–2017\)](#)



- [Peter Mansfield \(1933–2017\)](#)

# Subjects

- [History](#)
  - [Chemistry](#)
  - [Structural biology](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02040-5>

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

- COMMENT
- 26 July 2021

# Biodiversity needs every tool in the box: use OECMs

To conserve global biodiversity, countries must forge equitable alliances that support sustainability in traditional pastoral lands, fisheries-management areas, Indigenous territories and more.

- [Georgina G. Gurney](#)<sup>0</sup>,
- [Emily S. Darling](#)<sup>1</sup>,
- [Gabby N. Ahmadia](#)<sup>2</sup>,
- [Vera N. Agostini](#)<sup>3</sup>,
- [Natalie C. Ban](#)<sup>4</sup>,
- [Jessica Blythe](#)<sup>5</sup>,
- [Joachim Claudet](#)<sup>6</sup>,
- [Graham Epstein](#)<sup>7</sup>,
- [Estradivari](#)<sup>8</sup>,
- [Amber Himes-Cornell](#)<sup>9</sup>,
- [Harry D. Jonas](#)<sup>10</sup>,
- [Derek Armitage](#)<sup>11</sup>,
- [Stuart J. Campbell](#)<sup>12</sup>,
- [Courtney Cox](#)<sup>13</sup>,
- [Whitney R. Friedman](#)<sup>14</sup>,
- [David Gill](#)<sup>15</sup>,
- [Peni Lestari](#)<sup>16</sup>,
- [Sangeeta Mangubhai](#)<sup>17</sup>,
- [Elizabeth McLeod](#)<sup>18</sup>,
- [Nyawira A. Muthiga](#)<sup>19</sup>,
- [Josheena Naggea](#)<sup>20</sup>,

- [Ravaka Ranaivoson](#)<sup>21</sup>,
- [Amelia Wenger](#)<sup>22</sup>,
- [Irfan Yulianto](#)<sup>23</sup> &
- [Stacy D. Jupiter](#)<sup>24</sup>

1. [Georgina G. Gurney](#)

1. Georgina G. Gurney is a senior research fellow in environmental social science at the Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Australia.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Emily S. Darling

1. Emily S. Darling is director, Coral Reef Conservation, Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Gabby N. Ahmadia

1. Gabby N. Ahmadia is director, Marine Conservation Science, Ocean Conservation, World Wildlife Fund.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Vera N. Agostini

1. Vera N. Agostini is deputy director, Fisheries and Aquaculture Division, Food and Agriculture Organization of the United Nations.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Natalie C. Ban

1. Natalie C. Ban is associate professor in environmental studies at the University of Victoria, Canada.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Jessica Blythe

1. Jessica Blythe is assistant professor in environmental sustainability at Brock University, St Catharines, Canada.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Joachim Claudet

1. Joachim Claudet is a senior researcher at the National Center for Scientific Research, CRIODE, France.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Graham Epstein

1. Graham Epstein is a postdoctoral research associate at the School of Politics, Security and International Affairs at the University of Central Florida, Orlando, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## 9. Estradivari

1. Estradivari is a researcher at the Leibniz Center for Tropical Marine Research (ZMT), Germany, and a conservation research manager, World Wildlife Fund Indonesia.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## 10. Amber Himes-Cornell

1. Amber Himes-Cornell is a fisheries officer, Fisheries and Aquaculture Division, Food and Agricultural Organization of the United Nations.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## 11. Harry D. Jonas

1. Harry D. Jonas is an international lawyer at Future Law, Kota Kinabalu, Malaysia, and co-chair of the IUCN WCPA Specialist Group on Other Effective Area-based Conservation Measures.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## 12. Derek Armitage

1. Derek Armitage is a professor in the School of Environment, Resources and Sustainability, University of Waterloo, Canada.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## 13. Stuart J. Campbell

1. Stuart J. Campbell is senior director, RARE Indonesia.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Courtney Cox

1. Courtney Cox is senior director, Rare, Washington DC, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Whitney. R. Friedman

1. Whitney R. Friedman is a postdoctoral fellow at the University of California, Santa Barbara, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. David Gill

1. David Gill is an assistant professor of marine science and conservation at Duke University, Durham, North Carolina, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Peni Lestari

1. Peni Lestari is a socioeconomic marine specialist, Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Sangeeta Mangubhai

1. Sangeeta Mangubhai is director, Fiji Program, Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Elizabeth McLeod

1. Elizabeth McLeod is global reef lead, The Nature Conservancy, Arlington, Virginia, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Nyawira A. Muthiga

1. Nyawira A. Muthiga is director, Marine Conservation Program, Kenya Program, Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Josheena Naggea

1. Josheena Naggea is a PhD candidate at Stanford University, California, USA, studying conservation in her home country of Mauritius.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Ravaka Ranaivoson

1. Ravaka Ranaivoson is marine director, Madagascar Program, Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

#### 23. Amelia Wenger

1. Amelia Wenger is a research fellow in the School of Earth and environmental sciences at the University of Queensland, Brisbane, Australia, and a conservation scientist at the Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

#### 24. Irfan Yulianto

1. Irfan Yulianto is a researcher and lecturer at Institut Pertanian Bogor University, Indonesia, and a Senior Manager, Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

#### 25. Stacy D. Jupiter

1. Stacy D. Jupiter is Melanesia regional director at the Wildlife Conservation Society.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



Customary fishing-rights holders from Totoya Island, Fiji, marking a sacred reef area as a no-fishing zone.Credit: Keith Ellenbogen

Global support is growing for the  $30 \times 30$  movement — a goal to conserve 30% of the planet by 2030. In May, the G7 group of wealthy nations endorsed the commitment to this target that had been made by more than 50 countries in January. It is likely to be the headline goal when parties to the Convention on Biological Diversity (CBD) meet to discuss the latest global conservation agreement in May 2022 in Kunming, China.

So where do the sacred forests of Estonia or shipwrecks in North America's Great Lakes come in? What do these share with managed fishing grounds in Fiji and bighorn-sheep hunting areas in Mexico? All have the potential to be recognized using a conservation policy tool called other effective area-based conservation measures, or OECMs. Together with protected areas — from Malaysia's Taman Negara National Park to the Cerbère-Banyuls Marine Reserve in southern France — OECMs could help to achieve the 30% target.

Devised in 2010 and defined in 2018, the OECM tool is little known outside specialist circles. Less than 1% of the world's land and freshwater

environments and less than 0.1% of marine areas are currently covered under this designation. Meanwhile, biodiversity is in free fall and protected areas alone can't stem the loss. Both designations are among the international policy instruments being negotiated ahead of the CBD conference.

We call on the CBD parties and the conservation community of policymakers, scientists, practitioners and donors to study and use OECMs more, alongside protected areas. This policy tool can advance equitable and effective conservation if CBD parties stay true to the convention's intent to sustain biodiversity rather than 'achieve' area-based targets. But more groundwork must be laid to realize its potential.

Improvements are needed in research, policy and practice. Local managers and CBD parties need better ways to assess whether potential OECMs contribute to sustaining biodiversity, so that areas are properly designated. The conservation community needs to develop processes to ensure that OECM recognition strengthens, rather than displaces, existing local governance. And researchers need to articulate the value of OECMs to encourage policymakers to use them.

## Bigger toolkit

Protected areas have expanded rapidly in the past 10 years, and now cover 15.7% of the world's land and fresh water, and 7.7% of the marine realm. Defined by the CBD as areas designated or regulated and managed for biodiversity conservation, they are an essential conservation approach. But some have failed to be equitable or effective: aligning biodiversity goals with local values, needs and governance can be difficult in some contexts<sup>1,2</sup>. This conflict can lead to inequities, non-compliance and poor biodiversity outcomes.



### Indigenous rights vital to survival

OECMs can have an important and complementary role<sup>3</sup>. The tool recognizes managed areas that sustain biodiversity, irrespective of their objective. OECM recognition can support Indigenous and local communities in managing their lands and seas — be it for hunting, fishing or other cultural practices — while conserving nature. It opens up new conservation opportunities in landscapes where there is relatively light human usage, such as pastoralism with a low density of livestock. These regions make up nearly 56% of the world's lands, and contain more Key Biodiversity Areas — sites of global importance to biodiversity — than do remaining large wild areas<sup>4</sup>. So, management approaches that accommodate the ways people use landscapes and seascapes are crucial.

Some managed areas do not safeguard biodiversity<sup>5</sup>. But there is a wealth of evidence suggesting that many do. For example, a study of the Peruvian Amazon found that Indigenous peoples' territories were, on average, more effective than state-governed protected areas at preventing deforestation<sup>6</sup>. A review of 61 areas managed under territorial-use rights in fisheries in Chile found positive effects on biodiversity; some had levels of fish biomass and biodiversity that were comparable to those in a protected area that restricts all fishing<sup>7</sup>. And abandonment of agricultural management systems

involving low-intensity farming methods in Europe — such as traditional haymaking in Romania — has been linked repeatedly to biodiversity loss<sup>8</sup>.

Perhaps many of these could be recognized as OECMs (see ‘Conservation potential’). Doing so depends on the consent of the relevant governing bodies, and whether the managed area meets the CBD’s definition and criteria for OECMs, including demonstrated or expected biodiversity outcomes.

## CONSERVATION POTENTIAL

Areas managed by local people can conserve nature while sustaining livelihoods and cultural practices. Recognizing them as OECMs\* would allow their contribution to sustaining biodiversity to be counted.

### Amazon rainforest

Almost one-third of the rainforest is Indigenous territories, of which 80% is outside protected areas.

**30%** is Indigenous territories



### Mountain farmland in Europe

Almost one-fifth of the European Union's mountain area is highly biodiverse and depends on low-intensity farming practices.

**17%** is 'High Nature Value farmland'



### Community forests in Nepal

Almost one-quarter of the country's forests are governed by 33% of the rural population. They sustain healthy ecosystems and the well-being of inhabitants.

**23%** is community forestry areas



### Wildlife zones in Mexico

Wildlife Management Units (UMAs) cover nearly one-fifth of the country's territory. They are governed privately or communally for sustainable uses such as hunting.

**19%** is covered by UMAs



\*OECMs, other effective area-based conservation measures.

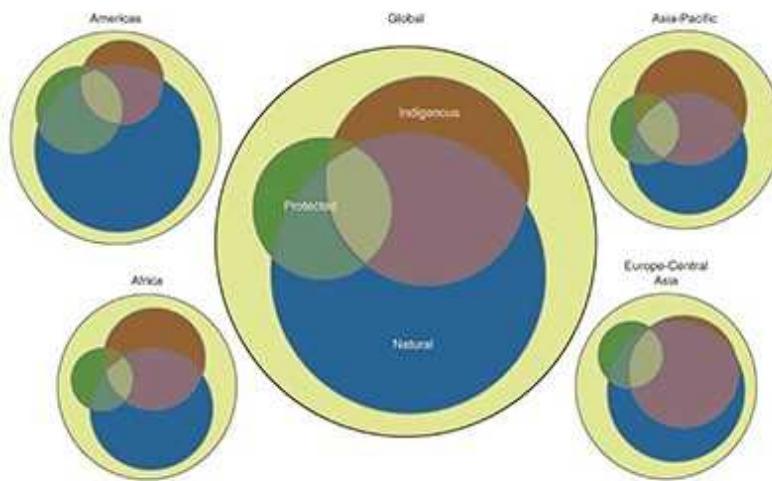
©nature

## Equity

OECMs can help to ensure that international conservation targets are legitimate to the many and diverse actors required to turn the tide on biodiversity loss.

Too often, the costs of conservation are felt locally while many of the benefits are shared globally — from carbon sequestration to preserving genetic resources. For instance, rainforest conservation, including a protected area, in the Ankeniheny-Zahamena Corridor in Madagascar meant that local farmers of vanilla, cloves and rice bore opportunity costs representing 27–84% of their average annual household income. The protection scheme is intended to cut 10 million tonnes of carbon dioxide emissions over 10 years<sup>9</sup>.

Such inequities can occur when protected areas do not prioritize local values and needs. Although protected areas can have multiple objectives, the widely followed guidance from the International Union for Conservation of Nature (IUCN) advises that nature conservation should retain priority over all other objectives. This can alienate people who manage areas for other reasons. Even in the instance of Indigenous Protected Areas in Australia, which have resulted in an array of social and biodiversity benefits, the IUCN definition can undermine Indigenous Australians' conceptualization of humans as part of nature, which underpins their governance systems<sup>2</sup>. This stands in contrast to the Western world view of humans as distinct from nature — a concept that is embedded in the IUCN definition and conservation more generally<sup>2,3</sup>.



[A spatial overview of the global importance of Indigenous lands for conservation](#)

However, OECMs need not have conservation as an objective. This means that they can be used to recognize the contributions of a myriad of actors who manage areas that sustain nature, regardless of why they do so.

Indigenous peoples, for instance, manage 37% of the world's natural lands<sup>10</sup> for many reasons, such as maintaining rights, harvesting and cultural identity<sup>2,10,11</sup>. Recognition of Indigenous territories as OECMs could help to overcome current challenges of insecure rights, insufficient funding and exclusion of these communities from decision-making<sup>12</sup>. For example, Indonesia has initiated revisions to its conservation laws to accommodate coastal OECMs, which could provide opportunities for Indigenous and local communities to gain legal recognition of their rights to use and manage fisheries.

OECMs can thus ensure a more equitable approach to conservation decision-making. They enable the participation of those who govern areas that sustain biodiversity but who are currently not involved in decision-making. For example, fisheries-management organizations have rebuilt some fish stocks, contributing to biodiversity and wider ecosystem health, yet the fisheries and conservation sectors are often divided<sup>13</sup>. OECMs can foster cooperation between sectors, and encourage the participation of fisheries-management organizations in conservation decision-making.

## Effectiveness

Collectively, alongside protected areas, OECMs can increase the effectiveness of the global conservation system in four key ways.

First, they support management that is tailored to its context<sup>14</sup>, and aligned with local values, governance and traditional knowledge systems. This fosters the local leadership, support and compliance that are key to biodiversity benefits<sup>14</sup>. For example, in Mo'orea, French Polynesia, protected areas that restricted all fishing did not meet fishers' needs, leading to non-compliance and relatively little change in the density and biomass of coral-reef fish<sup>15</sup>. Conversely, a management area in Labrador, Canada, implemented at the behest of crab fishers, maintained the fishery and increased the biomass of fish species such as Atlantic cod (*Gadus morhua*)

and other, non-target species<sup>16</sup>. This area seems likely to meet the OECM criteria.



Estonia's sacred groves are protected for their spiritual significance. Credit: Toomas Tuul/FOCUS/Universal Images Group via Getty

Second, OECMs, together with protected areas, can help to ensure a well-connected conservation network in which all elements of biodiversity are represented and in which ecological processes, such as species movements, are sustained. For instance, Kenya's wildlife conservancies provide geographical bridges between protected areas for the movement of wildlife such as zebras, and have potential to be recognized as OECMs.

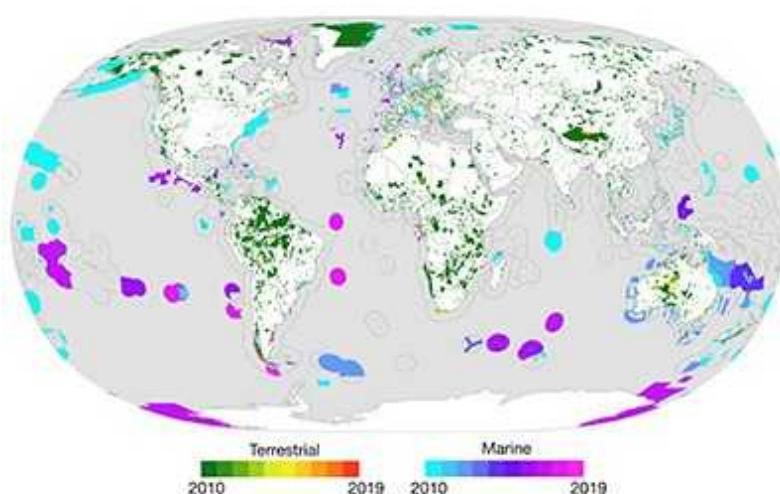
Third, OECMs can increase the diversity of tools in the global conservation system. This bolsters the system's resilience to social and biophysical shifts, including climate change<sup>14</sup>. Redundancy in governance arrangements can help to mitigate risks associated with the current reliance on government-led protected areas, which are vulnerable to shifts in national priorities. For example, in 2017, the Bears Ears National Monument, a protected area in Utah, was downsized by 85% to make way for oil and gas exploration under a former US presidential administration.

Fourth, OECMs help to bring conservation outcomes into focus. A key criterion for official designation is demonstrated or expected biodiversity

outcomes, such as the restoration of a crucial habitat. This is not the case for protected areas, where a focus on coverage has, in some cases, led to expansion with scant biodiversity gains<sup>4</sup>.

## Five steps

Key concerns remain about the misuse of OECM recognition. CBD parties might use it to meet commitments without actually tackling biodiversity loss. For example, in 2017, Canada increased the marine area it planned to report almost sixfold, by reclassifying 51 fishery closures as OECMs<sup>17</sup>. This decision was criticized on the grounds of insufficient evidence that these areas sustain biodiversity. Another concern is that, despite the focus on equity, any attempts to influence local governance could be perceived as a ‘land grab’ or ‘sea grab’ by external actors such as national governments, foreigners or international organizations. For example, the establishment of some privately owned protected areas in southern Chile has been suggested to have involved coercion and intimidation of smallholder farmers.



### Area-based conservation in the twenty-first century

The conservation community needs to take the following five steps to overcome these key challenges to using the OECM policy tool.

**Show that they work.** The 2019 [IUCN Guidelines for Recognizing and Reporting OECMs](#) provide clear criteria for identifying managed areas that are suitable for a full assessment against the CBD's definition. However, research is needed on how to meet the crucial criteria of demonstrated or expected *in situ* conservation of biodiversity. This is challenging and resource-intensive, especially because of the variety of actors involved. Ideas based in Western science might not align with the knowledge systems of all involved.

Guidelines should build on existing approaches for evaluation, such as the IUCN Green List for Protected and Conserved Areas and the Indicators of Resilience in Socio-ecological Production Landscapes (SEPLs). They should include recent advances focused on outcomes<sup>18</sup> and should be tailored to different types of managed area. To ensure that these are salient, credible and legitimate to those governing OECMs, they should be co-produced by groups such as rights holders, civil-society organizations, government and industry, as well as by academics from various disciplines. This transdisciplinary approach is growing rapidly, with examples ranging from management at the national level (such as New Zealand's Sustainable Seas National Science Challenge) to the monitoring of coral reefs as social-ecological systems<sup>19</sup>.



Pastoral lands in Africa are often governed to maintain sustainable grazing.Credit: Steve Pastor

**Strengthen existing local governance.** Many rights holders have raised concerns that formal recognition of their managed areas for conservation might infringe their rights. For example, few communities in Fiji have had their fisheries-management areas recognized under national conservation laws, because that currently requires the communities to waive their customary rights<sup>20</sup>.

Engaging with global conservation processes might also erode self-determination through the imposition of external world views<sup>2,3,12</sup>. Although OECMs open the door to recognizing diverse relations between humans and nature, it is crucial that the need for demonstrated or expected biodiversity outcomes does not diminish other priorities and values.

OECM recognition must strengthen existing local governance, rather than displace or substantially alter it. This will require guidelines to be informed by principles of procedural equity and tailored to different types of managed area. Their development should draw on existing approaches such as the Australian Indigenous-led Healthy Country Planning and Our Knowledge, Our Way guidelines, which have underpinned engagement with the national carbon sequestration scheme<sup>11</sup>.

**Secure funding.** Funding for recognizing and reporting OECMs should be made available to ensure costs are not a barrier or burden for under-resourced groups. A prominent role for OECMs in the next CBD agreement will help — this policy guides conservation investments from nations and donors.



### [Sixty years of tracking conservation progress using the World Database on Protected Areas](#)

Importantly, the diversity of managed areas that OECMs encompass can provide funding opportunities beyond conventional conservation funders, whose resources for protected-area funding are already overstretched. Conservation practitioners should engage private sectors that manage areas that could be recognized as OECMs, and access funding earmarked for other priorities such as health and development. For example, the Watershed Interventions for Systems Health project in Fiji, which aims to reduce waterborne diseases using nature-based solutions, is supported by both conservation and public-health funding.

Conservation donors and practitioners should co-design new funding strategies for OECMs with those governing these areas. This will help to ensure that local priorities are supported. For example, Coast Funds, a unique conservation trust fund, was developed by First Nations people in collaboration with conservation practitioners and the forestry industry to support stewardship of the Great Bear Rainforest and Haida Gwaii regions of British Columbia, Canada.

**Agree on metrics.** The record of progress towards the CBD's area-based target, the World Database on Protected Areas, assumes that all reported protected areas have biodiversity conservation as a main objective. But some

CBD parties report areas that have other primary objectives, such as sustainable harvesting<sup>20</sup>. This leads to inaccurate accounting at the global level, and to misunderstanding of management actually occurring on the ground. Canada, among others, is developing legislation that demarcates protected areas and OECMs. But it is not clear whether all CBD parties will do the same.

Policymakers need to agree on targets that are based on outcomes — not just coverage — for both OECMs and protected areas. These might include, for example, changes in the populations of multiple species relative to a reference point. In constructing these targets, the conservation community should be guided by the development and health sectors, which have long used outcome targets. For example, the United Nations Sustainable Development Goal 1.2 aims to reduce at least by half the proportion of people living in multidimensional, regionally-defined poverty by 2030. A common currency of outcomes could alleviate concerns that there is an uneven burden of proof for the OECM and protected-area tools. It could also prevent the misuse of either to meet targets based on area without actually sustaining biodiversity.

**Include OECMs in other environmental agreements.** Addressing the interrelated global challenges of biodiversity loss, climate change and sustainability requires the coordination of policy across sectors. Right now, OECMs appear only in CBD policy. But they could contribute to the mandates of other intergovernmental initiatives. Policymakers should include OECMs alongside protected areas in international agreements such as the Sustainable Development Goals, new global climate agreements being negotiated under the UN convention on climate, and the emerging UN treaty on marine biodiversity in areas beyond national jurisdiction.

New targets negotiated at the upcoming CBD meeting will set the global conservation agenda over the next decade. If the steps we outline here are implemented, OECMs could be central to the transformations needed for a sustainable future for the planet.

*Nature* **595**, 646-649 (2021)

doi: <https://doi.org/10.1038/d41586-021-02041-4>

# References

1. 1.

Oldekop, J. A., Holmes, G., Harris, W. E. & Evans, K. L. *Conserv. Biol.* **30**, 133–141 (2016).

[PubMed](#) [Article](#) [Google Scholar](#)

2. 2.

Lee, E. *Antipode* **48**, 355–374 (2016).

[Article](#) [Google Scholar](#)

3. 3.

Jonas, H. D., Barbuto, V., Jonas, H. C., Kothari, A. & Nelson, F. *PARKS* **20**, 111–128 (2014).

[Article](#) [Google Scholar](#)

4. 4.

Ellis, E. C. *One Earth* **1**, 163–167 (2019).

[Article](#) [Google Scholar](#)

5. 5.

Donald, P. F. *et al.* *Conserv. Lett.* **12**, e12659 (2019).

[Article](#) [Google Scholar](#)

6. 6.

Schleicher, J., Peres, C. A., Amano, T., Llactayo, W. & Leader-Williams, N. *Sci. Rep.* **7**, 11318 (2017).

[PubMed](#) [Article](#) [Google Scholar](#)

7. 7.

Gelcich, S., Martínez-Harms, M. J., Tapia-Lewin, S., Vasquez-Lavin, F. & Ruano-Chamorro, C. *Conserv. Lett.* **12**, e12637 (2019).

[Article](#) [Google Scholar](#)

8. 8.

Lomba, A. *et al.* *Front. Ecol. Environ.* **18**, 36–42 (2020).

[Article](#) [Google Scholar](#)

9. 9.

Poudyal, M. *et al.* *PeerJ* **6**,e5106 (2018).

[PubMed](#) [Article](#) [Google Scholar](#)

10. 10.

Garnett, S. T. *et al.* *Nature Sustain.* **1**, 369–374 (2018).

[Article](#) [Google Scholar](#)

11. 11.

Ansell, J. *et al.* *Int. J. Wildland Fire* **29**, 371–385 (2019).

[Article](#) [Google Scholar](#)

12. 12.

Corson, C. *et al.* *Conserv. Soc.* **12**, 190–202 (2014).

[Article](#) [Google Scholar](#)

13. 13.

Hilborn, R. *Nature* **535**, 224–226 (2016).

[PubMed](#) [Article](#) [Google Scholar](#)

14. 14.

Carlisle, K. & Gruby, R. L. *Policy Stud. J.* **47**, 927–952 (2019).

[Article](#) [Google Scholar](#)

15. 15.

Thiault, L. *et al.* *Ecosphere* **10**, e02576 (2019).

[Article](#) [Google Scholar](#)

16. 16.

Kincaid, K. & Rose, G. *Can. J. Fish. Aqua. Sci.* **74**, 1490–1502 (2017).

[Article](#) [Google Scholar](#)

17. 17.

Lemieux, C. J. & Gray, P. A. *J. Environ. Stud. Sci.* **10**, 483–491 (2020).

[Article](#) [Google Scholar](#)

18. 18.

Geldmann, J. *et al.* *Conserv. Lett.* <https://doi.org/10.1111/conl.12792> (2021).

[Article](#) [Google Scholar](#)

19. 19.

Gurney, G. G. *et al.* *Biol. Conserv.* **240**, 108298 (2019).

[Article](#) [Google Scholar](#)

20. 20.

Govan, H. & Jupiter, S. *PARKS* **19**, 73–80 (2013).

[Article](#) [Google Scholar](#)

[Download references](#)

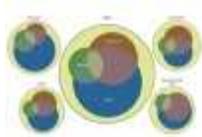
## Competing Interests

The authors declare no competing interests.

## Related Articles



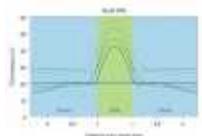
- [Indigenous rights vital to survival](#)



- [A spatial overview of the global importance of Indigenous lands for conservation](#)



- [Area-based conservation in the twenty-first century](#)



- [Marine conservation across protected area boundaries](#)



- [Sixty years of tracking conservation progress using the World Database on Protected Areas](#)

## Subjects

- [Biodiversity](#)
- [Policy](#)
- [Ecology](#)

**nature** careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

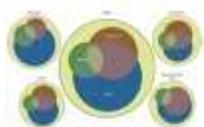
Leipzig, Germany

[Download PDF](#)

## Related Articles



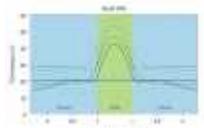
- [Indigenous rights vital to survival](#)



- [A spatial overview of the global importance of Indigenous lands for conservation](#)



- [Area-based conservation in the twenty-first century](#)



- [Marine conservation across protected area boundaries](#)



- [Sixty years of tracking conservation progress using the World Database on Protected Areas](#)

# Subjects

- [Biodiversity](#)
  - [Policy](#)
  - [Ecology](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02041-4>

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

- CORRESPONDENCE
- 27 July 2021

# **Scientists — reach across the aisle, Republicans could be back soon**

- [Imran Khan](#) 0
  - 1. [Imran Khan](#)
    - 1. University College London, UK.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

## [Download PDF](#)

The Democratic White House is seeking a US\$13.5-billion increase in US federal research and development budgets (*Nature* 594, 485; 2021). It claims to have put evidence at the heart of its COVID-19 strategy, and plans to decarbonize the economy. With the director of the Office of Science and Technology Policy now a cabinet-level position for the first time, President Joe Biden declares “science is back”.

But it would be a mistake for US researchers to just bask in the warm glow of a science-friendly government and jockey for the funding windfall.

Republicans could easily take control again. This might involve a clean sweep of all three branches of government. And it could start in just 18 months time, after the 2022 mid-term elections.

In the United States, just 20% of people on the political right trust scientists, compared with 62% on the left — the biggest such split in public opinion across 14 countries. Only 40% of Republican voters favour increases in science funding, compared with 60% of Democrats.

Now is the time to try to bridge the political divide, structurally and personally — not to huddle on one side of a growing chasm. Science can't afford to become a one-party system. We have only to look to the pandemic to see the cost of the political gulf growing wider.

*Nature* **595**, 650 (2021)

doi: <https://doi.org/10.1038/d41586-021-02046-z>

## Competing Interests

I.K. is Trustee, Nesta, and Main Panel A member, REF 2021

## Related Articles

- [See more letters to the editor](#)

## Subjects

- [Research management](#)
- [Policy](#)
- [Society](#)

**nature** careers

[Jobs](#) >

- [\*\*Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics\*\*](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [\*\*FACULTY POSITION IN MICROBIOME BIOCHEMISTRY\*\*](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [\*\*Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology\*\*](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

- [\*\*See more letters to the editor\*\*](#)

## Subjects

- [Research management](#)

- [Policy](#)
  - [Society](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02046-z>

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

- CORRESPONDENCE
- 27 July 2021

# Global climate models do not need more behavioural science

- [M. Granger Morgan](#)<sup>0</sup> &
- [Hadi Dowlatabadi](#)<sup>1</sup>

1. [M. Granger Morgan](#)

1. Carnegie Mellon University, Pittsburgh, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Hadi Dowlatabadi

1. University of British Columbia, Vancouver, Canada.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

## [Download PDF](#)

We disagree that it is useful to include more behavioural and political science in global integrated assessment models (IAMs; see [W. Peng \*et al.\*](#) [Nature 594, 174–176; 2021](#)). Many of the features the authors suggest have already been reflected in several IAMs (see, for instance, [M. G. Morgan and H. Dowlatabadi](#) [Climat. Change 34, 337–368; 1996](#)). But high degrees of uncertainty mean that these cannot provide useful quantitative conclusions about policy at the local level.

For instance, in 2000, we used autonomous agents with pro- and anti-climate policies in 13 global regions to study policy formation and dissolution under public pressure from extreme events and high taxes. The key insights were that uncertainties and regional diversity in decision-making and values are too broad and varied for global models to inform specifics in policies.

Regional adherence to global targets requires nuanced strategies at the sub-national scale, reflecting local politics and technical, economic and behavioural aspects of specific problems — just as we need downscaling for detailed climate projections. Consequently, we have spent the past two decades focusing on analysis for specific sector and transition challenges. We urge others to do the same.

*Nature* **595**, 650 (2021)

*doi:* <https://doi.org/10.1038/d41586-021-02055-y>

## Competing Interests

The authors declare no competing interests.

## Related Articles

- [See more letters to the editor](#)

## Subjects

- [Climate sciences](#)
- [Human behaviour](#)
- [Policy](#)
- [Politics](#)

**nature** careers

[Jobs](#) >

- [\*\*Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics\*\*](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [\*\*FACULTY POSITION IN MICROBIOME BIOCHEMISTRY\*\*](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [\*\*Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology\*\*](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

- [\*\*See more letters to the editor\*\*](#)

## Subjects

- [Climate sciences](#)

- [Human behaviour](#)
  - [Policy](#)
  - [Politics](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02055-y>.

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

- CORRESPONDENCE
- 27 July 2021

# Peru to punish bending of clinical-trial rules

- [Cesar Cabezas-Sanchez](#) ORCID: <http://orcid.org/0000-0001-5120-0713><sup>0</sup>,
- [Yamilee Hurtado-Roca](#) ORCID: <http://orcid.org/0000-0003-1993-6223><sup>1</sup> &
- [Victor Suárez-Moreno](#) ORCID: <http://orcid.org/0000-0002-7078-4731><sup>2</sup>

1. Cesar Cabezas-Sánchez

1. National Institute of Health (INS), Lima, Peru.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Yamilee Hurtado-Roca

1. National Institute of Health (INS), Lima, Peru.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. [Victor Suárez-Moreno](#)

1. National Institute of Health (INS), Lima, Peru.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

-



•

[Download PDF](#)

Peru's National Institute of Health (INS) is urgently reviewing its procedures for authorization and supervision of clinical trials for COVID-19 treatments and vaccines, after protocol deviations in one such trial (see [Nature 592, 174–175 \(2021\)](#) and [C. F. Cáceres et al. Nature 592, 685; 2021](#)). Now under investigation by a regulatory body, the incident dented public confidence in the integrity of our scientific community.

The INS aims to ensure that such irregularities can never happen again. It plans to create a National Bioethics Commission, to establish a sanctions

procedure and to upgrade training in ethics and scientific integrity for researchers.

There are currently 28 other clinical trials related to COVID-19 (4 vaccines and 24 treatments) in progress in Peru. All conform to international ethical and scientific standards.

*Nature* **595**, 650 (2021)

*doi:* <https://doi.org/10.1038/d41586-021-02056-x>

## Competing Interests

The authors declare no competing interests.

## Related Articles

- [See more letters to the editor](#)

## Subjects

- [SARS-CoV-2](#)
- [Vaccines](#)
- [Public health](#)
- [Ethics](#)
- [Government](#)

**nature** careers

## [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- **FACULTY POSITION IN MICROBIOME BIOCHEMISTRY**

Duke University

Durham, NC, United States

- **Bioinformatician (m/f/div)**

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- **Student Research Assistant / HIWI (m/f/x) for Literature Research in the Field of Soil Microbiology**

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

- **See more letters to the editor**

## Subjects

- [SARS-CoV-2](#)
  - [Vaccines](#)
  - [Public health](#)
  - [Ethics](#)
  - [Government](#)
-

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02056-x>

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

- CORRESPONDENCE
- 27 July 2021

# Europe's reformed agricultural policy disappoints

- [Jeroen J. L. Candel](#)<sup>0</sup>,
- [Sebastian Lakner](#)<sup>1</sup> &
- [Guy Pe'er](#)<sup>2</sup>

1. [Jeroen J. L. Candel](#)

1. Wageningen University, the Netherlands.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Sebastian Lakner

1. University of Rostock, Germany.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Guy Pe'er

1. German Centre for Integrative Biodiversity Research, Leipzig, Germany.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

## [Download PDF](#)

Last month, the European Union’s parliament, council and commission reached a compromise on reforms to the bloc’s Common Agricultural Policy. Commission vice-president Frans Timmermans hailed the deal as “a real shift in how we practice agriculture in Europe”. In our view, this €270-billion (US\$318-billion) deal is unlikely to achieve the desired improvements in climate impact and ecosystems over the period 2023–27.

Almost all the conditions on which subsidies depend have been watered down. In particular, the requirements to rotate crops and protect carbon-rich

soils and landscape features are particularly unambitious. The Eco-schemes instrument, intended to encourage farmers to provide more environmental services, leaves too much up to member states to decide, and there is little scope for the commission to intervene if they fall behind. Disappointingly, most of the budget will be spent on ‘business as usual’ instead of on future-proofing.

The size of the EU single market means that this failure to properly reform the policy will have global implications. It is now up to member states to ensure that they support the commission in realizing its Green Deal ambitions.

*Nature* **595**, 650 (2021)

doi: <https://doi.org/10.1038/d41586-021-02047-y>

## Competing Interests

J.J.L.C. is a member of the Dutch Council on Animal Affairs and a member of the supervisory board of the Transitiocoalitie Voedsel ('Transition coalition Food') foundation.

## Related Articles

- [See more letters to the editor](#)

## Subjects

- [Agriculture](#)
- [Environmental sciences](#)
- [Policy](#)

**nature** careers

[Jobs](#) >

- [\*\*Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics\*\*](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [\*\*FACULTY POSITION IN MICROBIOME BIOCHEMISTRY\*\*](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [\*\*Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology\*\*](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

- [\*\*See more letters to the editor\*\*](#)

## Subjects

- [Agriculture](#)

- [Environmental sciences](#)
  - [Policy](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02047-y>.

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

# Work

- **Why science needs a new reward and recognition system**

[ 15 July 2021]

Career Column • Researchers with children or carer roles have struggled more than others during the pandemic, amplifying existing inequalities, argue Edyta Swider-Cios and colleagues.

- **Chemical modelling with a sense of touch** [ 26 July 2021]

Where I Work • Computational chemist Mona Minkara simulates the molecular dynamics required for lung function and boosts access for visually impaired chemists.

- CAREER COLUMN
- 15 July 2021
- Correction [19 July 2021](#)

# Why science needs a new reward and recognition system

Researchers with children or carer roles have struggled more than others during the pandemic, amplifying existing inequalities, argue Edyta Swider-Cios and colleagues.

- [Edyta Swider-Cios](#)<sup>0</sup>,
- [Katalin Solymosi](#)<sup>1</sup> &
- [Mangala Srinivas](#)<sup>2</sup>

1. Edyta Swider-Cios

1. Edyta Swider-Cios is a researcher in the Department of Cognitive Neuropsychology, Tilburg University, the Netherlands, and is a former project manager at the Young Academy of Europe.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Katalin Solymosi

1. Katalin Solymosi is an assistant professor in the Department of Plant Anatomy, ELTE Eötvös Loránd University, Budapest, Hungary, a board member of the Young Academy of Europe and co-chair of the Hungarian Young Academy.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

### 3. Mangala Srinivas

1. Mangala Srinivas is a professor of cell biology and immunology at Wageningen University and Research, the Netherlands, chief science officer at Cenza Imaging in Amsterdam and outgoing chair of the Young Academy of Europe.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•

[Download PDF](#)



Agency: Getty

The chance of 'success' in an academic science career, which has historically been defined as obtaining a permanent position, is low<sup>1,2</sup>. Because of this, early-career researchers, including PhD students, postdocs and junior faculty members, are vulnerable to extra stresses<sup>3–5</sup>.

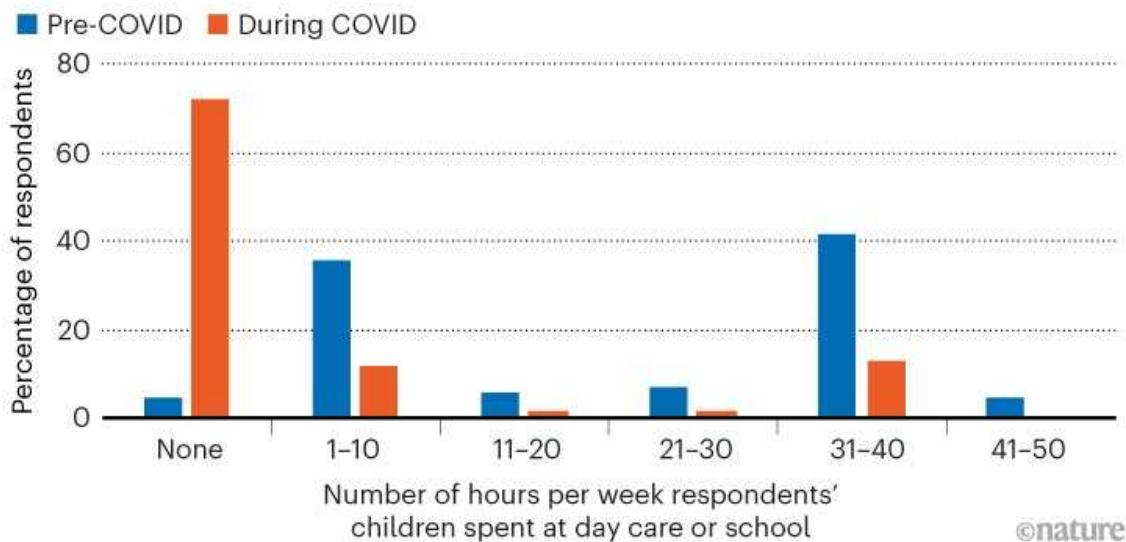
From June to October 2020, we surveyed 151 such researchers in different fields and from multiple countries in Europe to better understand how they were affected by pandemic-related lockdowns and associated relief efforts — such as extensions on grant or scholarship reporting and eligibility, and budget-neutral project extensions.

Not surprisingly, we found that the pandemic has amplified existing discrepancies among these researchers, especially between those with and without carer responsibilities. Some funders and academic institutions have provided deadline extensions or extended contracts. But these efforts might

offer advantages only to certain groups, owing to their eligibility conditions and how well publicized they are. It is alarming that these measures might increase existing inequalities in academia as the pandemic continues into a second year and a return to normality remains unpredictable.

## LOST SAFETY NET

The pandemic closed the doors of many day-care centres and schools, an important resource for almost all respondents with children in pre-pandemic times. The percentage of children who spent no time at a day-care provider or in a school building during the pandemic jumped from 5% to 70%.



©nature

Of our survey respondents, 68% were women and 31% were men (1% had a different gender identity or did not disclose). Roughly half the respondents are members of a Young Academy, such as a National Young Academy or the Young Academy of Europe, which suggests that they are successful in their field and have a leadership role. Given the relatively small number of participants and the complexity of this survey, we can draw only qualitative conclusions. Respondents who have been negatively affected were more likely to respond than were others, and one-third of our respondents are from Hungary, which could bias our results.

Our survey, which was financially supported by [the CALIPER project](#), produced three key findings (the full survey data are available [here](#)). Free-text responses further illustrated the conclusions we've drawn (see 'Free-text responses').

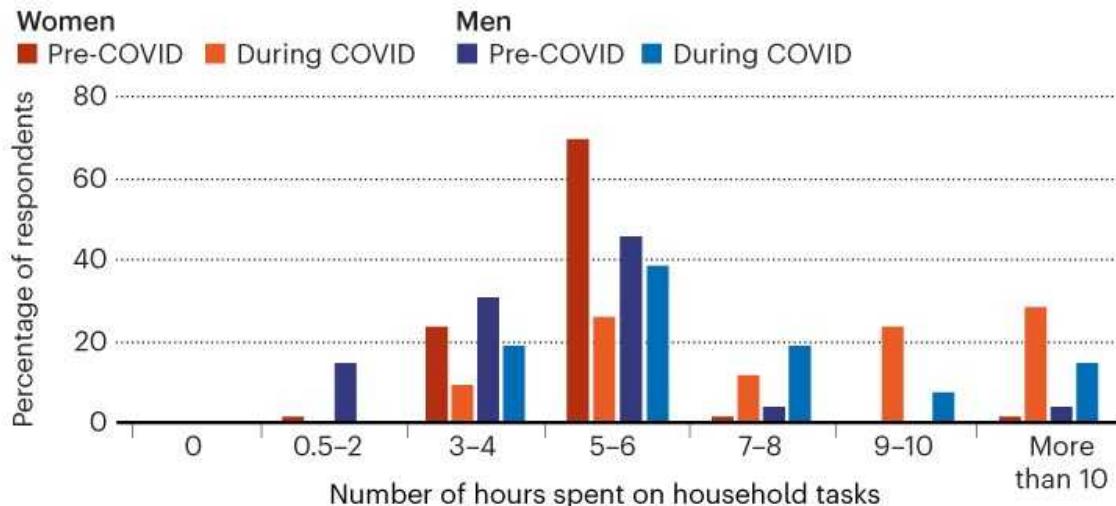
## **1. Reconciling work obligations with childcare responsibilities has been particularly challenging**

We found that researchers with children under the age of 10 struggled the most to balance home-schooling and professional work. Our participants did not have other significant carer responsibilities beyond those for their own children, although we presume that those who provide care for an older person or dependent family member would experience similar difficulties. The absence of in-person school or day care (see ‘Lost safety net’) resulted in these respondents spending up to 8 extra hours daily on childcare and household activities (see ‘Growing responsibilities’).

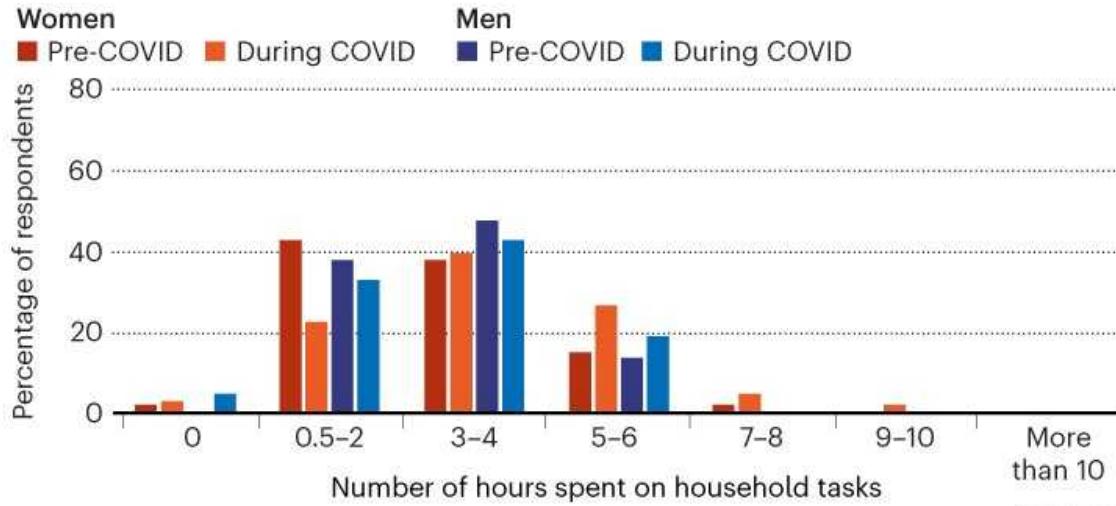
## GROWING RESPONSIBILITIES

Women with children were already devoting many hours to household tasks before the pandemic. During the pandemic, they took on an even greater share of the responsibilities. A majority reported spending at least 9 hours per day on domestic duties.

### With children



### Without children



©nature

## 2. Gender bias persists

Even before the pandemic, the average time spent on household activities differed between researchers with and without kids (see ‘Growing responsibilities’). Most scientist-parents spent 5–6 hours per day on household activities, whereas researchers without kids spent less than 3–4

hours. The pandemic further exacerbated these differences, especially for women with children. Our results indicate that women experienced more stress than men did during the lockdown (see ‘Stress factor’). There were even differences between genders in the amount of hours spent asleep (see ‘Work–life imbalance’). Other articles have also outlined how the pandemic has hit academic scientist–mums harder than their male or non-parent counterparts<sup>[6,7](#)</sup>.

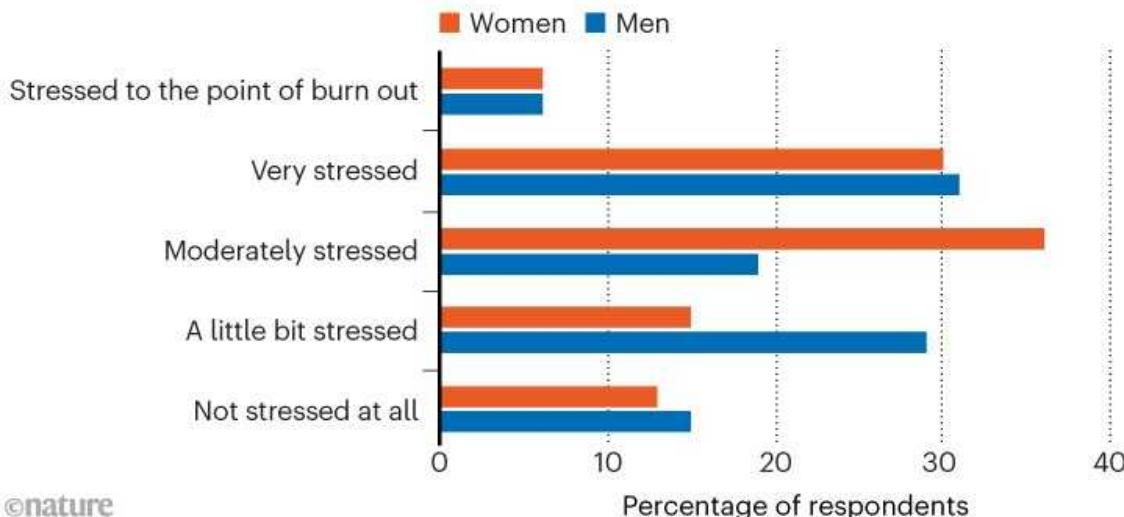
### **3. Some people benefited**

Those who did not suffer from restricted laboratory access or non-ideal working conditions at home might have experienced an advantage from working remotely. Many reported having fewer meetings, administrative tasks or disturbances, and noted that they were able to quietly focus on their work — an accomplishment that is almost impossible for scientist–parents with young children at home, especially before the kids’ bedtime.

However, those with no carer obligations often had to take on work responsibilities — particularly teaching and administrative work, but also experimental work, instrument maintenance and activities that required an in-person presence at the workplace, such as looking after cell, bacterial or fungal cultures, experimental animals or plants — from others who did have such obligations. These scientists might also have experienced a decrease in actual research time, as well as increased stress levels because of their higher workloads and feelings of isolation (and possibly increased exposure risk to SARS-CoV-2 when they had to go into their workplace).

## STRESS FACTOR

Women and men were equally likely to report feeling very stressed or burnt out. They were also equally likely to report not being stressed at all. But differences become apparent between those two extremes: women were more likely to report feeling 'moderately' stressed, whereas men were more likely to report being 'a little bit' stressed.



©nature

## Interventions

Funding agencies' mitigation measures have varied. Some have given blanket extensions to everyone by, for example, extending eligibility windows for a year (for example, a grant that allowed researchers to apply within three years of receiving their doctoral degree would be extended to four years), or extending existing projects, often at no extra cost. This is a strong start to helping researchers. Yet simple extensions for all are not ideal, because they could penalize those who cannot work from home effectively, and could give a boost to those who can do so. This disparity could worsen existing inequalities in the research enterprise and could prompt scientists from disadvantaged groups to reconsider their career options and think about leaving academia.



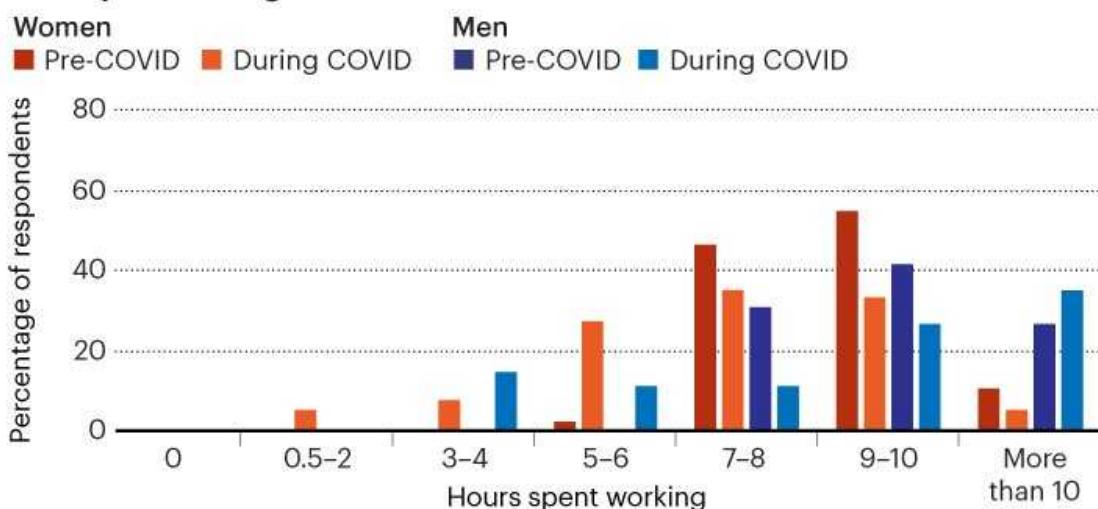
## [Collection: Science careers and mental health](#)

On the bright side, most academic institutions have normalized working from home. This allows for far greater flexibility, which in the long run is an advantage for scientists with carer responsibilities, and in general can improve work–life balance for everyone. Similarly, although online webinars and conferences have their own limitations, especially in terms of informal in-person networking, they are more accessible than are their on-site counterparts to scientists who have fewer financial resources or less mobility and, as such, can significantly increase the diversity of attendees. We hope that issues such as the unequal carer responsibilities, evaluation criteria that make it more difficult for women and minority groups to advance, and a lack of diversity and gender balance in scientific panels and among invited speakers and those in senior roles in academia will continue to receive attention<sup>8</sup> after the pandemic ends.

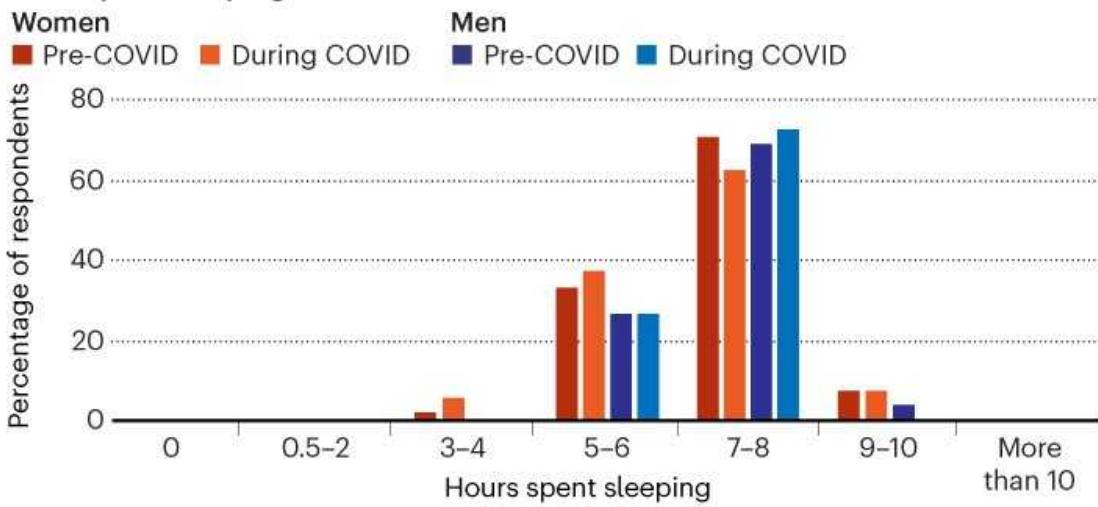
## WORK-LIFE IMBALANCE

The pandemic affected mothers and fathers differently. The proportion of women getting at least 7 hours of sleep each day dropped, as did the proportion who worked more than 10 hours per day. Slightly more men, conversely, got at least 7 hours of sleep during the pandemic.

### Time spent working



### Time spent sleeping



©nature

The consequences of the pandemic, and the difficulties that it has caused for many scientists, call for accelerating the development of new ways to recognize and reward academic researchers. To mitigate the problems we identify here, [which have also been brought up by others](#), we suggest that scientists should be appraised on their efforts and progress in light of their personal circumstances, and funders should move towards evaluating

narrative CVs, instead of using assessments focused only on impact factors and other quantitative performance indicators.

We already know that academia treats those in the sector unequally, penalizing carers, women and those from minority ethnic groups<sup>9–12</sup>. In our view, intervention is necessary, and existing solutions are welcome, but are insufficient on their own. The degree of support that a researcher receives — when they receive it at all — needs to be more nuanced than simple blanket extensions to everyone. It is necessary to gauge the impact that the lockdown has had on an individual and their work before awarding an extension. This would require a careful and personalized career-evaluation procedure.

We recognize that this creates more work for reviewers, evaluation committees and others, but we do not want the pandemic's effects to fortify the proverbial glass ceiling.

## Free-text responses

Free-text responses illustrate the struggles some respondents face in terms of childcare during the pandemic.

- “A number of leadership roles have been offered and awarded in my department during the lockdown, which I and others with similar (short-term) childcare difficulties were simply not in a position to apply for.”
- “I feel terribly disadvantaged. My children are 3 and 5, and require permanent attention. My husband is a key worker, I get a few hours a day [for work], essentially for teaching, and no progress on my research.”
- “The combination of home-schooling and a kindergarten-age child at home was extremely burdensome. No full-time work possible during that time.”

Free-text responses also highlighted respondents' experience of gender bias.

- “My male colleagues all seem to manage to pursue their research brilliantly, and that creates a terribly unfair bias for the next funding application or the professor promotion.”

- “Even in households with two parents of different genders, it’s quite remarkable how much more childcare fell to the woman — in my household, but many colleagues’ households as well.”

How scientist-parents’ pandemic experiences differed from those of their colleagues without children is also underscored in the responses.

- “I got a lot of e-mails from single/non-parent colleagues saying how nice the pandemic is and that we can finally concentrate on research (which was not at all my case, due to my son).”
- “Most of the childless researchers were happy to do ‘home office’; it was more peaceful without transport time to work. Doing home office, my research, my lectures online, with two children, wasn’t easy.”
- “I think the only researchers who have benefited from COVID-19 are those who don’t teach, don’t have children, and don’t have service roles. The workload for the rest of us has been enormous.”
- “Time and stress related to public transport disappeared.”
- “I have been more capable of focussing on the writing of my article due to fewer other activities and the lack of commuting.”

*Nature* **595**, 751-753 (2021)

doi: <https://doi.org/10.1038/d41586-021-01952-6>

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.](#)

## Updates & Corrections

- **Correction 19 July 2021:** An earlier version of this story erroneously stated that Mangala Srinivas is chief security officer at Cenza Imaging.

## References

1. 1.

Smaglik, P. *Nature* **562**, S49–S51 (2018).

[PubMed](#) [Article](#) [Google Scholar](#)

2. 2.

Castellacci, F. & Viñas-Bardolet, C. *Stud. Higher Educ.* <https://doi.org/10.1080/03075079.2019.1711041> (2020).

[Article](#) [Google Scholar](#)

3. 3.

Husby, A. & Modinos, G. *Nature* **580**, 185 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

4. 4.

Yan, W. *Science* <https://doi.org/10.1126/science.caredit.abc1291> (2020).

[Article](#) [Google Scholar](#)

5. 5.

Gewin, V. *Nature* **591**, 489–491 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

6. 6.

Langin, K. *Science* **371**, 660 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

7. 7.

UN Women. *Whose Time to Care? Unpaid Care and Domestic Work During COVID-19*. (UN Women, 2020).

8. 8.

Gewin, V. *Nature* **583**, 867–869 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

9. 9.

Larivière, V., Ni, C., Gingras, Y., Cronin, B. & Sugimoto, C. R. *Nature* **504**, 211–213 (2013).

[PubMed](#) [Article](#) [Google Scholar](#)

10. 10.

Wingfield, A. H. *Science* **369**, 351 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

11. 11.

Dworkin, J. D. *et al.* *Nature Neurosci.* **23**, 918–926 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

12. 12.

Holman, L., Stuart-Fox, D. & Hauser, C. E. *PLoS Biol.* **16**, e2004956 (2018).

[PubMed](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Competing Interests

The authors declare no competing interests.

## Related Articles

-  [Webcast: Parenting in the pandemic](#)
-  [How I managed my work and personal life as a sole parent during the pandemic](#)
-  [Coronavirus lockdown lessons from single-parent scientists](#)

## Subjects

- [Policy](#)
- [Funding](#)
- [Careers](#)

**nature** careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- **FACULTY POSITION IN MICROBIOME BIOCHEMISTRY**

Duke University

Durham, NC, United States

- **Bioinformatician (m/f/div)**

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- **Student Research Assistant / HIWI (m/f/x) for Literature Research in the Field of Soil Microbiology**

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles

-  [Webcast: Parenting in the pandemic](#)
-  [How I managed my work and personal life as a sole parent during the pandemic](#)
-  [Coronavirus lockdown lessons from single-parent scientists](#)

# Subjects

- [Policy](#)
  - [Funding](#)
  - [Careers](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01952-6>

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

- WHERE I WORK
- 26 July 2021

# Chemical modelling with a sense of touch

Computational chemist Mona Minkara simulates the molecular dynamics required for lung function and boosts access for visually impaired chemists.

- [Kendall Powell](#) <sup>0</sup>

1. Kendall Powell

1. Kendall Powell is a freelance writer in Boulder, Colorado.

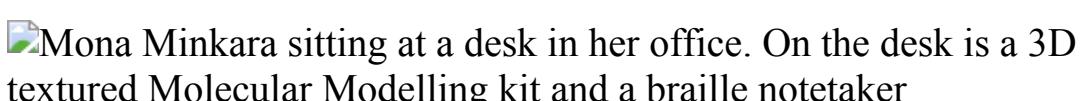
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





•



Mona Minkara is a computational chemist at Northeastern University in Boston, Massachusetts. Credit: Kayana Szymczak for *Nature*

[Download PDF](#)

My group wants to understand the components of lung surfactants — complex substances that keep our air sacs from collapsing. Bioengineers would like to make synthetic surfactants to treat lung disease, but we must understand the real ones first.

As a computational chemist, I use computer simulations to understand the molecular dynamics of how the lipids and proteins in the surfactants react to one another at the air–liquid interface.

My group studies the four proteins in pulmonary surfactant. Two of these flag down any foreign entity that enters the lungs, and we’re interested in their interactions with respiratory viruses. The proteins instruct immune cells to engulf and kill invaders. But how do they identify intruders? It’s like being a detective on the atomic scale to work out which atoms of the molecules are attracted to which others.

Does being blind help me to visualize molecular interactions in different ways from sighted chemists? That’s the million-dollar question. I think so. There is an often-unseen advantage to the scientific community including as many differently abled people from diverse backgrounds as possible. I want my lab group to be like that, so we can solve problems creatively.

I’ve collected various accessibility tools on my career path, and I believe in sharing them to catalyse knowledge. I helped to design the 3D-printed chemistry modelling kit shown here. Although anyone can work with the model, it is textured as well as colour-coded, so that a blind person can use it through the Braille-labelled key. The red oxygen module has tiny bumps, the purple boron module has wavy lines, and so on.

Academic science is not inclusive because we do not understand our differences. Not understanding is OK, but not acting to improve understanding is not OK. Everyone needs to see people with disabilities as capable, even if they are capable in different ways.

*Nature* **595**, 756 (2021)

doi: <https://doi.org/10.1038/d41586-021-02042-3>

## Related Articles

-  [Engineering a brighter future for refugees and female scientists in Uganda](#)
-  [My international career journey as a disabled scientist](#)
-  [How I self-advocate for researchers with disabilities and illnesses](#)

## Subjects

- [Careers](#)
- [Chemistry](#)
- [Scientific community](#)

**nature** careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [\*\*Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology\*\*](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- [\*\*Engineering a brighter future for refugees and female scientists in Uganda\*\*](#)



- [\*\*My international career journey as a disabled scientist\*\*](#)



- [How I self-advocate for researchers with disabilities and illnesses](#)

## Subjects

- [Careers](#)
- [Chemistry](#)
- [Scientific community](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02042-3>

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

# Research

- **[A protective signal between the brain's supporting cells in Alzheimer's disease](#)** [ 14 July 2021]

News & Views • In a mouse model of Alzheimer's disease, interleukin-3 protein released by cells called astrocytes activates microglia, the immune cells of the brain. These then cluster around disease-associated protein aggregates and help to clear them.

- **[Direct installation of boron groups offers boost to medicinal chemistry](#)** [ 26 July 2021]

News & Views • Compounds called borylated azines have untapped potential for organic synthesis, but have faced problems associated with their preparation, stability and reactivity. A new class of these compounds provides a solution.

- **[Sperm CatSper ion channel swims into sharper focus](#)** [ 21 July 2021]

News & Views • Activation of the CatSper membrane ion channel triggers enhanced sperm motility, and this protein complex is required for male fertility. Structural data now provide insights into the channel, and offer a few surprises, too.

- **[Earthquakes triggered by underground fluid injection modelled for a tectonically active oil field](#)** [ 28 July 2021]

News & Views • An analysis of the Val d'Agri oil field in Italy provides insight into how processes associated with wastewater disposal trigger earthquakes — and how such effects can be reduced to maintain the economic viability of mature oil fields.

- **[Light bending and X-ray echoes from behind a supermassive black hole](#)** [ 28 July 2021]

Article • One of the key predictions of general relativity, the bending of light around massive, compact objects, is observed for a supermassive black hole in the galaxy I Zwicky 1.

- **[Incoherent transport across the strange-metal regime of overdoped cuprates](#)** [ 28 July 2021]

Article • Measurements of high-field magnetotransport in overdoped cuprates indicate that the strange-metal regime exists beyond the critical doping, and that it has both coherent and incoherent contributions.

- **[Linear-in temperature resistivity from an isotropic Planckian scattering rate](#)** [ 28 July 2021]

Article • Angle-dependent magnetoresistance measurements of a strange-metal phase of a hole-doped cuprate show a well defined Fermi surface and an isotropic linear-in-temperature scattering rate that saturates at the Planckian limit.

- **Spectroscopic evidence for a gold-coloured metallic water solution** [ 28 July 2021]

Article • Spectroscopic measurements confirm that when water is adsorbed on drops of an alkali alloy at low pressure a gold-coloured metallic layer forms as electrons rapidly move from the drop into the water.

- **A radical approach for the selective C–H borylation of azines** [ 20 May 2021]

Article • Selective borylation of azines—nitrogen-containing aromatic heterocycles used in the synthesis of many pharmaceuticals—is made possible by forming a radical from an aminoborane using a photocatalyst.

- **A process-based approach to understanding and managing triggered seismicity** [ 28 July 2021]

Article • A multidisciplinary method for managing triggered seismicity is developed using detailed subsurface information to calibrate geomechanical and earthquake source physics models, and is applied to the Val d’Agri oil field in seismically active southern Italy.

- **Dynamical prefrontal population coding during defensive behaviours** [ 14 July 2021]

Article • Dorsomedial prefrontal cortex represents threat and defensive action initiation, and the dynamic interplay between those representations dictates the temporal evolution of avoidance response to threats.

- **Food cue regulation of AGRP hunger neurons guides learning** [ 14 July 2021]

Article • In response to food cues, a hypothalamic circuit in the mouse brain transiently inhibits neurons expressing agouti-related peptide, and this promotes learning of cue-initiated food-seeking tasks.

- **Astrocytic interleukin-3 programs microglia and limits Alzheimer’s disease** [ 14 July 2021]

Article • Interleukin-3 signalling from astrocytes to microglia readies microglia to defend against Alzheimer’s disease.

- **Spread of a SARS-CoV-2 variant through Europe in the summer of 2020** [ 07 June 2021]

Article • Analysis of the spread of the 20E (EU1) variant of SARS-CoV-2 through Europe suggests that international travel and insufficient containment, rather than increased transmissibility, led to a resurgence of infections.

- **Untangling introductions and persistence in COVID-19 resurgence in Europe** [ 30 June 2021]

Article • In many European countries, more than half of the SARS-CoV-2 lineages circulating in late summer 2020 resulted from new introductions, highlighting the threat of viral dissemination when restrictions are lifted.

- **Nasal delivery of an IgM offers broad protection from SARS-CoV-2 variants** [ 03 June 2021]

Article • An engineered IgM antibody administered intranasally in mice shows high prophylactic efficacy and therapeutic efficacy against SARS-CoV-2, and is also effective against multiple variants of concern that are resistant to IgG-based therapeutics.

- **Metabolic control of TFH cells and humoral immunity by phosphatidylethanolamine** [ 07 July 2021]

Article • Enzymes in the cytidine diphosphate–ethanolamine metabolic pathway, which promotes de novo synthesis of phosphatidylethanolamine, are shown to act as post-transcriptional mediators of the differentiation of T follicular helper (TFH) cells, by regulating the chemokine receptor CXCR5.

- **EGFR activation limits the response of liver cancer to lenvatinib** [ 21 July 2021]

Article • EGFR inhibition and lenvatinib treatment of liver cancer cells in vitro and in vivo mouse models has potent anti-proliferative effects, and lenvatinib plus gefitinib treatment of 12 patients with advanced liver cancer resulted in meaningful clinical responses.

- **Enhancer release and retargeting activates disease-susceptibility genes** [ 26 May 2021]

Article • Disruption of a promoter can release its partner enhancer to activate other promoters in the same contact domain, and this process, named ‘enhancer release and retargeting’, can often lead to gene alterations that cause disease.

- **Structural basis of early translocation events on the ribosome** [ 07 July 2021]

Article • Cryo-electron microscopy and single-molecule fluorescence methods are used to elucidate the mechanism of early translocation events on the bacterial ribosome.

- **Structure of a mammalian sperm cation channel complex** [ 05 July 2021]

Article • A structure of the sperm-specific CatSper complex features a number of additional components; together, these components and the CatSper complex are termed the CatSpermosome.

- NEWS AND VIEWS
- 14 July 2021

# A protective signal between the brain's supporting cells in Alzheimer's disease

In a mouse model of Alzheimer's disease, interleukin-3 protein released by cells called astrocytes activates microglia, the immune cells of the brain. These then cluster around disease-associated protein aggregates and help to clear them.

- [Jerika J. Barron](#)<sup>0</sup> &
- [Anna V. Molofsky](#)<sup>1</sup>

1. Jerika J. Barron

1. Jerika J. Barron is in the Department of Psychiatry and Behavioral Sciences and the Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, California 94158, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. [Anna V. Molofsky](#)

1. Anna V. Molofsky is in the Department of Psychiatry and Behavioral Sciences and the Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, California 94158, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

-



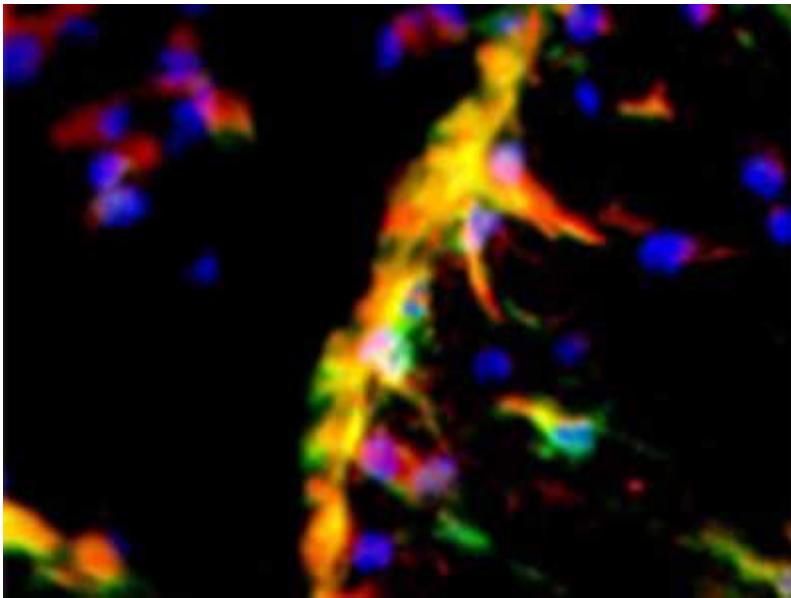
•

## [Download PDF](#)

Immune signalling associated with inflammation is sometimes considered a thing to be avoided. However, immune signalling can also be protective, quelling damage and disease-causing microorganisms (pathogens). In the brain, multiple cell types work together to maintain brain health, mediate inflammatory responses and optimize the function of the main output cells, neurons. [Writing in Nature](#), McAlpine *et al.*<sup>1</sup> uncover a signalling axis between two of these brain cell types, astrocytes and microglia. They demonstrate that this signalling, mediated by the immune protein

interleukin-3 (IL-3), limits disease progression and brain dysfunction in a model of Alzheimer's disease.

Alzheimer's disease (AD) is a devastating and prevalent neurodegenerative disorder that leads to the loss of brain cells and of the synaptic connections between neuronal cells, resulting in progressive cognitive decline. One hallmark of AD is the presence of disease-associated aggregates of different proteins in the brain: that is, plaques made up of the protein amyloid- $\beta$  (A $\beta$ ), and neuro-fibrillary tangles made up of the protein tau<sup>2</sup>.



[Read the paper: Astrocytic interleukin-3 programs microglia and limits Alzheimer's disease](#)

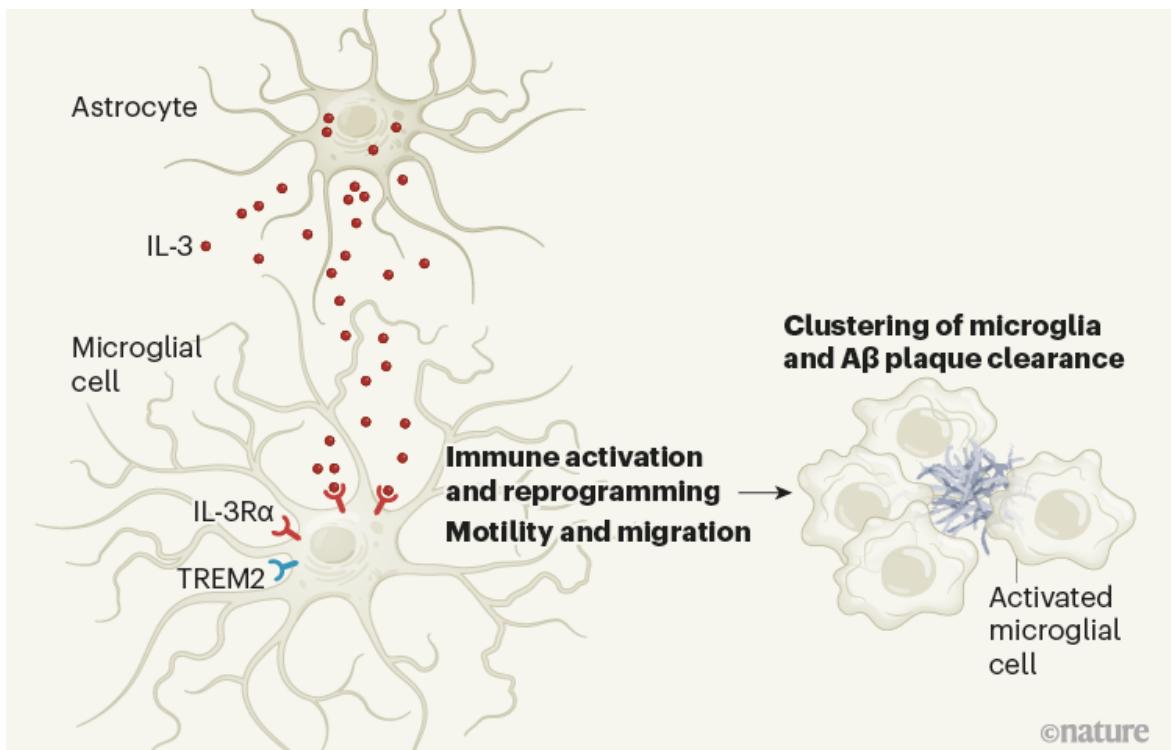
Normally, astrocytes and microglia help to maintain neuronal health and function by clearing debris, recycling neurotransmitter molecules and supporting communication across synapses<sup>3</sup>. In the brain of a person with AD, astrocytes and microglia become activated, produce inflammatory molecules and aggregate around the protein plaques. This microglial aggregation can be protective, by preventing loose (soluble) A $\beta$  protein from diffusing throughout the brain<sup>4</sup>. However, the signals that coordinate the functions of microglia are not fully understood. McAlpine *et al.* show that the astrocyte-produced cytokine IL-3 is one of these signals, and that it plays a central part in a model of AD.

Cytokines are soluble signalling proteins that are a key form of immune communication. They are involved in complex signalling loops that can initiate, intensify or resolve inflammation, recruit immune cells to where they are needed, and aid the clearance of pathogens and cellular debris. Their impact on the brain has long been of interest in neurological diseases in which inflammation is observed alongside impairments in neural function.

The roles of IL-3 in the brain have not been particularly well studied. However, it is known to regulate inflammation in multiple ways, such as by driving the proliferation of immune cells, including those that circulate in the blood<sup>5</sup>, and it has been associated with AD risk in studies of patient plasma<sup>6,7</sup>.

McAlpine *et al.* examined a model of AD in which mice carry five mutations that have been implicated in the disorder in humans. These AD mice develop A $\beta$  plaques and show progressive impairments in short-term memory with age<sup>8</sup>. When these AD-model mice also lacked IL-3, they showed an increased accumulation of A $\beta$  plaques, higher levels of soluble A $\beta$  and greater impairments in short-term and spatial memory.

To pinpoint the cellular sources of IL-3 in the brain, McAlpine *et al.* generated mice in which IL-3-producing cells are fluorescently labelled; this approach identified a subset of astrocytes as a major source of IL-3. AD mice in which IL-3 was deleted specifically from astrocytes showed increased accumulation of A $\beta$  plaques and more severely impaired short-term memory compared with AD mice. The authors observed a similar effect in AD mice that completely lacked IL-3, suggesting that astrocytes are the key cellular reservoir of the protein in this AD model (Fig. 1). Notably, the aggregation of microglia near A $\beta$  plaques was reduced in AD mice lacking astrocytic IL-3.



**Figure 1 | Interleukin-3 (IL-3) protein signalling in the brain can help to clear disease-associated protein aggregates.** McAlpine *et al.*<sup>1</sup> studied a mouse model of Alzheimer's disease (AD) in which disease-associated plaques consisting of amyloid- $\beta$  (A $\beta$ ) protein form in the brain. The authors revealed a signalling axis between two types of brain cell, astrocytes and microglia, that had not previously been implicated in AD. Astrocytes express and release IL-3, which activates IL-3R $\alpha$  on the surface of a subset of microglia that also express the cell-surface receptor TREM2. The IL-3 signal activates and reprograms the microglia, promoting their movement and clustering around A $\beta$  plaques. Microglia have a role in clearing A $\beta$  plaques<sup>4</sup> and, consistent with this, McAlpine *et al.* found that treating AD-model mice with brain injections of IL-3 resulted in reduced build-up of A $\beta$  plaques in these animals. (Figure adapted from Extended Data Fig. 9e in ref. 1.)

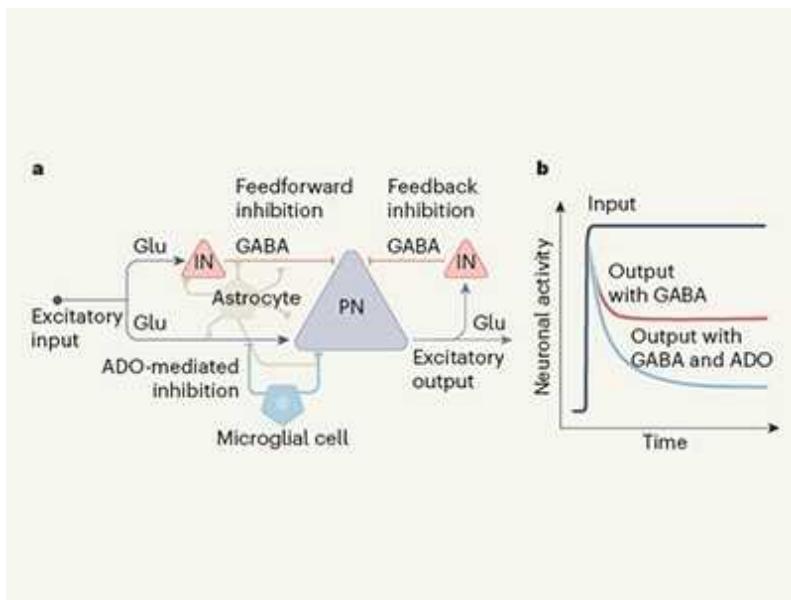
The authors next identified the targets of IL-3 in the brain. They found that microglia express the IL-3 receptor IL-3R $\alpha$ , and that levels of this receptor are substantially increased with age and in the AD model compared with young, wild-type mice. Deleting IL-3R $\alpha$  specifically from microglia in the

AD mice resulted in effects on A $\beta$  plaque burden and memory similar to those observed in AD mice lacking IL-3 in astrocytes.

Strikingly, McAlpine and colleagues found that injecting IL-3 into the brains of AD mice could reduce A $\beta$  build-up and stimulate the clustering of microglia around A $\beta$  plaques (Fig. 1). Continuous delivery of IL-3 into the brains of IL-3-deficient AD mice over four weeks resulted in a remarkable reduction in the size and amount of plaques and the levels of soluble A $\beta$ , as well as improvements in short-term memory relative to results seen in AD mice injected with an inactive control substance. This is a key finding with potential therapeutic implications.

Interest in the role of microglia in AD has increased dramatically since the discovery that a variant of the gene encoding the receptor protein TREM2, which is expressed by microglia, is associated with risk for AD<sup>9</sup>. McAlpine *et al.* show that *Il3ra* is enriched in a previously described subset of ‘disease-associated microglia’ that are activated through the TREM2 receptor<sup>10</sup>. Moreover, the authors found that deletion of *Trem2* prevented the increase in microglial expression of IL-3R $\alpha$  in their AD model, raising the question of whether *TREM2* mutations associated with AD risk in humans might prevent this protective IL-3-dependent response.

Indeed, the authors also found evidence that this pathway is at work in the human brain. In brain tissue from individuals who had died with AD, the authors observed astrocyte expression of IL-3 and higher microglial expression of IL-3R $\alpha$  than in the brains of age-matched individuals without AD. Moreover, the amount of microglial IL-3R $\alpha$  expression correlated with the length of time for which these individuals had been diagnosed with AD, as well as with the accumulation of A $\beta$  plaques.



### Brain's immune cells put the brakes on neurons

How does IL-3 promote the protective functions of microglia? The authors found that in AD mice lacking IL-3, microglia did not cluster around plaques, and plaque deposition was greater than in AD mice. In experiments with human microglia in culture, treating these cells with IL-3 promoted migration towards AD-associated protein aggregates.

It is important to keep in mind that IL-3 might offer protection through more than one mechanism. For example, IL-3 was particularly abundant in astrocytes at the blood–brain barrier, which controls the passage of proteins and cells from the circulatory system to the brain. Populations of immune cells called macrophages, which reside at the brain's borders, might have been targeted by some of the genetic tools used to manipulate IL-3R $\alpha$  expression. Such cells could also be affected by IL-3 to alter the entry of molecules or cells into the brain. Although McAlpine *et al.* examined some aspects of the integrity of the blood–brain barrier and found it to be intact, other unmeasured variables such as active transport of blood-borne molecules to the brain could be affected<sup>11</sup>.

In addition, it is possible that other brain cell types could express the IL-3 receptor in some contexts and respond to IL-3. Further dissecting the impact of IL-3 signalling on other cellular players will be a crucial next step.

Nonetheless, these findings are an exciting advance in understanding the role of astrocytes and microglia in AD, a disease that is notoriously difficult to treat and that currently lacks any curative or restorative therapies. Although caution should be exercised in translating these findings to the clinic, particularly given the role of IL-3 and IL-3R $\alpha$  in certain autoimmune disorders<sup>12</sup>, this study raises the intriguing idea that IL-3 or related molecules could have therapeutic potential in AD. Could this be one step towards personalized therapy for individuals with AD who carry a risk-associated *TREM2* variant? Further defining the roles of IL-3 in health and disease will be essential to fulfil the promise of McAlpine and colleagues' findings.

*Nature* **595**, 651-652 (2021)

doi: <https://doi.org/10.1038/d41586-021-01870-7>

## References

1. 1.

McAlpine, C. S. *et al.* *Nature* **595**, 701–706 (2021).

[Article](#) [Google Scholar](#)

2. 2.

Long, J. M. & Holtzman, D. M. *Cell* **179**, 312–339 (2019).

[PubMed](#) [Article](#) [Google Scholar](#)

3. 3.

Vainchtein, I. D. & Molofsky, A. V. *Trends Neurosci.* **43**, 144–154 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

4. 4.

Wang, Y. *et al.* *J. Exp. Med.* **213**, 667–675 (2016).

[PubMed](#) [Article](#) [Google Scholar](#)

5. 5.

Dougan, M., Dranoff, G. & Dougan, S. K. *Immunity* **50**, 796–811 (2019).

[PubMed](#) [Article](#) [Google Scholar](#)

6. 6.

Soares, H. D. *et al.* *Arch. Neurol.* **69**, 1310–1317 (2012).

[PubMed](#) [Article](#) [Google Scholar](#)

7. 7.

Ray, S. *et al.* *Nature Med.* **13**, 1359–1362 (2007).

[PubMed](#) [Article](#) [Google Scholar](#)

8. 8.

Oakley, H. *et al.* *J. Neurosci.* **26**, 10129–10140 (2006).

[PubMed](#) [Article](#) [Google Scholar](#)

9. 9.

Colonna, M. & Wang, Y. *Nature Rev. Neurosci.* **17**, 201–207 (2016).

[PubMed](#) [Article](#) [Google Scholar](#)

10. 10.

Keren-Shaul, H. *et al.* *Cell* **169**, 1276–1290 (2017).

[PubMed](#) [Article](#) [Google Scholar](#)

11. 11.

Yang, A. C. *et al.* *Nature* **583**, 425–430 (2020).

[PubMed](#) [Article](#) [Google Scholar](#)

12. 12.

Renner, K. *et al.* *JCI Insight* **1**, e87157 (2016).

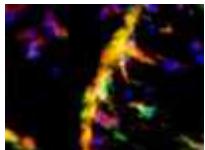
[PubMed](#) [Article](#) [Google Scholar](#)

[Download references](#)

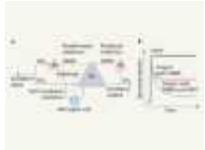
## Competing Interests

The authors declare no competing interests.

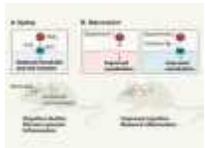
## Related Articles



- [Read the paper: Astrocytic interleukin-3 programs microglia and limits Alzheimer's disease](#)



- [Brain's immune cells put the brakes on neurons](#)



- [Reversal of immune-cell shutdown protects the ageing brain](#)

- [See all News & Views](#)

## Subjects

- [Alzheimer's disease](#)
- [Brain](#)
- [Immunology](#)
- [Neurodegeneration](#)

## nature careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

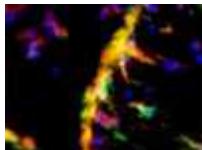
- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

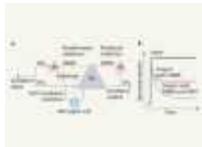
Leipzig, Germany

[Download PDF](#)

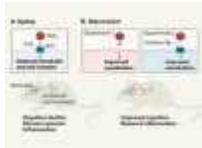
## Related Articles



- [\*\*Read the paper: Astrocytic interleukin-3 programs microglia and limits Alzheimer's disease\*\*](#)



- [\*\*Brain's immune cells put the brakes on neurons\*\*](#)



- [\*\*Reversal of immune-cell shutdown protects the ageing brain\*\*](#)
- [\*\*See all News & Views\*\*](#)

## Subjects

- [Alzheimer's disease](#)
- [Brain](#)
- [Immunology](#)
- [Neurodegeneration](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01870-7>

- NEWS AND VIEWS
- 26 July 2021

# Direct installation of boron groups offers boost to medicinal chemistry

Compounds called borylated azines have untapped potential for organic synthesis, but have faced problems associated with their preparation, stability and reactivity. A new class of these compounds provides a solution.

- [Christine M. Le](#) 0

## 1. [Christine M. Le](#)

1. Christine M. Le is in the Department of Chemistry, York University, Toronto, Ontario M3J 1P3, Canada.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



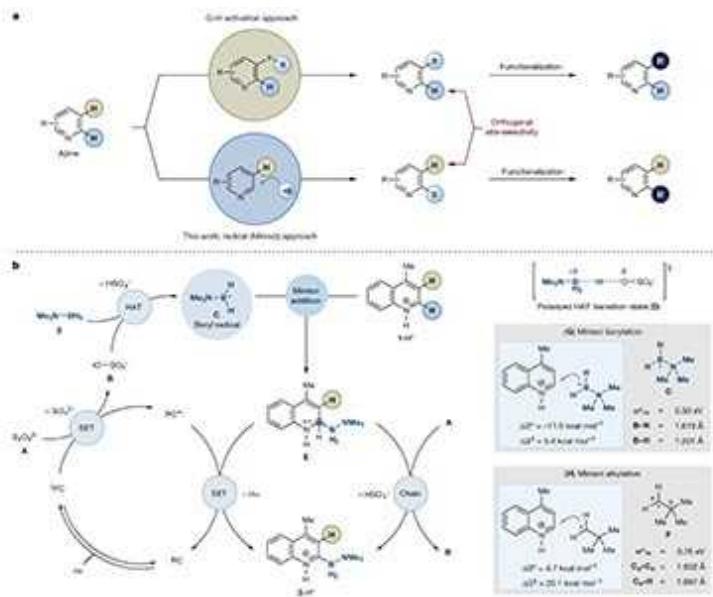


•

## [Download PDF](#)

The search for new pharmaceutical drugs by medicinal chemists relies on the synthesis of diverse compound libraries for biological testing. Of the reactions used for this purpose, those involving boron-containing organic molecules (known as organoboron compounds) are among the most popular because of the commercial availability and wide reactivity of these reagents<sup>1</sup>. [Writing in Nature](#), Kim *et al.*<sup>2</sup> report a method for the synthesis of organoboron compounds called borylated azines. The authors demonstrate that these compounds can participate in reactions common to other organoboron compounds, but offer distinct advantages, such as ease of

preparation and impressive stability. Importantly, these reagents will allow a full exploration of the therapeutic potential of molecules that have previously been difficult to prepare.

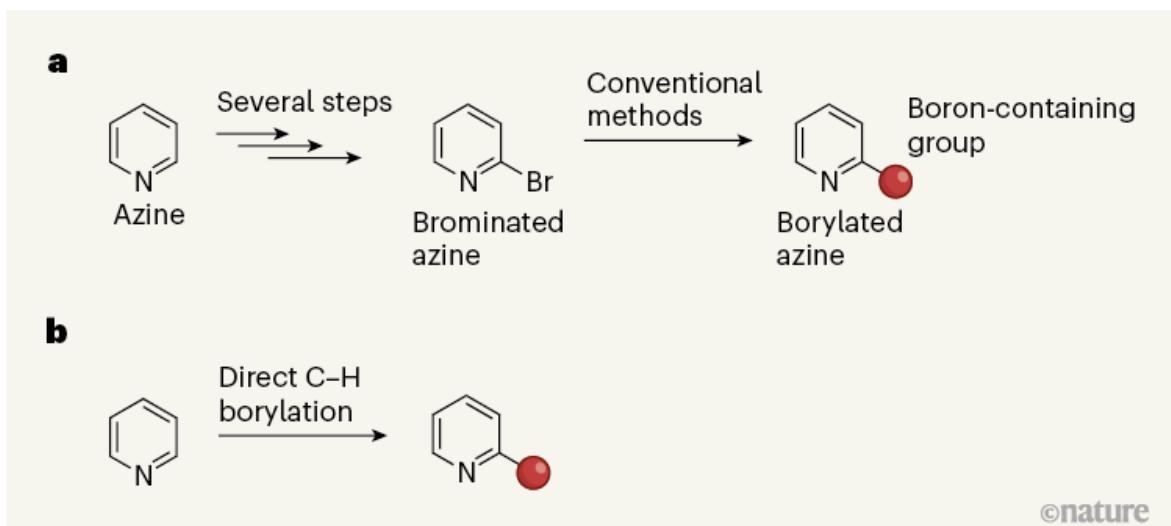


### [Read the paper: A radical approach for the selective C–H borylation of azines](#)

Azines are nitrogen-containing analogues of benzene rings and are present in many of the top-selling pharmaceuticals approved by the US Food and Drug Administration (see [go.nature.com/2dirpwf](http://go.nature.com/2dirpwf)). These medicines target several disease areas, including arthritis, diabetes and several types of cancer. If a boron-containing group is attached to an azine, the resulting borylated azine can be used as a reagent for the synthesis of many different azine-containing molecules — a crucial process for diversifying compound libraries. Kim *et al.* explored the use of transformations called C–H functionalization reactions to prepare borylated azines.

Although once regarded as an academic curiosity, C–H functionalization reactions are now a powerful methodology in organic synthesis. In these processes, a carbon–hydrogen (C–H) bond is converted into a carbon–X bond, where X can be any atom other than hydrogen. C–H functionalization can greatly streamline a synthetic procedure by reducing the number of steps it takes to get to a specific target molecule.

Borylated azines are typically synthesized from their bromine-containing analogues (Fig. 1a). But although these brominated starting materials are commercially available, they are more costly than the equivalent azines that lack a bromine atom. For example, 2-bromopyridine can be about 25 times more expensive per mole than pyridine, which reflects the cost of the synthetic steps needed to attach a bromine atom to pyridine<sup>3</sup>.

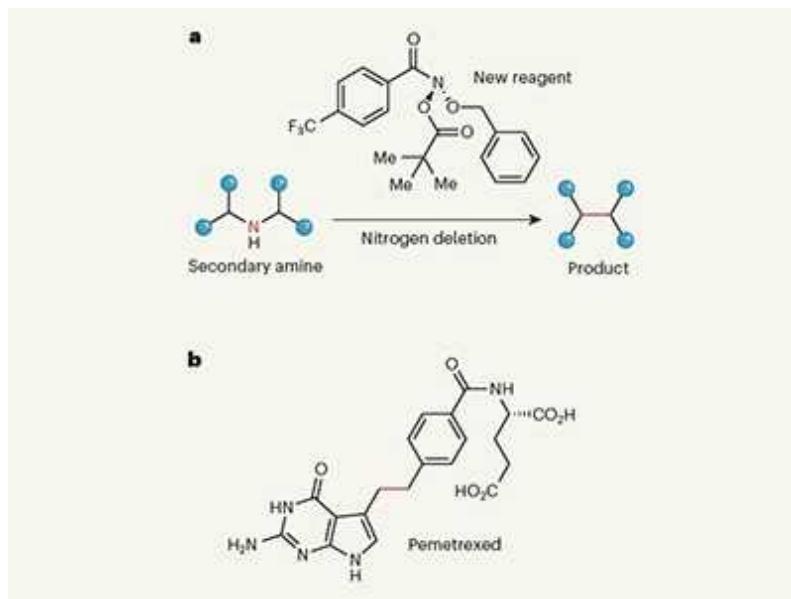


**Figure 1 | The synthesis of borylated azines.** **a**, Azines are analogues of benzene rings that contain one or more nitrogen atoms — the example shown is called a pyridine. Azines that bear a boron-containing group are called borylated azines, and are useful intermediates for organic synthesis. Conventional methods for making borylated azines involve adding a bromine atom to the starting azine (which might take several steps), and then replacing the bromine atom with the boron-containing group. **b**, Kim *et al.*<sup>2</sup> now report that a method known as C–H functionalization can be used to attach a boron-containing group directly at the  $\alpha$ -position of an azine. The resulting products are useful for making compounds with various other groups at the  $\alpha$ -position — a common molecular motif in pharmaceutical compounds.

A C–H functionalization reaction that directly introduces a boron-containing group to an azine can bypass these steps, saving valuable time and resources. If the conventional synthetic approach is like a bus route that has multiple transfers, C–H functionalization is like a non-stop express train. But before boarding this train, it is crucial to make sure that it stops at your desired

destination: azines often have several C–H bonds, which means that synthetic chemists must devise clever strategies to ensure that only the desired C–H bond reacts, rather than another.

The development of iridium-catalysed borylation methods was a notable advance in the field of C–H functionalization<sup>4</sup>. With azines, however, iridium-catalysed borylation occurs at a C–H bond distal to the nitrogen atom in the azine ring, mainly at the  $\beta$ -position (the second-nearest position to the nitrogen). These methods cannot provide products with a boron group adjacent to the nitrogen (the  $\alpha$ -position), which, in turn, means that the borylated products cannot be used to make molecules in which the azine ring is connected at the  $\alpha$ -position. This is a notable limitation, because many drug molecules contain azines attached at that site.

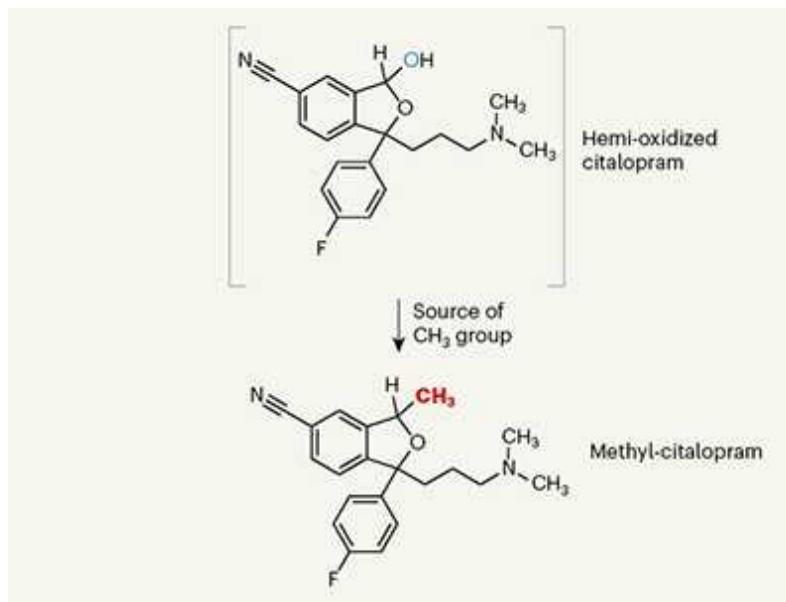


### Nitrogen deletion offers fresh strategy for organic synthesis

Kim and colleagues have now solved this problem by taking advantage of the inherent reactivity of azines towards radicals — atoms or molecules that contain an unpaired electron. This reactivity was heavily explored by the chemist Francesco Minisci in the 1960s with ‘alkyl’ radicals, which have the unpaired electron centred on a carbon atom. In the presence of acid, carbon-centred radicals react with azines predominantly at the  $\alpha$ -position, leading to attachment of a carbon atom to that site and loss of a hydrogen atom<sup>5</sup>.

In their work, Kim *et al.* show that boron-centred radicals can react similarly, allowing boron-containing groups to be installed at the  $\alpha$ -position and providing products that currently cannot be made using iridium-catalysed reactions (Fig. 1b). If the reactive  $\alpha$ -sites are blocked by attached chemical groups, borylation can occur at other positions around the azine ring. The authors' key discovery was that an amine-borane reagent ( $\text{Me}_3\text{N}\cdot\text{BH}_3$ , where Me is a methyl group,  $\text{CH}_3$ ), in the presence of an organic catalyst, acid and light, can form boron-centred radicals with the necessary reactivity for Minisci-type chemistry.

An added benefit of Kim and co-workers' chemistry is that it installs an amine-borane group ( $\text{BH}_2\text{-NMe}_3$ ) in the azine. The amine portion of the group ( $\text{NMe}_3$ ) confers stability on the borylated azine — without it, the boron atom has an incomplete shell of valence electrons and could undergo reactions that lead to the compound's decomposition. Indeed, the authors' borylated azines are stable enough to be stored in ambient conditions for several months, without the need for specialized storage.



### Manganese catalyst enables exploration of the magic methyl effect

Stable compounds can be unreactive, especially if the site of reactivity is blocked off — and in this case, the boron centre is blocked off by the attached amine group. Chemists might therefore be wondering whether

further reactions of the borylated azines are possible. The answer is yes: these compounds can participate in oxidations, and in various ‘cross-coupling’ reactions to form other types of bond (such as carbon–carbon, carbon–nitrogen and carbon–oxygen bonds), providing entry to medicinally relevant molecular scaffolds. The yields of some of the cross-coupling reactions are modest, but  $\alpha$ -borylated azines are generally challenging substrates for cross-coupling, owing to their tendency to decompose through protodeboronation (replacement of the boron group with a hydrogen).

It would now be useful to know whether these borylated azines can participate in reactions in which the amine-borane group is replaced with a fluorine atom<sup>6</sup>. Fluorine-containing groups are valuable motifs in medicinal chemistry, and therefore having access to fluorinated azines could be transformative in drug discovery. Moreover, if the authors’ chemistry can be used to incorporate fluorine into molecules as the final step of a synthesis, it might be possible to use it to label pharmaceutical compounds with a radioactive fluorine-18 atom — thereby allowing the compounds to be visualized in the body using positron emission tomography imaging<sup>7</sup>. Finally, given the balanced stability and reactivity profile of Kim and colleagues’ borylated azines, it will be interesting to see whether unique synthetic applications emerge for these reagents, particularly in situations for which conventional organoboron compounds fall short.

*Nature* **595**, 652–653 (2021)

doi: <https://doi.org/10.1038/d41586-021-02007-6>

## References

1. 1.

Lee, J. C. H. & Hall, D. G. in *Metal-Catalyzed Cross-Coupling Reactions and More* (eds de Meijere, A., Bräse, S. & Oestreich, M.) 65–132 (Wiley, 2014).

[Google Scholar](#)

2. 2.

Kim, J. H. *et al.* *Nature* **595**, 677–683 (2021).

[Article](#) [Google Scholar](#)

3. 3.

Allen, C. F. H. & Thirtle, J. R. *Org. Synth.* **26**, 16 (1946).

[PubMed](#) [Article](#) [Google Scholar](#)

4. 4.

Mkhalid, I. A. I., Barnard, J. H., Marder, T. B., Murphy, J. M. & Hartwig, J. F. *Chem. Rev.* **110**, 890–931 (2010).

[PubMed](#) [Article](#) [Google Scholar](#)

5. 5.

Duncton, M. A. J. *MedChemComm* **2**, 1135–1161 (2011).

[Article](#) [Google Scholar](#)

6. 6.

Pattison, G. *Org. Biomol. Chem.* **17**, 5651–5660 (2019).

[PubMed](#) [Article](#) [Google Scholar](#)

7. 7.

Tredwell, M. & Gouverneur, V. *Angew. Chem. Int. Edn* **51**, 11426–11437 (2012).

[Article](#) [Google Scholar](#)

[Download references](#)

## Competing Interests

The author declares no competing interests.

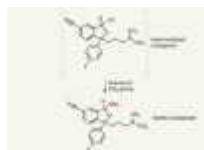
## Related Articles



- [Read the paper: A radical approach for the selective C–H borylation of azines](#)



- [Nitrogen deletion offers fresh strategy for organic synthesis](#)



- [Manganese catalyst enables exploration of the magic methyl effect](#)
- [See all News & Views](#)

## Subjects

- [Organic chemistry](#)
- [Synthesis](#)
- [Chemistry](#)

**nature** careers

## [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- **FACULTY POSITION IN MICROBIOME BIOCHEMISTRY**

Duke University

Durham, NC, United States

- **Bioinformatician (m/f/div)**

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

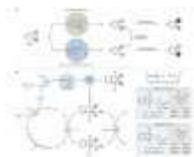
- **Student Research Assistant / HIWI (m/f/x) for Literature Research in the Field of Soil Microbiology**

Helmholtz Centre for Environmental Research (UFZ)

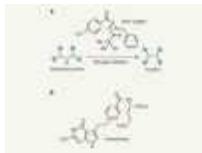
Leipzig, Germany

[Download PDF](#)

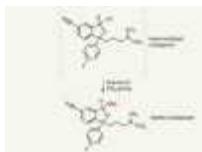
## Related Articles



- **Read the paper: A radical approach for the selective C–H borylation of azines**



- [Nitrogen deletion offers fresh strategy for organic synthesis](#)



- [Manganese catalyst enables exploration of the magic methyl effect](#)
- [See all News & Views](#)

## Subjects

- [Organic chemistry](#)
- [Synthesis](#)
- [Chemistry](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02007-6>

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

- NEWS AND VIEWS
- 21 July 2021

# Sperm CatSper ion channel swims into sharper focus

Activation of the CatSper membrane ion channel triggers enhanced sperm motility, and this protein complex is required for male fertility. Structural data now provide insights into the channel, and offer a few surprises, too.

- [David E. Clapham](#)    ORCID: <http://orcid.org/0000-0002-4459-9428><sup>0</sup>  
&
- [Raymond E. Hulse](#)    ORCID: <http://orcid.org/0000-0002-0110-3752><sup>1</sup>

## 1. [David E. Clapham](#)

1. David E. Clapham is at the Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA, and at Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia 20147, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## 2. [Raymond E. Hulse](#)

1. Raymond E. Hulse is at Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, Virginia 20147, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

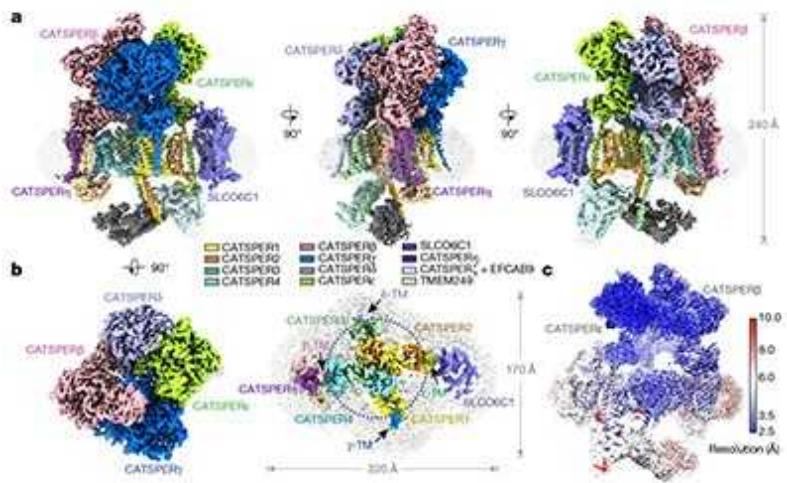




•

[Download PDF](#)

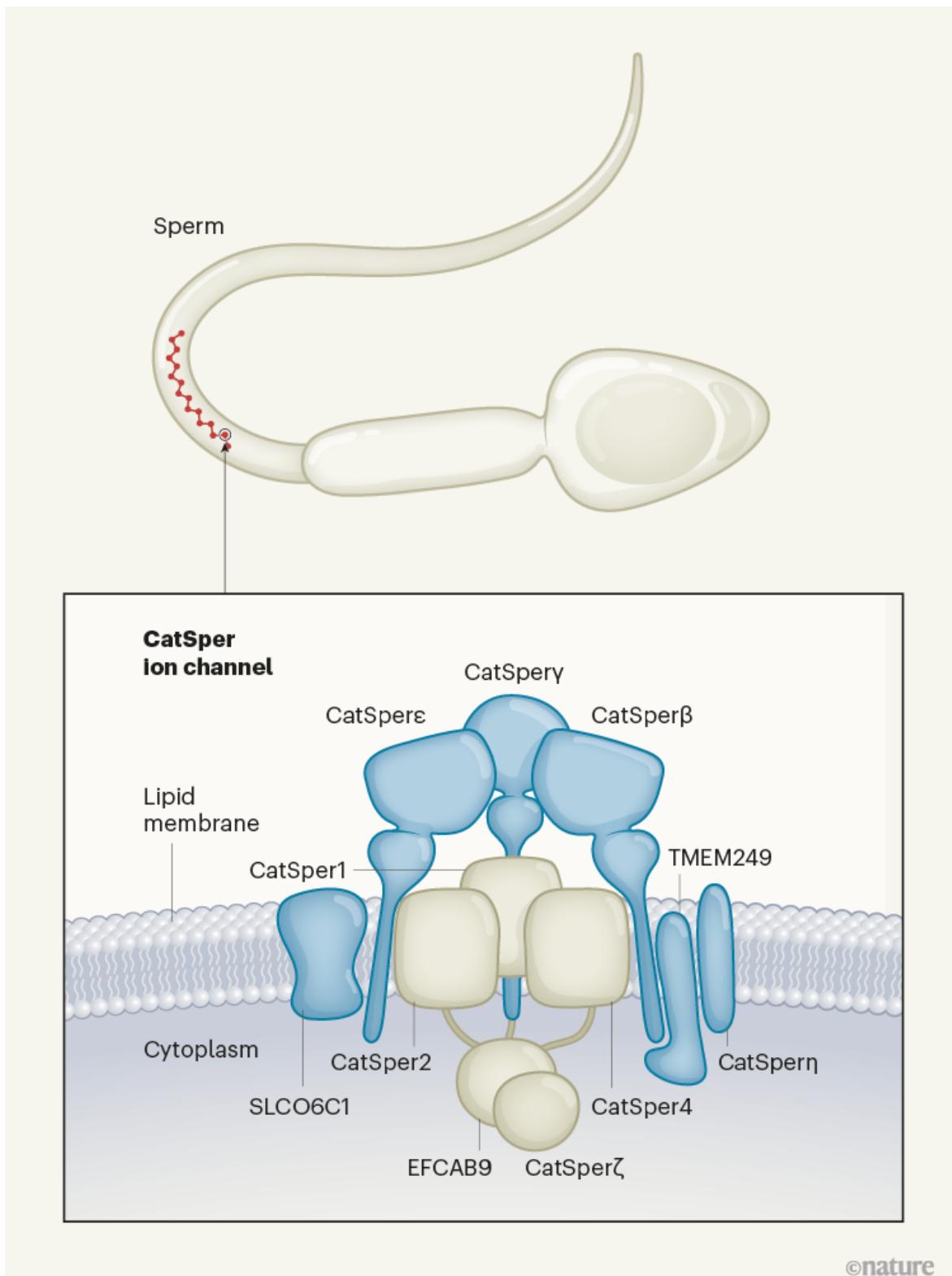
Many details about the structure of a key membrane ion channel required for normal sperm function have long remained mysterious. [Writing in Nature](#), Lin *et al.*<sup>1</sup> present data that fill some of the gaps in our knowledge.



## [Read the paper: Structure of a mammalian sperm cation channel complex](#)

Sperm and eggs are the vessels of our past and future, and they must overcome daunting challenges for the exchange of their genes to occur. A sperm cell enters a competitive race with millions of other sperm that are also trying to fertilize an egg cell. Each sperm has a long way to swim in the female reproductive tract, and this journey requires dramatic changes in the type of motion used, including a vigorous type of swimming movement that is triggered on encountering high pH. The fusion of a single sperm and egg during fertilization is another complex process that, after countless remarkable transformations, propagates a new genome in every cell of what will become an amazing, trillion-cell baby.

Not unlike anxious parents expecting their first child, scientists have also anticipated the arrival of structural insights into a sperm ion channel called CatSper (Fig. 1), which governs the hyperactive motility of sperm and is necessary for fertility<sup>2</sup>. Since its discovery 20 years ago<sup>3,4</sup>, efforts to determine the composition and structure of CatSper have been evolving. Early attempts at structure determination using high-resolution X-ray crystallography were stymied by the large number of protein subunits that CatSper contains (more than ten), and by problems encountered when, in structural investigations, these proteins are expressed in cells in which they are not normally found.



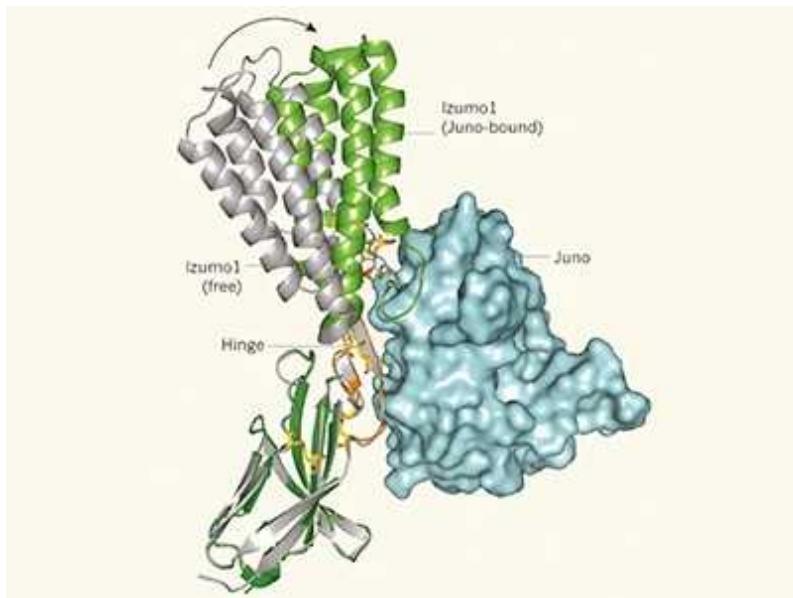
**Figure 1 | Structure of the CatSper ion channel.** Lin *et al.*<sup>1</sup> present structural data obtained using cryogenic electron microscopy. This channel,

arranged in a ‘racing-stripe’ pattern on the sperm tail, aids fertility by facilitating the entry of cations (positively charged calcium ions) needed for the hyperactive swimming motion of sperm. The channel’s central ion-conducting pore is formed from four subunits: CatSper1, CatSper2, CatSper3 and CatSper4 (CatSper3 is not shown to aid the visibility of other components). The authors report that some of the other proteins in the complex — CatSper $\beta$ , CatSper $\gamma$ , CatSper $\delta$  and CatSper $\epsilon$  — form a tent-like ‘pavilion’ over the pore (CatSper $\delta$  is not shown to aid visibility of other components). One of the proteins shown to be associated with the complex is SLCO6C1. Having a transporter of anions (negatively charged ions) associated with a cation channel is surprising. The newly discovered proteins associated with the channel include CatSper $\eta$  and TMEM249. The structure also indicates the location of EFCAB9 and CatSper $\zeta$ , which are proteins already known to be associated with the complex.

An ancestral version of this channel first arose in single-celled organisms called uniflagellates<sup>5</sup>. The strong evolutionary pressures exerted on CatSper over time have driven its complexity — the CatSper complex is arguably the most elaborate ion channel known. The proteins CatSper1, CatSper2, CatSper3 and CatSper4 form a tetramer that is the ion channel’s pore, and at least six other proteins are known to be associated with this complex (CatSper $\beta$ , CatSper $\gamma$ , CatSper $\delta$ , CatSper $\epsilon$ , CatSper $\zeta$  and EFCAB9)<sup>2,6</sup>. CatSper responds to high pH, and perhaps to other factors as well, when sperm reach the upper cervix, or utero-tubal junction, of the female reproductive tract. The activation of CatSper, which is associated with the entry of calcium ions into sperm<sup>7</sup>, completely transforms the motion of the sperm’s ‘tail’ (flagellum), generating a robust, asymmetrical beating pattern called hyperactivated motility. CatSper activity has a weak voltage dependence<sup>7</sup> (sensitivity to the voltage across the cell membrane), perhaps due to the different number of positively charged amino acids in each of the pore-forming unit’s voltage sensors.

Lin and colleagues now report structural data for CatSper. The authors generated mice that had a version of CatSper1 tagged with green fluorescent protein, together with copies of a peptide called FLAG. They used this tagged version of CatSper1 to isolate the CatSper complex from mouse

sperm cells, and now present the first structure of CatSper to be generated using cryogenic electron microscopy.



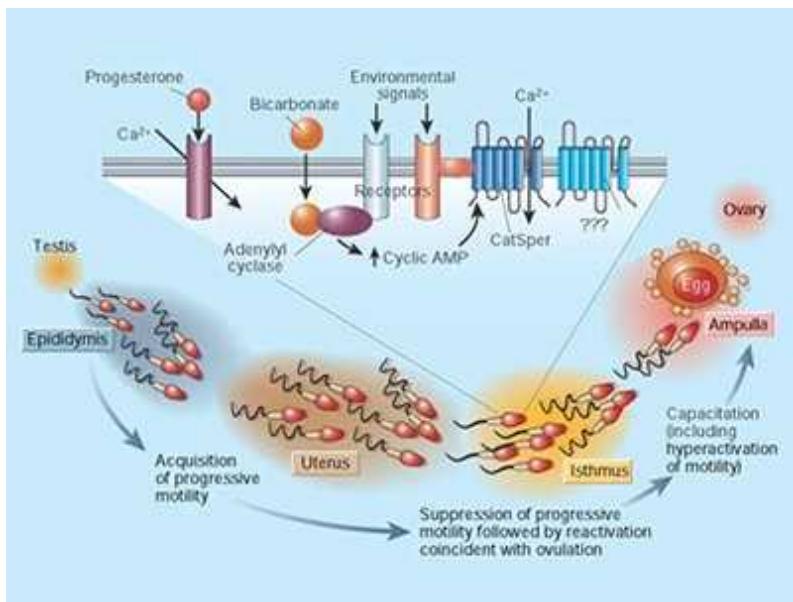
### When sperm meets egg

The structure reveals that four accessory transmembrane proteins (CatSper $\beta$ , CatSper $\gamma$ , CatSper $\delta$  and CatSper $\epsilon$ ) form a tent-like ‘pavilion’ over the pore (Fig. 1), with their pole-like, transmembrane-spanning supports positioned outside the pore, flanking the voltage sensors of the adjacent pore subunit. CatSper $\beta$  and CatSper $\epsilon$  are in contact with their voltage sensors, CatSper4 and CatSper2, respectively. The pavilion’s large interface, and particularly the presence of certain characteristic protein domains (for example, seven-bladed  $\beta$ -propeller domains), suggests regions that might bind to soluble ligand molecules. Notably, the multiple accessory components perturb the symmetry of the core tetrameric pore-forming subunits. This differs from the symmetry typically observed for related ion channels, and the discovery might reveal insights into the mechanisms involved in channel function.

The structure of CatSper brings several surprises. Structural modelling (and supporting evidence obtained using mass spectrometry) reveals that a protein called SLCO6C1, which is a type of transporter known as an organic anion transporter, links to CatSper $\epsilon$  and to the voltage-sensing domain of CatSper2. This observation is consistent with other work<sup>8,9</sup>. Another newly discovered protein association with CatSper is that of the transmembrane

protein CatSper $\eta$  — another previously unknown interactor — with the protein TMEM249 and an as-yet-unidentified cytoplasmic protein.

The cytoplasmic components are the least clearly visualized part of the structure, with the subcomplex of CatSper $\zeta$  and EFCAB9 revealing an unknown protein (consisting of  $\alpha$ -helices) and another unknown entity sandwiched between the elongated S6 segments of CatSper2 and CatSper3. Two lobes of EFCAB9 containing calcium-binding motifs (termed EF-hands) fit into the region where EFCAB9 and CatSper $\zeta$  interact. These domains might bind to calcium<sup>8</sup>, thereby linking calcium entry into sperm by means of intermediates to downstream proteins, such as the motor protein dynein. A major unanswered question is how the calcium signal results in an increased frequency and degree of bending of the sperm flagellar tail during swimming motions.



### Swimming with sperm

Exciting complementary data from other research groups, obtained using cryo-electron tomography and microscopy, show that the CatSper complex is assembled in a zigzag pattern<sup>9</sup> in each of the four quadrilateral ‘racing stripes’ along the sperm tail<sup>10</sup>. An EFCAB9–CatSper $\zeta$  pair bridges two staggered channel units (rotated by 180°) as a building block for the zigzag assembly, and it was proposed<sup>9</sup> that this longitudinal nanodomain allows the

simultaneous opening of the array of CatSper channels along the flagellum to allow rapid signalling in the sperm tail.

Lin and colleagues' structural data, together with these contributions, greatly increase our knowledge of CatSper, and suggest potential answers to, or ways of addressing, long-standing questions. For example, to what, if anything, do the extracellular pavilion structures bind? If binding partners exist for these structures, does their binding affect CatSper voltage sensitivity or calcium entry, or cause structural changes that propagate along CatSper and affect cytoplasmic components of the complex? And does CatSper activation alter the activity of the anion transporter?

These seminal contributions will speed up the identification of molecules that could form targets for the development of male-specific contraceptives (and species-specific forms of such contraceptives). Finally, the discoveries relating to CatSper might point the way to the treatment of male infertility associated with mutations in the genes encoding components of this complex.

*Nature* **595**, 654-655 (2021)

doi: <https://doi.org/10.1038/d41586-021-01945-5>

## References

1. 1.

Lin, S., Ke, M., Zhang, Y., Yan, Z. & Wu, J. *Nature* **595**, 746–750 (2021).

[Article](#) [Google Scholar](#)

2. 2.

Qi, H. *et al.* *Proc. Natl Acad. Sci. USA* **104**, 1219–1223 (2007).

[PubMed](#) [Article](#) [Google Scholar](#)

3. 3.

Ren, D. *et al.* *Nature* **413**, 603–609 (2001).

[PubMed](#) [Article](#) [Google Scholar](#)

4. 4.

Quill, T. A., Ren, D., Clapham, D. E. & Garbers, D. L. *Proc. Natl Acad. Sci. USA* **98**, 12527–12531 (2001).

[PubMed](#) [Article](#) [Google Scholar](#)

5. 5.

Cai, X., Wang, X. & Clapham, D. E. *Mol. Biol. Evol.* **31**, 2735–2740 (2014).

[PubMed](#) [Article](#) [Google Scholar](#)

6. 6.

Wang, H., McGoldrick, L. L. & Chung, J.-J. *Nature Rev. Urol.* **18**, 46–66 (2021).

[PubMed](#) [Article](#) [Google Scholar](#)

7. 7.

Kirichok, Y., Navarro, B. & Clapham, D. E. *Nature* **439**, 737–740 (2006).

[PubMed](#) [Article](#) [Google Scholar](#)

8. 8.

Hwang, J. Y. *et al.* *Cell* **177**, 1480–1494 (2019).

[PubMed](#) [Article](#) [Google Scholar](#)

9. 9.

Zhao, Y. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.06.19.448910> (2021).

10. 10.

Chung, J.-J. *et al.* *Cell* **157**, 808–822 (2014).

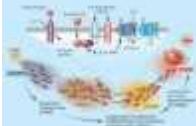
[PubMed](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Competing Interests

The authors declare no competing interests

## Related Articles

-  [Read the paper: Structure of a mammalian sperm cation channel complex](#)
-  [When sperm meets egg](#)
-  [Swimming with sperm](#)
- [See all News & Views](#)

## Subjects

- [Structural biology](#)
- [Cell biology](#)

## nature careers

### [Jobs](#) >

- [Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [FACULTY POSITION IN MICROBIOME BIOCHEMISTRY](#)

Duke University

Durham, NC, United States

- [Bioinformatician \(m/f/div\)](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

## Related Articles



- [Read the paper: Structure of a mammalian sperm cation channel complex](#)



- [When sperm meets egg](#)



- [Swimming with sperm](#)

- [See all News & Views](#)

## Subjects

- [Structural biology](#)
- [Cell biology](#)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01945-5>

- NEWS AND VIEWS
- 28 July 2021

# Earthquakes triggered by underground fluid injection modelled for a tectonically active oil field

An analysis of the Val d'Agri oil field in Italy provides insight into how processes associated with wastewater disposal trigger earthquakes — and how such effects can be reduced to maintain the economic viability of mature oil fields.

- [Mirko van der Baan](#) 0

1. [Mirko van der Baan](#)

1. Mirko van der Baan is in the Department of Physics, University of Alberta, Edmonton, Alberta T6G 2E1, Canada.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)





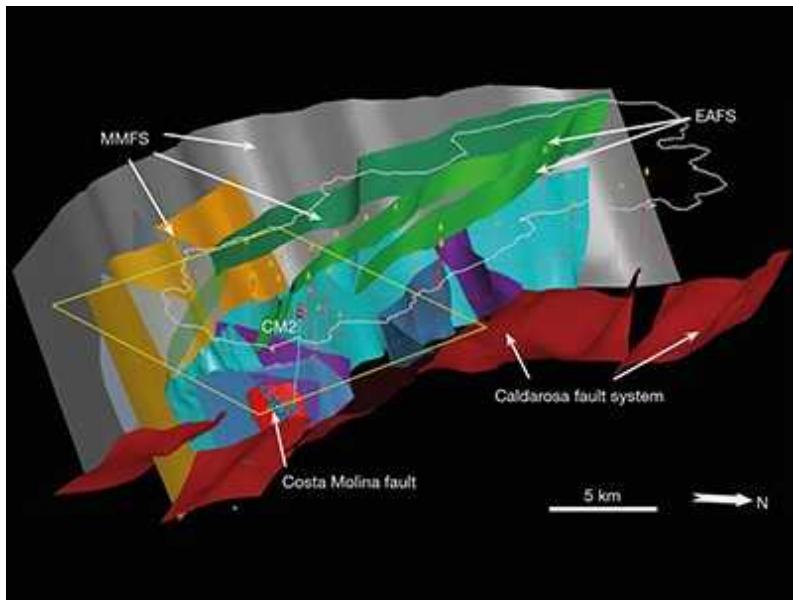
•

## [Download PDF](#)

Gas and oil extraction generates wealth — it can significantly boost the gross domestic product of a country. But water is also extracted with the hydrocarbons, and is often reinjected into the ground for disposal.

Unfortunately, large-scale fluid injection can induce earthquakes<sup>1</sup>, potentially leading to the termination of extraction before the full economic potential of an oil field has been realized. [Writing in Nature](#), Hager *et al.*<sup>2</sup> describe a multidisciplinary process to manage earthquake hazard in an active oil field, maintaining the economic viability of a field that uses fluid

injection for water disposal, while minimizing the likelihood of seismic activity that is sufficiently strong to be felt by humans.

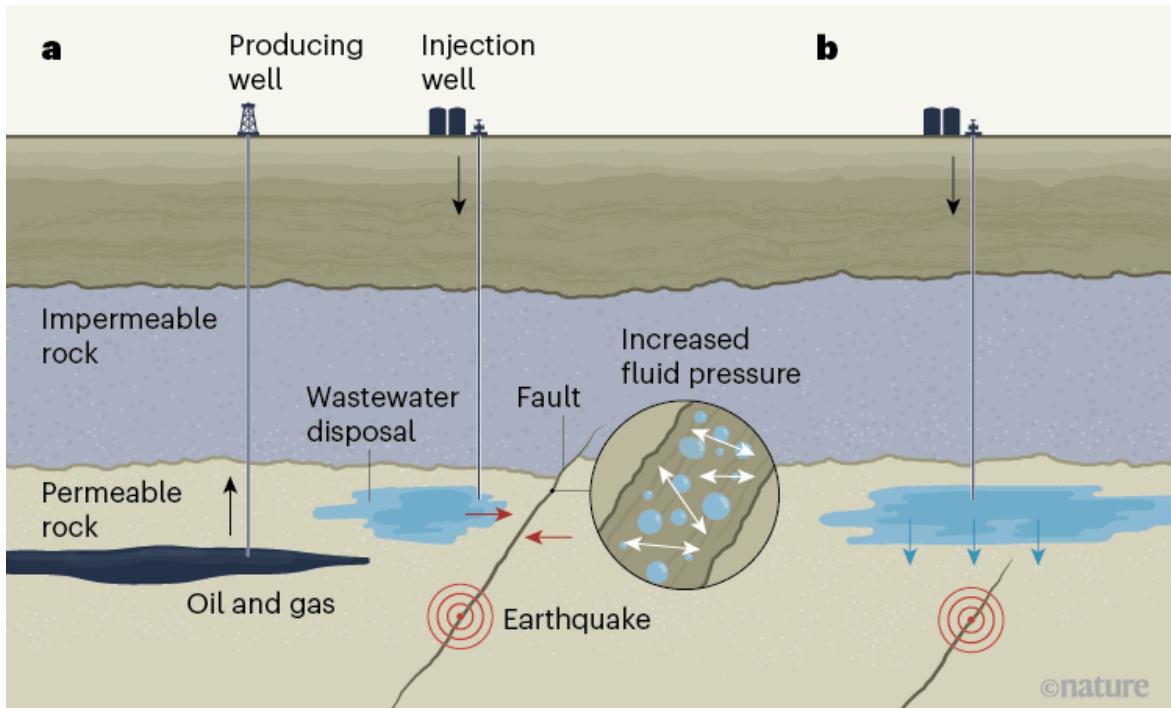


[Read the paper: A process-based approach to understanding and managing triggered seismicity](#)

Worldwide hydrocarbon extraction of both natural gas and petroleum liquids has increased steadily (see [go.nature.com/3hnqdat](https://go.nature.com/3hnqdat)) since the mid-1980s (although the COVID-19 pandemic has interrupted this long-term trend). The water produced during the process must be treated, recycled or disposed of, because it is salty and contaminated by hydrocarbons and other organic and inorganic compounds<sup>3</sup>. Moreover, wastewater production tends to increase with the maturity of an oil field. This effect, combined with the sustained increase in hydrocarbon extraction, means that the disposal of wastewater is a growing global challenge.

Although most earthquakes are caused by tectonic forces, they can also be triggered by fluid injection into bedrock, most commonly when fluids penetrate pre-existing faults. The associated increase in fluid pressures reduces frictional resistance to slip, which, in turn, can reactivate the fault and trigger an earthquake (Fig. 1). Fluid injection and hydrocarbon extraction can also cause large changes in volume or mass underground that exert stresses on nearby, pre-existing faults, resulting in seismic activity<sup>1,4</sup>. In the past few years, various regions have undergone significant changes in

earthquake-recurrence patterns owing to large-scale fluid injection — including Oklahoma in the United States<sup>5</sup>, the Sichuan Basin in China<sup>6</sup> and the Western Canadian Sedimentary Basin<sup>7</sup>. Such changes have been observed for fluid injection associated both with water disposal<sup>1,5</sup> and with hydraulic fracturing (fracking)<sup>8</sup>.



**Figure 1 | Mechanisms through which earthquakes can be triggered by fluid injection at oil and gas fields.** **a**, Crude oil and gas extracted by a producing well contains contaminated wastewater, which is often disposed of underground by an injection well. If the collected wastewater connects to a nearby fault, the increased fluid pressure in the fault can reduce frictional resistance to slip — potentially reactivating the fault and causing an earthquake. Tectonic forces (red arrows) can contribute to this process. **b**, Fluid injection and hydrocarbon extraction can also cause large changes in volume or mass that exert stresses (blue arrows) on underlying faults, resulting in seismic activity. Hager *et al.*<sup>2</sup> report a computational model of the Val d’Agri oil field in Italy that combines multiple data sources to simulate the effects of fluid injection and tectonic activity on seismicity in that region. They use the model to estimate the maximum amount of fluid injection that can be tolerated without triggering earthquakes sufficiently strong to be felt by humans. (Adapted from ref. 1.)

Fluid-flow simulations are typically used to investigate correlations between earthquake patterns and fluid injection<sup>9</sup>. This approach provides insight into the underlying drivers of earthquake occurrence in tectonically quiet areas. However, fluid simulations alone are probably insufficient for developing strategies to manage seismicity, particularly in areas in which tectonic earthquakes are common, because the dominant cause in such regions is explicitly ignored.

Hager *et al.* have developed a multidisciplinary approach to earthquake mitigation in Italy's Val d'Agri field, which is located in a tectonically active area. Val d'Agri is the largest onshore oil field in Western Europe.

Extraction started in 1993, and the field now accounts for more than half of Italy's oil production. Wastewater disposal started in 2006 and led to about 300 small seismic events (maximum local magnitude 2.2, which is too small to be felt). Historically, an average of about four tectonic earthquakes of moment magnitude equal to or greater than 5.5 (strong enough to shake and possibly damage buildings) occur each century within 100 kilometres of the fluid-injection site. So, what injection rate is safe — that is, unlikely to trigger substantial seismic activity?



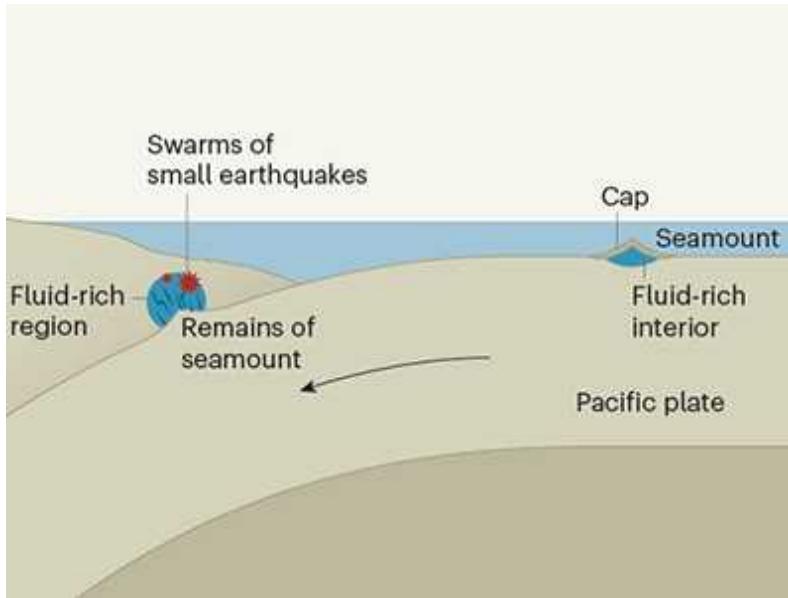
### Determining whether the worst earthquake has passed

The authors developed a multi-step, process-based approach to address this question. First, they produced a 3D structural model of the Val d'Agri

region,  $80 \times 50 \times 10$  km in size, containing 22 known major faults, surface topography, the top of the hydrocarbon reservoir and the rock that seals it off. The model contains the entire Val d'Agri oil field and includes 24 active hydrocarbon-producing wells. Fluid flow is coupled with geomechanical processes in the model to replicate the effects of external tectonic forces, hydrocarbon extraction and fluid injection from 1993 to 2016. The model parameters were estimated and calibrated using many data sources, including GPS data, records of well pressures and reflection seismology (which uses reflected seismic waves to determine the structure of Earth's subsurface).

Next, the authors constructed a smaller 3D model,  $13 \times 13 \times 15$  km in size, incorporating 17 faults and the producing wells in that region. This model focuses on the oil field's fluid-injection well and the associated earthquake locations. Coupled modelling of fluid flow and geomechanical processes was again carried out using the larger model to constrain the behaviour of the smaller model at its boundaries. The smaller model was then used to evaluate local stress and slip conditions at faults over time.

The simulations show that stresses have stabilized in most of the area around the injector well, because hydrocarbon extraction has reduced fluid pressures and therefore increased resistance to slip on most faults. Conversely, fluids have penetrated a fault near the injector well, causing the observed small-magnitude seismicity in that area. The authors then combined the output of the smaller model with earthquake-physics models. They found that the results matched available observations of fluid flow within the hydrocarbon reservoir and observed seismicity patterns — including the dependence on past injection rates of the location, timing and evolution of the earthquakes. The calibrated model suggests that a rate of fluid injection of 2,000 cubic metres per day, which corresponds to 50% of current total wastewater production, is unlikely to trigger noticeable seismicity above the tectonic background rate, whereas small events are increasingly likely to be triggered at higher rates of 2,500 and 3,000 cubic metres per day.



### Fluid-rich extinct volcanoes cause small earthquakes beneath New Zealand

Hager and colleagues' work is unusual for several reasons. Their method relies on the availability of detailed data and expert knowledge of many aspects of the region and its wells. Unfortunately, it is unusual to have such detailed information. The authors' results are the product of a highly fruitful partnership between academia and industry. As such, the findings provide insights that might result in the development of new industrial practices for managing and mitigating seismicity triggered by hydrocarbon extraction.

For instance, an ongoing case of triggered seismicity is the Groningen field in the Netherlands, the largest gas field in Europe. A gradual rise in seismicity since 1991 in this region caused property damage and led to increasingly vocal public discontent, resulting in the decision to terminate extraction in 2022. This will leave around 20% of potentially recoverable gas, worth about €70 billion (US\$83 billion), in the ground (see [go.nature.com/3blwz2c](https://go.nature.com/3blwz2c)). Application of Hager and co-workers' method to this region might enable the remaining gas to be extracted without causing further damage.

If the authors' approach can be extended to seismicity associated with hydrocarbon extraction, as well as that associated with wastewater injection, it might help to manage and mitigate the associated environmental impacts if used at the nascent stage of seismicity. Their method might also be suitable

for managing earthquakes associated with the sequestration of carbon dioxide<sup>10</sup> and engineered geothermal systems<sup>11,12</sup>.

Phenomena associated with hydrocarbon extraction are often contentious<sup>13</sup>, but Hager *et al.* have developed a process for managing and mitigating one of the most important adverse effects: induced seismicity. It is to be hoped that this will help the oil and gas industry to manage the balance between the economic viability and the environmental effects of extraction.

*Nature* **595**, 655–656 (2021)

doi: <https://doi.org/10.1038/d41586-021-01997-7>

## References

1. 1.

Ellsworth, W. L. *Science* **341**, 1225942 (2013).

[PubMed](#) [Article](#) [Google Scholar](#)

2. 2.

Hager, B. H. *et al.* *Nature* **595**, 684–689 (2021).

[Article](#) [Google Scholar](#)

3. 3.

Igunnu, E. T. & Chen, G. Z. *Int. J. Low-Carbon Technol.* **9**, 157–177 (2014).

[Article](#) [Google Scholar](#)

4. 4.

Segall, P. *Geology* **17**, 942–946 (1989).

[Article](#) [Google Scholar](#)

5. 5.

Langenbruch, C. & Zoback, M. D. *Sci. Adv.* **2**, e1601542 (2016).

[PubMed](#) [Article](#) [Google Scholar](#)

6. 6.

Lei, X. *et al.* *Sci. Rep.* **7**, 7971 (2017).

[PubMed](#) [Article](#) [Google Scholar](#)

7. 7.

Atkinson, G. M. *et al.* *Seismol. Res. Lett.* **87**, 631–647 (2016).

[Article](#) [Google Scholar](#)

8. 8.

Schultz, R. *et al.* *Rev. Geophys.* **58**, e2019RG000695 (2020).

[Article](#) [Google Scholar](#)

9. 9.

Keranen, K. M., Weingarten, M., Abers, G. A., Bekins, B. A. & Ge, S. *Science* **345**, 448–451 (2014).

[PubMed](#) [Article](#) [Google Scholar](#)

10. 10.

Zoback, M. D. & Gorelick, S. M. *Proc. Natl Acad. Sci. USA* **109**, 10164–10168 (2012).

[PubMed](#) [Article](#) [Google Scholar](#)

11. 11.

Grigoli, F. *et al.* *Science* **360**, 1003–1006 (2018).

[PubMed](#) [Article](#) [Google Scholar](#)

12. 12.

Kim, K.-H. *et al.* *Science* **360**, 1007–1009 (2018).

[PubMed](#) [Article](#) [Google Scholar](#)

13. 13.

Howarth, R. W., Ingraffea, A. & Engelder, T. *Nature* **477**, 271–275 (2011).

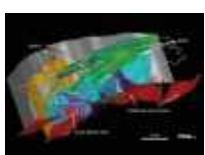
[PubMed](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Competing Interests

The author receives research funding from the hydrocarbon industry.

## Related Articles



- [Read the paper: A process-based approach to understanding and managing triggered seismicity](#)



- [Determining whether the worst earthquake has passed](#)



- [\*\*Fluid-rich extinct volcanoes cause small earthquakes beneath New Zealand\*\*](#)
- [\*\*See all News & Views\*\*](#)

## Subjects

- [Solid Earth sciences](#)
- [Seismology](#)
- [Geophysics](#)
- [Environmental sciences](#)

**nature** careers

## [Jobs](#) >

- [\*\*Chair Professor/Professor/Associate Professor/Assistant Professor in the Department of Physics\*\*](#)

City University of Hong Kong (CityU)

Hong Kong, China

- [\*\*FACULTY POSITION IN MICROBIOME BIOCHEMISTRY\*\*](#)

Duke University

Durham, NC, United States

- [\*\*Bioinformatician \(m/f/div\)\*\*](#)

Max Planck Institute for Plant Breeding Research (MPIPZ)

Cologne, Germany

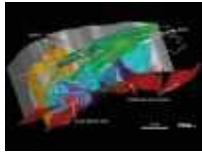
- [Student Research Assistant / HIWI \(m/f/x\) for Literature Research in the Field of Soil Microbiology](#)

Helmholtz Centre for Environmental Research (UFZ)

Leipzig, Germany

[Download PDF](#)

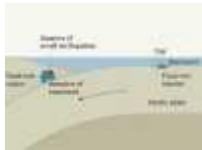
## Related Articles



- [Read the paper: A process-based approach to understanding and managing triggered seismicity](#)



- [Determining whether the worst earthquake has passed](#)



- [Fluid-rich extinct volcanoes cause small earthquakes beneath New Zealand](#)
- [See all News & Views](#)

## Subjects

- [Solid Earth sciences](#)
- [Seismology](#)

- [Geophysics](#)
  - [Environmental sciences](#)
- 

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01997-7>

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

- Article
- [Published: 28 July 2021](#)

# Light bending and X-ray echoes from behind a supermassive black hole

- [D. R. Wilkins](#) ORCID: [orcid.org/0000-0002-4794-5998<sup>1</sup>](https://orcid.org/0000-0002-4794-5998),
- [L. C. Gallo<sup>2</sup>](#),
- [E. Costantini<sup>3,4</sup>](#),
- [W. N. Brandt<sup>5,6,7</sup>](#) &
- [R. D. Blandford<sup>1</sup>](#)

*Nature* volume **595**, pages 657–660 (2021) [Cite this article](#)

- 1918 Accesses
- 1396 Altmetric
- [Metrics details](#)

## Subjects

- [Compact astrophysical objects](#)
- [General relativity and gravity](#)
- [High-energy astrophysics](#)

## Abstract

The innermost regions of accretion disks around black holes are strongly irradiated by X-rays that are emitted from a highly variable, compact

corona, in the immediate vicinity of the black hole<sup>1,2,3</sup>. The X-rays that are seen reflected from the disk<sup>4</sup>, and the time delays, as variations in the X-ray emission echo or ‘reverberate’ off the disk<sup>5,6</sup>, provide a view of the environment just outside the event horizon. I Zwicky 1 (I Zw 1) is a nearby narrow-line Seyfert 1 galaxy<sup>7,8</sup>. Previous studies of the reverberation of X-rays from its accretion disk revealed that the corona is composed of two components: an extended, slowly varying component extending over the surface of the inner accretion disk, and a collimated core, with luminosity fluctuations propagating upwards from its base, which dominates the more rapid variability<sup>9,10</sup>. Here we report observations of X-ray flares emitted from around the supermassive black hole in I Zw 1. X-ray reflection from the accretion disk is detected through a relativistically broadened iron K line and Compton hump in the X-ray emission spectrum. Analysis of the X-ray flares reveals short flashes of photons consistent with the re-emergence of emission from behind the black hole. The energy shifts of these photons identify their origins from different parts of the disk<sup>11,12</sup>. These are photons that reverberate off the far side of the disk, and are bent around the black hole and magnified by the strong gravitational field. Observing photons bent around the black hole confirms a key prediction of general relativity.

## 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.

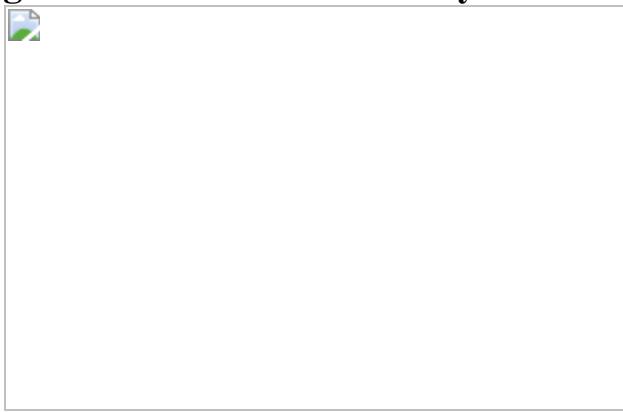
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

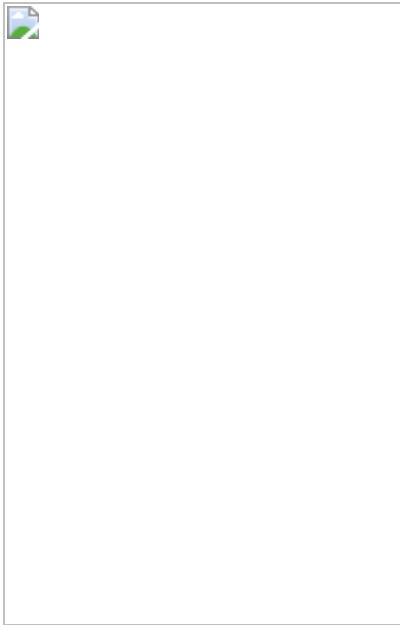
**Fig. 1: X-ray light curves of the AGN I Zw 1.**



**Fig. 2: Schematic of the X-ray reverberation model.**



**Fig. 3: Modelled response of the reverberating iron K fluorescence line to a flare of continuum emission.**



**Fig. 4: Features of the reverberation response function detected during the X-ray flares.**



## Data availability

The data used in this study from NuSTAR and XMM-Newton are publicly available. NuSTAR observations can be accessed via the NASA High Energy Astrophysics Science Archive Research Center (<https://heasarc.gsfc.nasa.gov>). This work includes data obtained during NuSTAR observation ID 60501030002. XMM-Newton observations can be accessed via the XMM-Newton Science Archive (<http://nxsa.esac.esa.int/nxsa-web>). The primary analysis was conducted on observation IDs 0851990101 and 0851990201. Observation IDs 0110890301, 0300470101, 0743050301 and 0743050801 were included in the measurement of the average reverberation timescale.

## Code availability

X-ray spectroscopic analysis was conducted using xspec, which is freely available as part of the Heasoft package (<https://heasarc.gsfc.nasa.gov/docs/software/heasoft>). The emcee MCMC sampling code is available at <http://emcee.readthedocs.io>, and the X-spec implementation by J. Sanders is available at [https://github.com/jeremysanders/xspec\\_emcee](https://github.com/jeremysanders/xspec_emcee). The relxill X-ray reflection model is available at <http://www.sternwarte.uni-erlangen.de/~dauser/research/relxill/>. Analysis of light curves and X-ray reverberation was conducted using pyLag, publicly available at <https://github.com/wilkinsdr/pyLag>. The CUDA Kerr code used for reverberation modelling is available on request from the corresponding author.

## References

1. 1.

Galeev, A. A., Rosner, R. & Vaiana, G. S. Structured coronae of accretion disks. *Astrophys. J.* **229**, 318–326 (1979).

[ADS](#) [Article](#) [Google Scholar](#)

2. 2.

Haardt, F. & Maraschi, L. A two-phase model for the X-ray emission from Seyfert galaxies. *Astrophys. J. Lett.* **380**, L51 (1991).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Merloni, A. & Fabian, A. C. Thunderclouds and accretion discs: a model for the spectral and temporal variability of Seyfert 1 galaxies. *Mon. Not. R. Astron. Soc.* **328**, 958–968 (2001).

[ADS](#) [Article](#) [Google Scholar](#)

4. 4.

George, I. M. & Fabian, A. C. X-ray reflection from cold matter in active galactic nuclei and X-ray binaries. *Mon. Not. R. Astron. Soc.* **249**, 352–367 (1991).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Uttley, P., Cackett, E. M., Fabian, A. C., Kara, E. & Wilkins, D. R. X-ray reverberation around accreting black holes. *Astron. Astrophys. Rev.* **22**, 72 (2014).

[ADS](#) [Article](#) [Google Scholar](#)

6. 6.

Fabian, A. C. et al. Broad line emission from iron K- and L-shell transitions in the active galaxy 1H0707–495. *Nature* **459**, 540–542 (2009).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Boller, T., Brandt, W. N. & Fink, H. Soft X-ray properties of narrow-line Seyfert 1 galaxies. *Astron. Astrophys.* **305**, 53 (1996).

[ADS](#) [Google Scholar](#)

8. 8.

Gallo, L. C. Investigating the nature of narrow-line Seyfert 1 galaxies with high-energy spectral complexity. *Mon. Not. R. Astron. Soc.* **368**, 479–486 (2006).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Wilkins, D. R. et al. Revealing structure and evolution within the corona of the Seyfert galaxy I Zw 1. *Mon. Not. R. Astron. Soc.* **471**, 4436–4451 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Gallo, L. C., Brandt, W. N., Costantini, E. & Fabian, A. C. A longer XMM-Newton look at I Zwicky 1—distinct modes of X-ray spectral variability. *Mon. Not. R. Astron. Soc.* **377**, 1375–1382 (2007).

[ADS](#) [Article](#) [Google Scholar](#)

11. 11.

Reynolds, C. S., Young, A. J., Begelman, M. C. & Fabian, A. C. X-ray iron line reverberation from black hole accretion disks. *Astrophys. J.* **514**, 164–179 (1999).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Wilkins, D. R., Cackett, E. M., Fabian, A. C. & Reynolds, C. S. Towards modelling X-ray reverberation in AGN: piecing together the extended corona. *Mon. Not. R. Astron. Soc.* **458**, 200–225 (2016).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Harrison, F. A. et al. The Nuclear Spectroscopic Telescope Array (NuSTAR) high-energy X-ray mission. *Astrophys. J.* **770**, 103 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

14. 14.

Jansen, F. et al. XMM-Newton observatory. I. The spacecraft and operations. *Astron. Astrophys.* **365**, L1–L6 (2001).

[ADS](#) [Article](#) [Google Scholar](#)

15. 15.

Risaliti, G. et al. A rapidly spinning supermassive black hole at the centre of NGC 1365. *Nature* **494**, 449–451 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Fabian, A. C., Rees, M. J., Stella, L. & White, N. E. X-ray fluorescence from the inner disc in Cygnus X-1. *Mon. Not. R. Astron. Soc.* **238**, 729–736 (1989).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Wilkins, D. R. et al. Flaring from the supermassive black hole in Mrk 335 studied with Swift and NuSTAR. *Mon. Not. R. Astron. Soc.* **454**, 4440–4451 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Strüder, L. et al. The European photon imaging camera on XMM-Newton: the pn-CCD camera. *Astron. Astrophys.* **365**, L18–L26 (2001).

[ADS](#) [Article](#) [Google Scholar](#)

19. 19.

Dauser, T. et al. Normalizing a relativistic model of X-ray reflection. Definition of the reflection fraction and its implementation in relxill. *Astron. Astrophys.* **590**, A76 (2016).

[Article](#) [Google Scholar](#)

20. 20.

García, J. & Kallman, T. R. X-ray reflected spectra from accretion disk models. I. Constant density atmospheres. *Astrophys. J.* **718**, 695–706 (2010).

[ADS](#) [Article](#) [Google Scholar](#)

21. 21.

García, J., Kallman, T. R. & Mushotzky, R. F. X-ray reflected spectra from accretion disk models. II. Diagnostic tools for X-ray observations. *Astrophys. J.* **731**, 131 (2011).

[ADS](#) [Article](#) [Google Scholar](#)

22. 22.

García, J. et al. X-Ray reflected spectra from accretion disk models. III. A complete grid of ionized reflection calculations. *Astrophys. J.* **768**, 146 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

23. 23.

Dauser, T., Wilms, J., Reynolds, C. S. & Brenneman, L. W. Broad emission lines for a negatively spinning black hole. *Mon. Not. R. Astron. Soc.* **409**, 1534–1540 (2010).

[ADS](#) [Article](#) [Google Scholar](#)

24. 24.

Silva, C. V. et al. The variability of the warm absorber in I Zwicky 1 as seen by XMM-Newton. *Mon. Not. R. Astron. Soc.* **480**, 2334–2342 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Costantini, E., Gallo, L. C., Brandt, W. N., Fabian, A. C. & Boller, T. A longer XMM-Newton look at I Zwicky 1: physical conditions and variability of the ionized absorbers. *Mon. Not. R. Astron. Soc.* **378**, 873–880 (2007).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

26. 26.

García, J. A. et al. The effects of high density on the X-ray spectrum reflected from accretion discs around black holes. *Mon. Not. R. Astron. Soc.* **462**, 751–760 (2016).

[ADS](#) [Article](#) [Google Scholar](#)

27. 27.

Done, C., Davis, S. W., Jin, C., Blaes, O. & Ward, M. Intrinsic disc emission and the soft X-ray excess in active galactic nuclei. *Mon. Not. R. Astron. Soc.* **420**, 1848–1860 (2012).

[ADS](#) [Article](#) [Google Scholar](#)

28. 28.

Wilkins, D. R. & Fabian, A. C. Understanding X-ray reflection emissivity profiles in AGN: locating the X-ray source. *Mon. Not. R. Astron. Soc.* **424**, 1284–1296 (2012).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Wilkins, D. R. & Fabian, A. C. The origin of the lag spectra observed in AGN: reverberation and the propagation of X-ray source fluctuations. *Mon. Not. R. Astron. Soc.* **430**, 247–258 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

30. 30.

Press, W. H. & Schechter, P. Remark on the statistical significance of flares in Poisson count data. *Astrophys. J.* **193**, 437–442 (1974).

[ADS](#) [Article](#) [Google Scholar](#)

31. 31.

Emmanoulopoulos, D., McHardy, I. M. & Papadakis, I. E. Generating artificial light curves: revisited and updated. *Mon. Not. R. Astron. Soc.* **433**, 907–927 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

32. 32.

Iwasawa, K., Miniutti, G. & Fabian, A. C. Flux and energy modulation of redshifted iron emission in NGC 3516: implications for the black hole mass. *Mon. Not. R. Astron. Soc.* **355**, 1073–1079 (2004).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

33. 33.

Narayan, R. & Quataert, E. Black hole accretion. *Science* **307**, 77 (2005).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

34. 34.

Huang, Y.-K. et al. Reverberation mapping of the narrow-line Seyfert 1 galaxy I Zwicky 1: black hole mass. *Astrophys. J.* **876**, 102 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

35. 35.

Vestergaard, M. & Peterson, B. M. Determining central black hole masses in distant active galaxies and quasars. II. Improved optical and UV scaling relationships. *Astrophys. J.* **641**, 689–709 (2006).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

36. 36.

Collin, S., Kawaguchi, T., Peterson, B. M. & Vestergaard, M. Systematic effects in measurement of black hole masses by emission-line reverberation of active galactic nuclei: Eddington ratio and inclination. *Astron. Astrophys.* **456**, 75–90 (2006).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

This work was supported by the NASA NuSTAR and XMM-Newton Guest Observer programmes under grants 80NSSC20K0041 and 80NSSC20K0838. D.R.W. received additional support from a Kavli Fellowship at Stanford University. W.N.B. acknowledges support from the V. M. Willaman Endowment. Computing for this project was performed on the Sherlock cluster. D.R.W. thanks Stanford University and the Stanford Research Computing Center for providing computational resources and support.

## Author information

## Affiliations

1. Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA, USA  
D. R. Wilkins & R. D. Blandford
2. Department of Astronomy and Physics, Saint Mary's University, Halifax, Nova Scotia, Canada  
L. C. Gallo
3. SRON, Netherlands Institute for Space Research, Utrecht, The Netherlands  
E. Costantini
4. Anton Pannekoek Institute for Astronomy, University of Amsterdam, Amsterdam, The Netherlands  
E. Costantini
5. Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA, USA  
W. N. Brandt
6. Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA, USA  
W. N. Brandt
7. Department of Physics, The Pennsylvania State University, University Park, PA, USA  
W. N. Brandt

## Authors

1. D. R. Wilkins

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. L. C. Gallo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. E. Costantini

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. W. N. Brandt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. R. D. Blandford

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

D.R.W. performed the data analysis and reverberation modelling. L.C.G. contributed to the analysis of the XMM-Newton spectra. W.N.B., E.C. and R.D.B. contributed to the interpretation and discussion of the results.

## Corresponding author

Correspondence to [D. R. Wilkins](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

### [Extended Data Fig. 1 The X-ray spectrum of I Zw 1.](#)

**a**, The soft and hard X-ray spectra of I Zw 1 obtained by XMM-Newton and NuSTAR during the 2020 observations, fitted with a model describing the emission in the 3–50-keV band as the combination of directly observed continuum emission and its reflection from the disk, described by relxill. **b**, Ratio of the X-ray spectrum in the 3–78-keV energy range measured with NuSTAR to the best-fitting power-law continuum model, showing the residual features that arise because of the X-rays reflected from the inner accretion disk: the broad iron K line (around 6.4 keV) with redshifted wing, and the Compton hump centred at 25 keV. Error bars represent  $1\sigma$  uncertainties due to Poisson noise.

### [Extended Data Fig. 2 The relative response times of the X-ray emission at different energies to variations in luminosity.](#)

The time lag is computed from the cross-spectrum, averaged over Fourier frequency components from  $5 \times 10^{-4}$  to  $7 \times 10^{-4}$  Hz from all XMM-Newton observations of I Zw 1 between 2002 and 2020. The lag versus energy spectrum shows the delay of the response of the iron K line and soft X-ray excess, which reverberate from the accretion disk, with respect to the

continuum, which is most dominant in the 1–2-keV energy band. Error bars show the  $1\sigma$  confidence interval.

### [Extended Data Fig. 3 Detection of short peaks in the X-ray light curves.](#)

**a**, The raw light curves obtained from the XMM-Newton EPIC pn camera in the 3–5-keV, 5–7-keV and 6–8-keV energy bands, in 100-s time bins. The blue line shows the smoothed light-curve model with respect to which the significance of short peaks is assessed. The panel beneath each light curve shows the reduction in the  $\chi^2$  statistic over this model, when a short Gaussian peak is added in different time bins with different normalization. Shaded regions correspond to  $1\sigma$  (blue),  $2\sigma$  (yellow) and  $3\sigma$  (red) detections. The bottom panel shows the maximum reduction in  $\chi^2$  when a flare is added in each time bin. The peaks of interest that are offset between the bands appear at 6.8 ks and 18.8 ks. **b**, The same, with light curves in each energy band summed between the first and second flare, aligning the 3–5-keV peaks at 6.8 ks and 18.8 ks.

### [Extended Data Fig. 4 The coherence between the 3–5-keV and 1–3-keV light curves.](#)

The coherence describes the fraction of the variability in each light curve, at different Fourier frequencies, that can be predicted based on a linear transformation (for example the reverberation response) of the variability in the other light curve. The observed coherence is compared with predictions of pairs of random, red noise light curves between which a fraction  $u$  of variability is uncorrelated. Much of the drop in coherence at high frequencies is due to Poisson noise, and the high values of the coherence measured at low frequencies set tight constraints on the level of uncorrelated variability that can exist between the light curves that is not due to Poisson noise.

### **Extended Data Table 1 Parameters of the X-ray continuum and reflection from the accretion disk**

[Full size table](#)

# Supplementary information

## Supplementary Information

This file contains alternative explanations for the short-timescale variability and notes regarding the mass of the black hole.

## Peer Review File

# Rights and permissions

## Reprints and Permissions

# About this article



## Cite this article

Wilkins, D.R., Gallo, L.C., Costantini, E. *et al.* Light bending and X-ray echoes from behind a supermassive black hole. *Nature* **595**, 657–660 (2021). <https://doi.org/10.1038/s41586-021-03667-0>

## Download citation

- Received: 14 August 2020
- Accepted: 24 May 2021
- Published: 28 July 2021
- Issue Date: 29 July 2021

- DOI: <https://doi.org/10.1038/s41586-021-03667-0>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03667-0>

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

- Article
- [Published: 28 July 2021](#)

# Incoherent transport across the strange-metal regime of overdoped cuprates

- [J. Ayres](#) [ORCID: orcid.org/0000-0003-3466-4430](#)<sup>1,2,3 na1</sup>,
- [M. Berben](#)<sup>2,3 na1</sup>,
- [M. Čulo](#)<sup>2,3 nAff10</sup>,
- [Y.-T. Hsu](#)<sup>2,3</sup>,
- [E. van Heumen](#) [ORCID: orcid.org/0000-0001-9536-578X](#)<sup>4,5</sup>,
- [Y. Huang](#)<sup>4</sup>,
- [J. Zaanen](#)<sup>6</sup>,
- [T. Kondo](#) [ORCID: orcid.org/0000-0002-3912-5172](#)<sup>7</sup>,
- [T. Takeuchi](#)<sup>8</sup>,
- [J. R. Cooper](#)<sup>9</sup>,
- [C. Putzke](#) [ORCID: orcid.org/0000-0002-6205-9863](#)<sup>1</sup>,
- [S. Friedemann](#) [ORCID: orcid.org/0000-0002-0606-4991](#)<sup>1</sup>,
- [A. Carrington](#) [ORCID: orcid.org/0000-0001-8624-6148](#)<sup>1</sup> &
- [N. E. Hussey](#) [ORCID: orcid.org/0000-0002-8590-8979](#)<sup>1,2,3</sup>

[Nature](#) volume **595**, pages 661–666 (2021) [Cite this article](#)

- 993 Accesses
- 52 Altmetric
- [Metrics details](#)

## Subjects

- [Electronic properties and materials](#)
- [Superconducting properties and materials](#)

## Abstract

Strange metals possess highly unconventional electrical properties, such as a linear-in-temperature resistivity<sup>1,2,3,4,5,6</sup>, an inverse Hall angle that varies as temperature squared<sup>7,8,9</sup> and a linear-in-field magnetoresistance<sup>10,11,12,13</sup>. Identifying the origin of these collective anomalies has proved fundamentally challenging, even in materials such as the hole-doped cuprates that possess a simple bandstructure. The prevailing consensus is that strange metallicity in the cuprates is tied to a quantum critical point at a doping  $p^*$  inside the superconducting dome<sup>14,15</sup>. Here we study the high-field in-plane magnetoresistance of two superconducting cuprate families at doping levels beyond  $p^*$ . At all dopings, the magnetoresistance exhibits quadrature scaling and becomes linear at high values of the ratio of the field and the temperature, indicating that the strange-metal regime extends well beyond  $p^*$ . Moreover, the magnitude of the magnetoresistance is found to be much larger than predicted by conventional theory and is insensitive to both impurity scattering and magnetic field orientation. These observations, coupled with analysis of the zero-field and Hall resistivities, suggest that despite having a single band, the cuprate strange-metal region hosts two charge sectors, one containing coherent quasiparticles, the other scale-invariant ‘Planckian’ dissipators.

## 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.

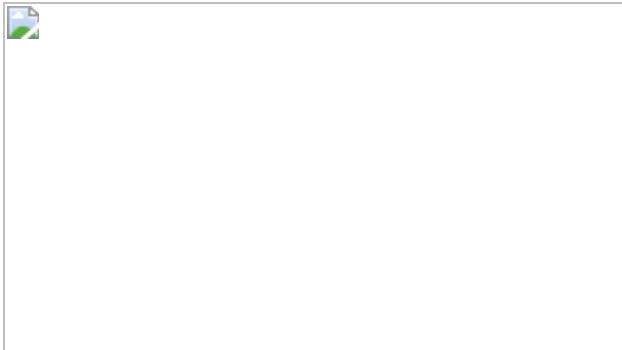
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

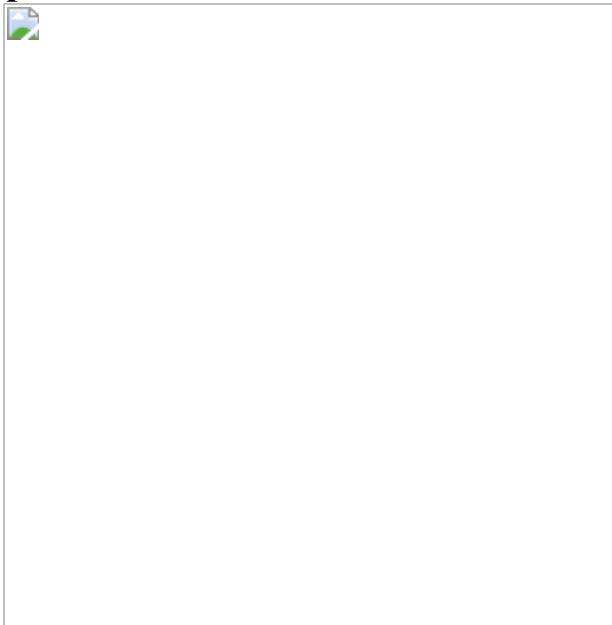
**Fig. 1: The strange-metal regime of overdoped cuprates.**



**Fig. 2: Quadrature scaling of the in-plane MR in heavily overdoped cuprates.**



**Fig. 3: Evidence for incoherent transport in heavily overdoped cuprates.**



## Data availability

The data that support the plots within this paper and other findings of this study are available from the Bristol data repository, data.bris, at  
<https://doi.org/10.5523/bris.150s0lqyd3eh61zsiaj8cj5vqd>.

## References

1. 1.

- Martin, S. et al. Normal-state transport properties of  $\text{Bi}_2\text{Sr}_2\text{CuO}_{6+\delta}$  crystals. *Phys. Rev. B* **41**, 846–849 (1990).

[ADS](#) [CAS](#) [Google Scholar](#)

2. 2.

Custers, J. et al. The break-up of heavy electrons at a quantum critical point. *Nature* **424**, 524–527 (2003).

[ADS](#) [CAS](#) [Google Scholar](#)

3. 3.

Cooper, R. A. et al. Anomalous criticality in the electrical resistivity of  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . *Science* **323**, 603–607 (2009).

[ADS](#) [CAS](#) [Google Scholar](#)

4. 4.

Bruin, J. A. N., Sakai, H., Perry, R. S. & Mackenzie, A. P. Similarity of scattering rates in metals showing  $T$ -linear resistivity. *Science* **339**, 804–807 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Legros, A. et al. Universal  $T$ -linear resistivity and Planckian dissipation in overdoped cuprates. *Nat. Phys.* **15**, 142–147 (2019).

[CAS](#) [Google Scholar](#)

6. 6.

Licciardello, S. et al. Electrical resistivity across a nematic quantum critical point. *Nature* **567**, 213–217 (2019).

[ADS](#) [CAS](#) [Google Scholar](#)

7. 7.

Chien, T., Wang, Z. & Ong, N. Effect of Zn impurities on the normal-state Hall angle in single crystal  $\text{YBa}_2\text{Cu}_{3-x}\text{Zn}_x\text{O}_{7-\delta}$ . *Phys. Rev. Lett.* **67**, 2088–2091 (1991).

[ADS](#) [CAS](#) [Google Scholar](#)

8. 8.

Nakajima, Y. et al. Non-Fermi-liquid behavior in the magnetotransport of  $\text{CeMIn}_5$  (M: Co and Rh): striking similarity between quasi-two-dimensional heavy fermion and high- $T_c$  cuprates. *J. Phys. Soc. Jpn.* **76**, 024703 (2007).

[ADS](#) [Google Scholar](#)

9. 9.

Liu, R. H. et al. Anomalous transport properties and phase diagram of the FeAs-based  $\text{SmFeAsO}_{1-x}\text{F}_x$  superconductors. *Phys. Rev. Lett.* **101**, 087001 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

10. 10.

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

[Google Scholar](#)

11. 11.

Sarkar, T., Mandal, P. R., Poniatowski, N. R., Chan, M. K. & Greene, R. L. Correlation between scale-invariant normal-state resistivity and superconductivity in an electron-doped cuprate. *Sci. Adv.* **5**, eaav6753 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Giraldo-Gallo, P. et al. Scale-invariant magnetoresistance in a cuprate superconductor. *Science* **361**, 479–481 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

13. 13.

Licciardello, S. et al. Coexistence of orbital and quantum critical magnetoresistance in  $\text{FeSe}_{1-x}\text{S}_x$ . *Phys. Rev. Res* **1**, 023011 (2019).

[CAS](#) [Google Scholar](#)

14. 14.

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

[ADS](#) [CAS](#) [Google Scholar](#)

15. 15.

Michon, B. et al. Thermodynamic signatures of quantum criticality in cuprate superconductors. *Nature* **567**, 218–222 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Emery, V. J. & Kivelson, S. A. Superconductivity in bad metals. *Phys. Rev. Lett.* **74**, 3253–3256 (1995).

[ADS](#) [CAS](#) [Google Scholar](#)

17. 17.

van der Marel, D. et al. Quantum critical behaviour in a high- $T_c$  superconductor. *Nature* **425**, 271–274 (2003).

[ADS](#) [Google Scholar](#)

18. 18.

Zaanen, J. Why the temperature is high. *Nature* **430**, 512–513 (2004).

[ADS](#) [CAS](#) [Google Scholar](#)

19. 19.

Hartnoll, S. A. Theory of universal incoherent metallic transport. *Nat. Phys.* **11**, 54–61 (2015).

[CAS](#) [Google Scholar](#)

20. 20.

Zaanen, J. Planckian dissipation, minimal viscosity and the transport in cuprate strange metals. *SciPost Phys.* **6**, 061 (2019).

[ADS](#) [MathSciNet](#) [CAS](#) [Google Scholar](#)

21. 21.

Chen, S.-D. et al. Incoherent strange metal sharply bounded by a critical doping in Bi2212. *Science* **366**, 1099–1102 (2019).

[ADS](#) [CAS](#) [Google Scholar](#)

22. 22.

Hussey, N. E., Buhot, J. & Licciardello, S. A tale of two metals: contrasting criticalities in the pnictides and hole-doped cuprates. *Rep. Prog. Phys.* **81**, 052501 (2018).

[ADS](#) [MathSciNet](#) [CAS](#) [Google Scholar](#)

23. 23.

Tallon, J. L., Storey, J. G., Cooper, J. R. & Loram, J. W. Locating the pseudogap closing point in cuprate superconductors: absence of entrant or reentrant behavior. *Phys. Rev. B* **101**, 174512 (2020).

[ADS](#) [CAS](#) [Google Scholar](#)

24. 24.

Hussey, N. E., Gordon-Moys, H., Kokalj, J. & McKenzie, R. H. Generic strange-metal behaviour of overdoped cuprates. *J. Phys. Conf. Ser.* **449**, 012004 (2013).

[CAS](#) [Google Scholar](#)

25. 25.

Putzke, C. et al. *Reduced Hall carrier density in the overdoped strange metal regime of cuprate super-conductors*. *Nat. Phys.* <https://doi.org/10.1038/s41567-021-01197-0> (2021).

26. 26.

Ando, Y. et al. Evolution of the Hall coefficient and the peculiar electronic structure of the cuprate superconductors. *Phys. Rev. Lett.* **92**, 197001 (2004).

[ADS](#) [Google Scholar](#)

27. 27.

Božović, I., He, X., Wu, J. & Bollinger, A. T. Dependence of the critical temperature in overdoped copper oxides on superfluid density. *Nature* **536**, 309–311 (2016).

[ADS](#) [Google Scholar](#)

28. 28.

McKenzie, R. H. et al. Violation of Kohler's rule by the magnetoresistance of a quasi-two-dimensional organic metal. *Phys. Rev. B* **57**, 11854–11857 (1998).

[ADS](#) [CAS](#) [Google Scholar](#)

29. 29.

Kiritsis, E. & Li, L. Quantum criticality and DBI magneto-resistance. *J. Phys. A* **50**, 115402 (2017).

[ADS](#) [MathSciNet](#) [MATH](#) [Google Scholar](#)

30. 30.

Patel, A. A., McGreevy, J., Arovas, D. P. & Sachdev, S. Magnetotransport in a model of a disordered strange metal. *Phys. Rev. X* **8**, 021049 (2018).

[CAS](#) [Google Scholar](#)

31. 31.

Boyd, C. & Phillips, P. W. Single-parameter scaling in the magnetoresistance of optimally doped  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . *Phys. Rev. B* **100**, 155139 (2019).

[ADS](#) [CAS](#) [Google Scholar](#)

32. 32.

Singleton, J. Temperature scaling behavior of the linear magnetoresistance observed in high-temperature superconductors. *Phys. Rev. Mater.* **4**, 061801 (2020).

[CAS](#) [Google Scholar](#)

33. 33.

Bangura, A. F. et al. Fermi surface and electronic homogeneity of the overdoped cuprate superconductor  $Tl_2Ba_2CuO_{6+\delta}$  as revealed by quantum oscillations. *Phys. Rev. B* **82**, 140501 (2010).

[ADS](#) [Google Scholar](#)

34. 34.

Wise, W. D. et al. Imaging nanoscale Fermi-surface variations in an inhomogeneous superconductor. *Nat. Phys.* **5**, 213–216 (2009).

[CAS](#) [Google Scholar](#)

35. 35.

Hayes, I. M. et al. Magnetoresistance scaling reveals symmetries of the strongly correlated dynamics in  $BaFe_2(As_{1-x}P_x)_2$ . *Phys. Rev. Lett.* **121**, 197002 (2018).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Hussey, N. E., Abdel-Jawad, M., Carrington, A., Mackenzie, A. P. & Balicas, L. A coherent three-dimensional Fermi surface in a high-transition temperature superconductor. *Nature* **425**, 814–817 (2003).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Platé, M. et al. Fermi surface and quasiparticle excitations of overdoped  $Tl_2Ba_2CuO_{6+\delta}$ . *Phys. Rev. Lett.* **95**, 077001 (2005).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

38. 38.

Abdel-Jawad, M. et al. Anisotropic scattering and anomalous normal-state transport in a high-temperature superconductor. *Nat. Phys.* **2**, 821–825 (2006).

[CAS](#) [Google Scholar](#)

39. 39.

Hayes, I. M. et al. Superconductivity and quantum criticality linked by the Hall effect in a strange metal. *Nat. Phys.* **17**, 58–62 (2021).

[CAS](#) [Google Scholar](#)

40. 40.

Knolle, J. & Cooper, N. R. Anomalous de Haas–van Alphen effect in InAs/GaSb quantum wells. *Phys. Rev. Lett.* **118**, 176801 (2017).

[ADS](#) [Google Scholar](#)

41. 41.

Tyler, A. W. *An Investigation into the Magnetotransport Properties of Layered Superconducting Perovskites*. PhD thesis, Univ. Cambridge (1997).

42. 42.

Presland, M. R., Tallon, J. L., Buckley, R. G., Liu, R. S. & Flower, N. E. General trends in oxygen stoichiometry effects on  $T_c$  in Bi and Tl superconductors. *Physica C* **176**, 95–105 (1991).

[ADS](#) [CAS](#) [Google Scholar](#)

43. 43.

Ono, S. & Ando, Y. Evolution of the resistivity anisotropy in  $\text{Bi}_2\text{Sr}_{2-x}\text{La}_x\text{CuO}_{6+\delta}$  single crystals for a wide range of hole doping. *Phys. Rev. B* **67**, 104512 (2003).

[ADS](#) [Google Scholar](#)

44. 44.

Shibauchi, T. et al. Field-induced quantum critical route to a Fermi liquid in high-temperature superconductors. *Proc. Natl Acad. Sci. USA* **105**, 7120–7123 (2008).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

45. 45.

French, M. M. J. & Hussey, N. E. Orbital origin of field-induced quantum criticality in overdoped  $Tl_2Ba_2CuO_{6+x}$ . *Proc. Natl Acad. Sci. USA* **105**, E58 (2008).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Mackenzie, A. P., Julian, S. R., Sinclair, D. C. & Lin, C. T. Normal-state magnetotransport in superconducting  $Tl_2Ba_2CuO_{6+\delta}$  to millikelvin temperatures. *Phys. Rev. B* **53**, 5848–5855 (1996).

[ADS](#) [CAS](#) [Google Scholar](#)

47. 47.

Hussey, N. E. et al. Angular dependence of the *c*-axis normal state magnetoresistance in single crystal  $Tl_2Ba_2CuO_{6+\delta}$ . *Phys. Rev. Lett.* **76**, 122–125 (1996).

[ADS](#) [CAS](#) [Google Scholar](#)

48. 48.

French, M. M. J., Analytis, J. G., Carrington, A., Balicas, L. & Hussey, N. E. Tracking anisotropic scattering in overdoped  $Tl_2Ba_2CuO_{6+\delta}$

above 100 K. *New J. Phys.* **11**, 055057 (2009).

[ADS](#) [Google Scholar](#)

49. 49.

Rourke, P. M. C. et al. A detailed de Haas–van Alphen effect study of the overdoped cuprate  $Tl_2Ba_2CuO_{6+\delta}$ . *New J. Phys.* **12**, 105009 (2010).

[ADS](#) [Google Scholar](#)

50. 50.

Ong, N. P. Geometric interpretation of the weak-field Hall conductivity in two-dimensional metals with arbitrary Fermi surface. *Phys. Rev. B* **43**, 193–201 (1991).

[ADS](#) [CAS](#) [Google Scholar](#)

51. 51.

Pippard, A. B. *Magnetoresistance in Metals* (Cambridge Univ. Press, 1989).

52. 52.

Clarke, D. C., Strong, S. P. & Anderson, P. W. Conductivity between Luttinger liquids in the confinement regime and *c*-axis conductivity in the cuprate superconductors. *Phys. Rev. Lett.* **74**, 4499–4502 (1995).

[ADS](#) [CAS](#) [Google Scholar](#)

53. 53.

Grissonnanche, G. et al. Measurement of the Planckian scattering rate. Preprint at <https://arxiv.org/abs/2011.13054> (2020).

54. 54.

Kondo, T., Takeuchi, T., Tsuda, S. & Shin, S. Electrical resistivity and scattering processes in  $(\text{Bi},\text{Pb})_2(\text{Sr},\text{La})_2\text{CuO}_{6+\delta}$  studied by angle-resolved photoemission spectroscopy. *Phys. Rev. B* **74**, 224511 (2006).

[ADS](#) [Google Scholar](#)

55. 55.

Doiron-Leyraud, N. et al. Quantum oscillations and the Fermi surface in an underdoped high- $T_c$  superconductor. *Nature* **447**, 565–568 (2007).

[ADS](#) [CAS](#) [Google Scholar](#)

56. 56.

Bangura, A. F. et al. Small Fermi surface pockets in underdoped high-temperature superconductors: observation of Shubnikov–de Haas oscillations in  $\text{YBa}_2\text{Cu}_4\text{O}_8$ . *Phys. Rev. Lett.* **100**, 047004 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

57. 57.

Barišić, N. et al. Universal quantum oscillations in the underdoped cuprate superconductors. *Nat. Phys.* **9**, 761–764 (2013).

[Google Scholar](#)

58. 58.

Knolle, J. & Cooper, N. R. Excitons in topological Kondo insulators: theory of thermodynamic and transport anomalies in  $\text{SmB}_6$ . *Phys. Rev. Lett.* **118**, 176801 (2017).

[ADS](#) [Google Scholar](#)

59. 59.

Li, G. et al. Two-dimensional Fermi surfaces in Kondo insulator SmB<sub>6</sub>. *Science* **346**, 1208–1212 (2014).

[ADS](#) [CAS](#) [Google Scholar](#)

60. 60.

Tan, B. S. et al. Unconventional Fermi surface in an insulating state. *Science* **349**, 287–290 (2015).

[ADS](#) [CAS](#) [Google Scholar](#)

61. 61.

Xiang, Z. et al. Quantum oscillations of electrical resistivity in an insulator. *Science* **362**, 65–69 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

62. 62.

Yasui, K. & Kita, T. Theory of the de Haas–van Alphen effect in type-II superconductors. *Phys. Rev. B* **66**, 184516 (2002).

[ADS](#) [Google Scholar](#)

63. 63.

Hartnoll, S. A. & Hofman, D. M. Generalized Lifshitz–Kosevich scaling at quantum criticality from the holographic correspondence. *Phys. Rev. B* **81**, 155125 (2010).

[ADS](#) [Google Scholar](#)

64. 64.

Chan, M. K. et al. In-plane magnetoresistance obeys Kohler’s rule in the pseudogap phase of cuprate superconductors. *Phys. Rev. Lett.* **113**, 177005 (2014).

[ADS](#) [CAS](#) [Google Scholar](#)

65. 65.

Harris, J. M. et al. Violation of Kohler's rule in the normal-state magnetoresistance of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  and  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . *Phys. Rev. Lett.* **75**, 1391–1394 (1995).

[ADS](#) [CAS](#) [Google Scholar](#)

66. 66.

Mirzaei, S. I. et al. Spectroscopic evidence for Fermi liquid-like energy and temperature dependence of the relaxation rate in the pseudogap phase of the cuprates. *Proc. Natl Acad. Sci. USA* **110**, 5774–5778 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

67. 67.

Proust, C., Vignolle, B., Levallois, J., Adachi, S. & Hussey, N. E. Fermi liquid behavior of the in-plane resistivity in the pseudogap state of  $\text{YBa}_2\text{Cu}_4\text{O}_8$ . *Proc. Natl Acad. Sci. USA* **113**, 13654–13659 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

68. 68.

Ayres, J. *Correlated Electron Systems Under Extreme Conditions: High Fields, High Pressures, Low Temperatures*. PhD thesis, Univ. Bristol (2020); <https://research-information.bris.ac.uk/en/studentTheses/correlated-electron-systems-under-extreme-conditions>.

69. 69.

Gotlieb, K. et al. Revealing hidden spin–momentum locking in a high-temperature cuprate superconductor. *Science* **362**, 1271–1275 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

70. 70.

Fuseya, Y. et al. Origin of the large anisotropic  $g$  factor of holes in bismuth. *Phys. Rev. Lett.* **115**, 216401 (2015).

[ADS](#) [Google Scholar](#)

71. 71.

Ma, M. et al. Prominent role of spin-orbit coupling in FeSe revealed by inelastic neutron scattering. *Phys. Rev. X* **7**, 021025 (2017).

[Google Scholar](#)

72. 72.

Moses, P. & McKenzie, R. H. Comparison of coherent and weakly incoherent transport models for the interlayer magnetoresistance of layered Fermi liquids. *Phys. Rev. B* **60**, 7998–8011 (1999).

[ADS](#) [CAS](#) [Google Scholar](#)

73. 73.

Sandeman, K. & Schofield, A. J. Model of anisotropic scattering in a quasi-two-dimensional metal. *Phys. Rev. B* **63**, 094510 (2001).

[ADS](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank M. Allan, J. G. Analytis, I. Božović, M. S. Golden, B. Goutéraux, C. Pépin, K. Schalm, H. Stoof and S. Vandoren for insightful discussions during the course of this work. We also thank S. Smit and L. Bawden for initial characterization of some of the Bi2201 single crystals and L. Malone for assistance in the growth of the Tl2201 single crystals. J.A. acknowledges the support of the EPSRC-funded CMP-CDT (ref. EP/L015544/1) and an EPSRC Doctoral Prize Fellowship (ref. EP/T517872/1). A.C. acknowledges support of the EPSRC (ref. EP/R011141/1). We also acknowledge the support of the High Field Magnet Laboratory (HFML) at Radboud University, member of the European Magnetic Field Laboratory (EMFL – also supported by the EPSRC, ref. EP/N01085X/1), and the former Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organisation for Scientific Research (NWO) (grant no. 16METL01, ‘Strange Metals’). Finally, part of this work was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement nos 835279-Catch-22 and 715262-HPSuper).

## Author information

### Author notes

1. M. Čulo

Present address: Institut za fiziku, PO Box 304, Zagreb, Croatia

2. These authors contributed equally: J. Ayres, M. Berben

### Affiliations

1. H. H. Wills Physics Laboratory, University of Bristol, Bristol, UK

J. Ayres, C. Putzke, S. Friedemann, A. Carrington & N. E. Hussey

2. High Field Magnet Laboratory (HFML-EMFL), Radboud University, Nijmegen, Netherlands

J. Ayres, M. Berben, M. Čulo, Y.-T. Hsu & N. E. Hussey

3. Institute for Molecules and Materials, Radboud University, Nijmegen, Netherlands

J. Ayres, M. Berben, M. Čulo, Y.-T. Hsu & N. E. Hussey

4. Van der Waals–Zeeman Institute, University of Amsterdam, Amsterdam, Netherlands

E. van Heumen & Y. Huang

5. QSoft, Science, Park 123, Amsterdam, Netherlands

E. van Heumen

6. Instituut-Lorentz for Theoretical Physics, Leiden University, Leiden, Netherlands

J. Zaanen

7. Institute for Solid State Physics, University of Tokyo, Kashiwa, Japan

T. Kondo

8. Toyota Technological Institute, Nagoya, Japan

T. Takeuchi

9. Department of Physics, University of Cambridge, Cambridge, UK

J. R. Cooper

## Authors

1. J. Ayres

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. M. Berben

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. M. Čulo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Y.-T. Hsu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. E. van Heumen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Y. Huang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. J. Zaanen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. T. Kondo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. T. Takeuchi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. J. R. Cooper

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. C. Putzke

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. S. Friedemann

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. A. Carrington

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. N. E. Hussey

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

J.A., M.B., S.F., A.C. and N.E.H. conceived the overall project. J.A., M.B., M.Č., Y.-T.H., C.P. and N.E.H. performed the high-field measurements. Y.H., E.v.H., J.R.C., C.P., T.K. and T.T. grew and characterized the single-crystal samples. J.A. and A.C. performed the SCTIF calculations. J.A., M.B., J.Z. and N.E.H. wrote the manuscript with input from all of the co-authors.

## Corresponding authors

Correspondence to [J. Ayres](#) or [M. Berben](#) or [N. E. Hussey](#).

# Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Nicholas Breznay 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 Zero-field resistivities of Tl2201 and Bi2201.](#)

**a, b**, Zero-field, ambient-pressure resistivity  $\rho_{ab}(T)$  curves for representative Tl2201 (**a**) and Bi2201 (**b**) crystals investigated in this study. Note the super-linear  $T$  dependence for all samples. The spread in absolute magnitudes of  $\rho_{ab}(T)$  is higher in the Tl2201 crystals owing to the fact that they were mounted for pressure measurements and as such, their absolute resistivities were harder to quantify accurately.

### [Extended Data Fig. 2 Quadrature scaling in overdoped Tl2201.](#)

**a**,  $\rho_{ab}(H, T)$  as measured in Tl2201 with  $T_c = 26.5$  K.  $H^2$  behaviour cedes to a  $H$ -linear resistivity at high fields. **b, c**, Scaling plots of  $[\rho_{ab}(H, T) - \rho_{ab}(0, 0)]/T$  versus  $H/T$  for overdoped Tl2201 ( $T_c = 26.5$  K). As shown in **c**, there is a clear breakdown of the scaling at low  $H/T$ . **d, e**, Scaling plots of  $[\rho_{ab}(H, T) - \rho_{ab}(0, T)]/T$  versus  $H/T$  for the same sample where  $\rho(0, T) =$

$\mathcal{F}(T) = \rho_0 + A_g T + BT^2$ . Note that  $A_g$  does not correspond to  $A$ , the full  $T$ -linear coefficient of the zero-field resistivity, since part of that is contained within the quadrature form. The inclusion of these additional  $T$ -dependent terms makes the data collapse over the full range of  $T$ . Taking the derivative with respect to  $H$  (as done in the main text) provides another means of isolating the quadrature MR from  $\mathcal{F}(T)$ . The dashed lines in all panels represent the quadrature expression  $\langle(\Delta \{\rho\}_{ab}(H)=\alpha \{k\}_{\{\rm rm\} B}\} T\sqrt{1+\{\beta \{\mu\}_0 H/T\}^2}\rangle$  ( $\rho_0 = 15.5 \mu\Omega \text{ cm}$ ,  $A_g = 0.14 \mu\Omega \text{ cm K}^{-1}$ ,  $B = 0.003 \mu\Omega \text{ cm K}^{-2}$ ,  $\alpha k_B = 0.04 \mu\Omega \text{ cm K}^{-1}$ ,  $\gamma \mu_B = 0.20 \mu\Omega \text{ cm T}^{-1}$ ). **f**, The derivatives with respect to magnetic field of the measured curves shown in **a**. **g**, When plotted against  $H/T$ , the derivatives presented in **f** collapse onto a universal curve (with the exception of those sections of each field sweep that are in the mixed state).

### Extended Data Fig. 3 Success of ADMR-derived modelling of the in-plane transport of overdoped Tl2201.

**a**, The  $c$ -axis ADMR of Tl2201 with  $T_c = 15$  K measured at 50 K and at various (labelled) azimuthal angles taken from ref. <sup>48</sup>. **b**, Projection of the in-plane Fermi surface derived from the ADMR fitting. **c**, Schematic showing the isotropic  $T^2$  component (black solid line) and anisotropic  $T + T^2$  component (red solid line) of the scattering rate as deduced from the ADMR fitting. **d**, Black dots:  $\rho_{ab}(T)$  data for overdoped Tl2201 ( $T_c = 15$  K) in which superconductivity has been suppressed by a magnetic field ( $\mathbf{H} \parallel c$ )<sup>46</sup> and corresponding simulation based on the ADMR fitting<sup>48</sup>. The difference in the residual resistivities is probably because different samples have been used in the two studies<sup>46,48</sup>. **e**, Corresponding simulation for  $R_H(T)$ <sup>48</sup>. **f**, Simulation of  $R_H(H) = \rho_{xy}(H)/H$  at various temperatures as indicated. **g**, Same simulation data plotted versus  $H/\rho(0)$  where here,  $\rho(0)$  is the zero-field resistivity at each temperature. **h**,  $R_H(H)$  versus  $H/\rho(0)$  data taken from ref. <sup>25</sup>. For overdoped Tl2201 ( $T_c = 25$  K) for comparison with the simulation in **g**. The larger absolute values of  $R_H$  in **h** relative to **g** are due to the fact that the high-field data in **h** are taken on a sample with a higher  $T_c$  value where the anisotropy in  $\tau^{-1}(\phi)$  is expected to be larger.

## Extended Data Fig. 4 Failure of ADMR-derived modelling to reproduce quadrature scaling.

**a**, The field dependence of the longitudinal resistivity ( $\rho(T)$ ) determined with the SCTIF using parameters derived from the ADMR parameterization for overdoped Tl2201. **b**,  $\Delta\rho$ —the change in  $\rho_{xx}$  with field—at selected temperatures. **c**, Corresponding derivative plots of  $\rho(H)$  showing distinctly non-quadrature behaviour. **d**, As a consequence, the data fail to collapse when plotted against  $\mu_0H/T$ .

## Extended Data Fig. 5 Failure of the SCTIF to reproduce both the MR and Hall response.

Simulations of the MR and Hall responses within the SCTIF given different parameterizations of  $v_F(\phi)$  and  $\tau^{-1}(\phi, T)$ . In each simulation, the experimentally determined Fermi surface ( $k_F(\phi)$ ) has been used. Note that the SCTIF is slow to converge at low fields and so the simulations do not extend all the way to  $H = 0$ . Simulation 1, ADMR-derived parameterization of overdoped Tl2201 albeit with no anisotropy at  $T = 0$  and an anisotropic term that increases strictly linearly with  $T$ . Simulation 2, A scenario incorporating the  $v_F$  anisotropy derived from tight-binding modelling of ARPES measurements<sup>37</sup>. Simulation 3, A scenario in which the anisotropy ratio of  $\tau^{-1}(\phi)$  is strictly  $T$ -independent in order to generate an MR with a maximum slope that is also independent of temperature (reminiscent of quadrature scaling). Simulation 4, A scenario in which a similar  $\tau^{-1}(\phi)$  parameterization to that used to model Nd-LSCO<sup>53</sup> is applied to overdoped Tl2201. Simulation 5, Simulation for overdoped Bi2201 with an enhanced anisotropy in  $v_F$  and  $\tau^{-1}(T = 0)$  consistent with ARPES<sup>54</sup>. Column 1, The Fermi surface parameterizations  $k_F(\phi)$  and  $v_F(\phi)$ . Column 2,  $\tau^{-1}(\phi, T)$ . Column 3,  $d\rho/d(\mu_0H)$  versus  $H$ . Column 4,  $d\rho/d(\mu_0H)$  versus  $H/T$ . Column 5,  $R_H(H, T)$ .

## Extended Data Fig. 6 Kohler versus quadrature scaling in Bi2201.

**a**,  $\Delta\rho_{ab}/\rho_{ab}(0)$  plotted versus  $(H/\rho_{ab}(0))^2$  for a Bi2201 sample with  $T_c = 13$  K. In a system that shows Kohler scaling, these curves would collapse. Clearly, that is not the case here. **b**,  $\Delta\rho_{ab}$  plotted versus  $H^2$  for the same Bi2201 sample. The dotted lines are fits to the function  $f(x) = A(\mu_0 H)^2$  in the regions where the MR is strictly quadratic. Note that the quadrature form of the MR is only purely quadratic in the zero-field limit whereas fits to the data are taken at finite field ranges. Our simulations have shown that fitting up to  $\mu_0 H_{\max} = \beta T$  (with  $\beta$  as given in Fig. 2 of the main text) agrees with the zero-field limit within a few percent and falls within our experimental error. **c**,  $T$  dependence of  $A\rho_{ab}(0)$  (with  $A$  taken from the fits in b) compared to the square of the Hall coefficient  $\langle\{R\}\rangle_{\{\{rm{H}\}\}}^2$ . **d**, Temperature derivative of  $\rho_{ab}(T)$  for the same sample. Note that the onset of superconducting fluctuations appears only below 30 K. **e**, The product  $AT$  plotted over the full temperature range. The dotted line is a guide to the eye.

### Extended Data Table 1 TI2201 samples studied

[Full size table](#)

### Extended Data Table 2 Bi2201 samples studied

[Full size table](#)

### Extended Data Table 3 ADMR simulation parameters

[Full size table](#)

## Supplementary information

[Peer Review File](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Ayres, J., Berben, M., Čulo, M. *et al.* Incoherent transport across the strange-metal regime of overdoped cuprates. *Nature* **595**, 661–666 (2021). <https://doi.org/10.1038/s41586-021-03622-z>

### [Download citation](#)

- Received: 17 November 2020
- Accepted: 04 May 2021
- Published: 28 July 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03622-z>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Advertisement

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

- Article
- [Published: 28 July 2021](#)

# Linear-in temperature resistivity from an isotropic Planckian scattering rate

- [Gaël Grissonnanche](#) [ORCID: orcid.org/0000-0002-5115-3125](#)<sup>1,2,3</sup>,
- [Yawen Fang](#) [ORCID: orcid.org/0000-0001-6964-5817](#)<sup>2</sup>,
- [Anaëlle Legros](#) [ORCID: orcid.org/0000-0002-7938-8807](#)<sup>1,4</sup>,
- [Simon Verret](#)<sup>1</sup>,
- [Francis Laliberté](#) [ORCID: orcid.org/0000-0002-9952-6330](#)<sup>1</sup>,
- [Clément Collignon](#) [ORCID: orcid.org/0000-0002-1697-4331](#)<sup>1</sup>,
- [Jianshi Zhou](#)<sup>5</sup>,
- [David Graf](#) [ORCID: orcid.org/0000-0001-6195-0462](#)<sup>6</sup>,
- [Paul A. Goddard](#)<sup>7</sup>,
- [Louis Taillefer](#) [ORCID: orcid.org/0000-0002-8729-2194](#)<sup>1,8</sup> &
- [B. J. Ramshaw](#) [ORCID: orcid.org/0000-0002-3222-5007](#)<sup>2,8</sup>

[Nature](#) volume **595**, pages 667–672 (2021) [Cite this article](#)

- 961 Accesses
- 58 Altmetric
- [Metrics details](#)

## Subjects

- [Electronic properties and materials](#)
- [Phase transitions and critical phenomena](#)

# Abstract

A variety of ‘strange metals’ exhibit resistivity that decreases linearly with temperature as the temperature decreases to zero<sup>1,2,3</sup>, in contrast to conventional metals where resistivity decreases quadratically with temperature. This linear-in-temperature resistivity has been attributed to charge carriers scattering at a rate given by  $\hbar/\tau = \alpha k_B T$ , where  $\alpha$  is a constant of order unity,  $\hbar$  is the Planck constant and  $k_B$  is the Boltzmann constant. This simple relationship between the scattering rate and temperature is observed across a wide variety of materials, suggesting a fundamental upper limit on scattering—the ‘Planckian limit’<sup>4,5</sup>—but little is known about the underlying origins of this limit. Here we report a measurement of the angle-dependent magnetoresistance of  $\text{La}_{1.6-x}\text{Nd}_{0.4}\text{Sr}_x\text{CuO}_4$ —a hole-doped cuprate that shows linear-in-temperature resistivity down to the lowest measured temperatures<sup>6</sup>. The angle-dependent magnetoresistance shows a well defined Fermi surface that agrees quantitatively with angle-resolved photoemission spectroscopy measurements<sup>7</sup> and reveals a linear-in-temperature scattering rate that saturates at the Planckian limit, namely  $\alpha = 1.2 \pm 0.4$ . Remarkably, we find that this Planckian scattering rate is isotropic, that is, it is independent of direction, in contrast to expectations from ‘hotspot’ models<sup>8,9</sup>. Our findings suggest that linear-in-temperature resistivity in strange metals emerges from a momentum-independent inelastic scattering rate that reaches the Planckian limit.

## 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

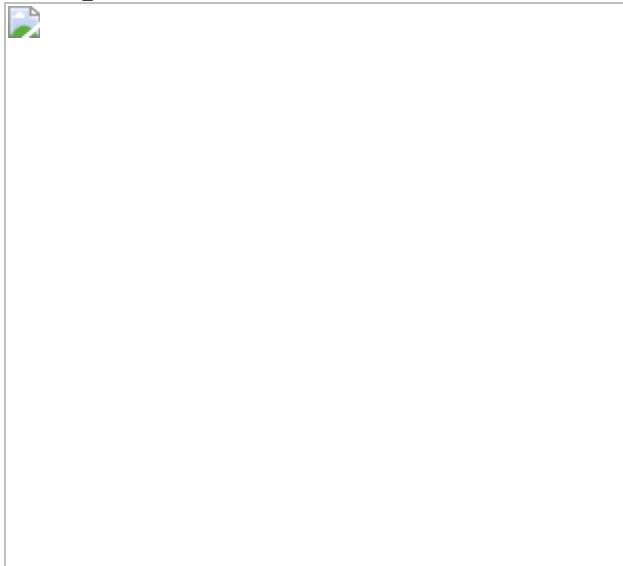
[Rent or Buy](#)

All prices are NET prices.

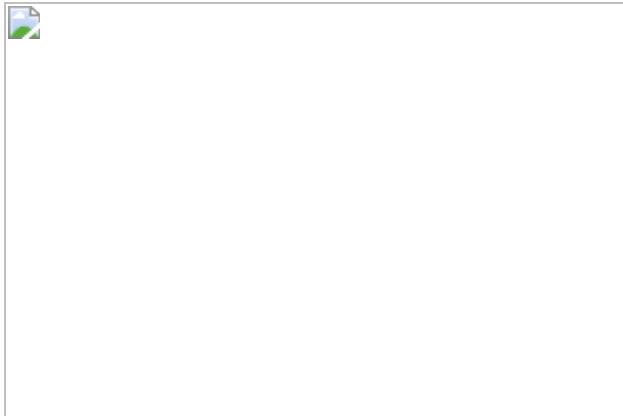
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

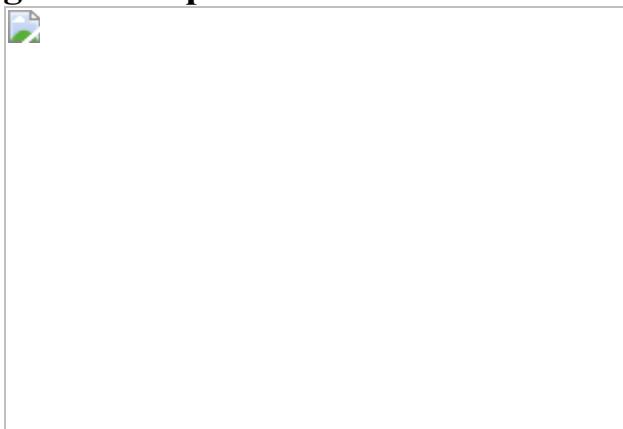
**Fig. 1: *T*-linear resistivity and the angle-dependent magnetoresistance technique.**



**Fig. 2: ADMR and quasiparticle scattering rate of Nd-LSCO at  $p = 0.24$ .**



**Fig. 3: Transport coefficients of Nd-LSCO at  $p=0.24$ .**



**Fig. 4: Comparison of two overdoped cuprates: Nd-LSCO and Tl2201.**

## Data availability

The experimental data presented in this paper are available at <http://wrap.warwick.ac.uk/152398/>. The results of the conductivity simulations are available from the corresponding authors upon reasonable request.

# Code availability

The code used to compute the conductivity is available from the corresponding authors upon reasonable request.

## References

1. 1.

Martin, S., Fiory, A. T., Fleming, R. M., Schneemeyer, L. F. & Waszczak, J. V. Normal-state transport properties of  $\text{Bi}_{2+x}\text{Sr}_{2-y}\text{CuO}_{6+\delta}$  crystals. *Phys. Rev. B* **41**, 846–849 (1990).

[ADS](#) [CAS](#) [Google Scholar](#)

2. 2.

Löhneysen, H. V. et al. Non-Fermi-liquid behavior in a heavy-fermion alloy at a magnetic instability. *Phys. Rev. Lett.* **72**, 3262–3265 (1994).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Doiron-Leyraud, N. et al. Correlation between linear resistivity and  $T_c$  in the Bechgaard salts and the pnictide superconductor  $\text{Ba}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$ . *Phys. Rev. B* **80**, 214531 (2009).

[ADS](#) [Google Scholar](#)

4. 4.

Bruin, J. A. N., Sakai, H., Perry, R. S. & Mackenzie, A. P. Similarity of scattering rates in metals showing  $T$ -linear resistivity. *Science* **339**, 804–807 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Legros, A. et al. Universal  $T$ -linear resistivity and Planckian dissipation in overdoped cuprates. *Nat. Phys.* **15**, 142–147 (2019).

[CAS](#) [Google Scholar](#)

6. 6.

Daou, R. et al. Linear temperature dependence of resistivity and change in the Fermi surface at the pseudogap critical point of a high- $T_c$  superconductor. *Nat. Phys.* **5**, 31–34 (2009).

[CAS](#) [Google Scholar](#)

7. 7.

Matt C. E. et al. Electron scattering, charge order, and pseudogap physics in  $\text{La}_{1.6-x}\text{Nd}_{0.4}\text{Sr}_x\text{CuO}_4$ : an angle-resolved photoemission spectroscopy study. *Phys. Rev. B* **92**, 134524 (2015).

[ADS](#) [Google Scholar](#)

8. 8.

Hlubina, R. & Rice, T. M. Resistivity as a function of temperature for models with hot spots on the Fermi surface. *Phys. Rev. B* **51**, 9253–9260 (1995).

[ADS](#) [CAS](#) [Google Scholar](#)

9. 9.

Stojković, B. P. & Pines, D. Theory of the longitudinal and Hall conductivities of the cuprate superconductors. *Phys. Rev. B* **55**, 8576–8595 (1997).

[ADS](#) [Google Scholar](#)

10. 10.

Gurvitch, M. & Fiory, A. T. Resistivity of  $\text{La}_{1.825}\text{Sr}_{0.175}\text{CuO}_4$  and  $\text{YBa}_2\text{Cu}_3\text{O}_7$  to 1100 K: absence of saturation and its implications. *Phys. Rev. Lett.* **59**, 1337–1340 (1987).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

11. 11.

Varma, C. M., Littlewood, P. B., Schmitt-Rink, S., Abrahams, E. & Ruckenstein, A. E. Phenomenology of the normal state of Cu-O high-temperature superconductors. *Phys. Rev. Lett.* **63**, 1996–1999 (1989).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Cao, Y. et al. Strange metal in magic-angle graphene with near Planckian dissipation. *Phys. Rev. Lett.* **124**, 076801 (2020).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

13. 13.

Parcollet, O. & Georges, A. Non-Fermi-liquid regime of a doped Mott insulator. *Phys. Rev. B* **59**, 5341–5360 (1999).

[ADS](#) [CAS](#) [Google Scholar](#)

14. 14.

Davison, R. A., Schalm, K. & Zaanen, J. Holographic duality and the resistivity of strange metals. *Phys. Rev. B* **89**, 245116 (2014).

[ADS](#) [Google Scholar](#)

15. 15.

Hartnoll, S. A. Theory of universal incoherent metallic transport. *Nat. Phys.* **11**, 54–61 (2015).

[CAS](#) [Google Scholar](#)

16. 16.

Patel, A. A. & Sachdev, S. Theory of a Planckian metal. *Phys. Rev. Lett.* **123**, 066601 (2019).

[ADS](#) [MathSciNet](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

17. 17.

Cha, P., Wentzell, N., Parcollet, O., Georges, A. & Kim, E.-A. Linear resistivity and Sachdev–Ye–Kitaev (SYK) spin liquid behavior in a quantum critical metal with spin-1/2 fermions. *Proc. Natl Acad. Sci. USA* **117**, 18341–18346 (2020).

[MathSciNet](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

18. 18.

Corson, J., Orenstein, J., Oh, S., O’Donnell, J. & Eckstein, J. N. Nodal quasiparticle lifetime in the superconducting state of  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ . *Phys. Rev. Lett.* **85**, 2569 (2000).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

19. 19.

Kaminski, A. et al. Momentum anisotropy of the scattering rate in cuprate superconductors. *Phys. Rev. B* **71**, 014517 (2005).

[ADS](#) [Google Scholar](#)

20. 20.

Collignon, C. et al. Fermi-surface transformation across the pseudogap critical point of the cuprate superconductor  $\text{La}_{1.6-x}\text{Nd}_{0.4}\text{Sr}_x\text{CuO}_4$ . *Phys. Rev. B* **95**, 224517 (2017).

[ADS](#) [Google Scholar](#)

21. 21.

Chambers, R. G. The kinetic formulation of conduction problems. *Proc. Phys. Soc. A* **65**, 458–459 (1952).

[ADS](#) [MATH](#) [Google Scholar](#)

22. 22.

Prange, R. E. & Kadanoff, L. P. Transport theory for electron–phonon interactions in metals. *Phys. Rev.* **134**, A566–A580 (1964).

[ADS](#) [MATH](#) [Google Scholar](#)

23. 23.

Abrahams, E. & Varma, C. M. Hall effect in the marginal Fermi liquid regime of high- $T_c$  superconductors. *Phys. Rev. B* **68**, 094502 (2003).

[ADS](#) [Google Scholar](#)

24. 24.

Horio, M. et al. Three-dimensional Fermi surface of overdoped La-based cuprates. *Phys. Rev. Lett.* **121**, 077004 (2018).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

25. 25.

Michon, B. et al. Thermodynamic signatures of quantum criticality in cuprate superconductors. *Nature* **567**, 218–222 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Narduzzo, A. et al. Violation of the isotropic mean free path approximation for overdoped  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . *Phys. Rev. B* **77**, 220502 (2008).

[ADS](#) [Google Scholar](#)

27. 27.

Abrahams, E. & Varma, C. M. What angle-resolved photoemission experiments tell about the microscopic theory for high-temperature superconductors. *Proc. Natl Acad. Sci. USA* **97**, 5714–5716 (2000).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

28. 28.

Chang, J. et al. Anisotropic breakdown of Fermi liquid quasiparticle excitations in overdoped  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . *Nat. Commun.* **4**, 2559 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Pelc, D. et al. Resistivity phase diagram of cuprates revisited. *Phys. Rev. B* **102**, 075114 (2020).

[ADS](#) [CAS](#) [Google Scholar](#)

30. 30.

Kovtun, P. K., Son, D. T. & Starinets, A. O. Viscosity in strongly interacting quantum field theories from black hole physics. *Phys. Rev. Lett.* **94**, 111601 (2005).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

31. 31.

Giraldo-Gallo, P. et al. Scale-invariant magnetoresistance in a cuprate superconductor. *Science* **361**, 479–481 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

32. 32.

Peierls, R. On the theory of the magnetic change in resistance. *Ann. Phys.* **10**, 97–110 (1931).

[Google Scholar](#)

33. 33.

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

[Google Scholar](#)

34. 34.

Nakamae, S. et al. Electronic ground state of heavily overdoped nonsuperconducting  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . *Phys. Rev. B* **68**, 100502 (2003).

[ADS](#) [Google Scholar](#)

35. 35.

Cooper, R. A. et al. Anomalous criticality in the electrical resistivity of  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . *Science* **323**, 603–607 (2009).

[ADS](#) [CAS](#) [Google Scholar](#)

36. 36.

Abdel-Jawad, M. et al. Anisotropic scattering and anomalous normal-state transport in a high-temperature superconductor. *Nat. Phys.* **2**, 821–825 (2006).

[CAS](#) [Google Scholar](#)

37. 37.

Proust, C., Boaknin, E., Hill, R. W., Taillefer, L. & Mackenzie, A. P. Heat transport in a strongly overdoped cuprate: Fermi liquid and a pure *d*-wave BCS superconductor. *Phys. Rev. Lett.* **89**, 147003 (2002).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

38. 38.

Ramshaw, B. J. et al. Broken rotational symmetry on the Fermi surface of a high- $T_c$  superconductor. *npj Quantum Mater.* **2**, 8 (2017).

[ADS](#) [Google Scholar](#)

39. 39.

Newville, M., Stensitzki, T., Allen, D. B. & Ingargiola, A. LMFIT: non-linear least-square minimization and curve-fitting for python. *Zenodo* <https://zenodo.org/record/11813#.YNRimOhKhPY> (2014).

40. 40.

Chakravarty, S., Sudbø, A., Anderson, P. W. & Strong, S. Interlayer tunneling and gap anisotropy in high-temperature superconductors. *Science* **261**, 337–340 (1993).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

41. 41.

Helm, T. *Electronic Properties of Electron-Doped Cuprate Superconductors Probed by High-Field Magnetotransport*. PhD thesis,

Technical Univ. Munich (2013).

42. 42.

Fournier, P. et al. Insulator–metal crossover near optimal doping in  $\text{Pr}_{2-x}\text{Ce}_x\text{CuO}_4$ : anomalous normal-state low temperature resistivity. *Phys. Rev. Lett.* **81**, 4720–4723 (1998).

[ADS](#) [CAS](#) [Google Scholar](#)

43. 43.

Bangura, A. F. et al. Fermi surface and electronic homogeneity of the overdoped cuprate superconductor  $\text{Tl}_2\text{Ba}_2\text{CuO}_{6+\delta}$  as revealed by quantum oscillations. *Phys. Rev. B* **82**, 140501 (2010).

[ADS](#) [Google Scholar](#)

44. 44.

Analytis, J. G., Abdel-Jawad, M., Balicas, L., French, M. M. J. & Hussey, N. E. Angle-dependent magnetoresistance measurements in  $\text{Tl}_2\text{Ba}_2\text{CuO}_{6+\delta}$  and the need for anisotropic scattering. *Phys. Rev. B* **760**, 104523 (2007).

[ADS](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We acknowledge helpful discussions with J. Analytis, D. Chowdhury, N. Doiron-Leyraud, N. Hussey, M. Kartsovnik, S. Kivelson, D.-H. Lee, P. A. Lee, S. Lewin, A. Maharaj, K. Modic, C. Murthy, S. Musser, C. Proust, S. Sachdev, A. Shekhter, S. Todadri, A.-M. Tremblay and C. Varma. A portion of this work was performed at the National High Magnetic Field Laboratory, which is supported by the National Science Foundation

Cooperative Agreement No. DMR-1644779 and the State of Florida. P.A.G. acknowledges that this project is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 681260). J.-S.Z. was supported by an NSF grant (MRSEC DMR-1720595). L.T. acknowledges support from the Canadian Institute for Advanced Research (CIFAR) as a Fellow and funding from the Natural Sciences and Engineering Research Council of Canada (NSERC; PIN: 123817), the Fonds de recherche du Québec - Nature et Technologies (FRQNT), the Canada Foundation for Innovation (CFI) and a Canada Research Chair. This research was undertaken thanks in part to funding from the Canada First Research Excellence Fund. Part of this work was funded by the Gordon and Betty Moore Foundation's EPiQS Initiative (Grant GBMF5306 to L.T.) B.J.R. and Y.F. acknowledge funding from the National Science Foundation under grant no. DMR-1752784.

## Author information

### Affiliations

1. Département de physique, Institut quantique, RQMP, Université de Sherbrooke, Sherbrooke, Québec, Canada

Gaël Grissonnanche, Anaëlle Legros, Simon Verret, Francis Laliberté, Clément Collignon & Louis Taillefer

2. Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY, USA

Gaël Grissonnanche, Yawen Fang & B. J. Ramshaw

3. Kavli Institute at Cornell for Nanoscale Science, Ithaca, NY, USA

Gaël Grissonnanche

4. SPEC, CEA, CNRS-UMR 3680, Université Paris-Saclay, Gif-sur-Yvette, France

Anaëlle Legros

5. Materials Science and Engineering Program, Department of Mechanical Engineering, University of Texas at Austin, Austin, TX, USA

Jianshi Zhou

6. National High Magnetic Field Laboratory, Tallahassee, FL, USA

David Graf

7. Department of Physics, University of Warwick, Coventry, UK

Paul A. Goddard

8. Canadian Institute for Advanced Research, Toronto, Ontario, Canada

Louis Taillefer & B. J. Ramshaw

## Authors

1. Gaël Grisonnanche

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Yawen Fang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Anaëlle Legros

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Simon Verret

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Francis Laliberté

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Clément Collignon

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Jianshi Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. David Graf

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Paul A. Goddard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Louis Taillefer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. B. J. Ramshaw

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

A.L., P.A.G., L.T. and B.J.R. conceived the experiment. J.Z. grew the sample. A.L., F.L. and C.C. performed the sample preparation and characterization. G.G., Y.F., A.L., D.G., P.A.G. and B.J.R. performed the ADMR measurements at the National High Magnetic Field Laboratory in Tallahassee. G.G., Y.F., S.V. and B.J.R. performed the data analysis and simulations. G.G., L.T. and B.J.R. wrote the manuscript with input from all other co-authors. L.T. and B.J.R. supervised the project.

## Corresponding authors

Correspondence to [Louis Taillefer](#) or [B. J. Ramshaw](#).

## Ethics declarations

### Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Antoine Georges 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 ADMR experimental set up.](#)

**a**, A photograph of the sample on the rotator. The two samples here are mounted on a G-10 wedge to provide an azimuthal angle  $\phi$  of 30°. Additional wedges provided angles of  $\phi = 15^\circ$  and  $\phi = 45^\circ$ . **b**, ADMR as a

function of  $\theta$  angle from  $-15^\circ$  to  $110^\circ$  and  $\phi = 0$  at  $T = 20$  K for Nd-LSCO  $p = 0.24$ , showing the symmetry of the data about these two angles.

### Extended Data Fig. 2 Calculated and measured Sommerfeld coefficients of Nd-LSCO.

**a**, The Sommerfeld coefficient  $\gamma$  for Nd-LSCO as a function of doping. The measured values (red circles) are obtained from measurements of the electronic specific heat  $C_{\text{el}}/T$  at  $T = 10$  K (ref. [25](#)). For the calculated  $\gamma$  (black dashed, dotted and solid lines), we use the tight-binding parameters from our ADMR analysis for three different values of  $t$ , as indicated. The grey band represents the region of consistency between the calculations and the data. **b**, Electronic specific heat  $C_{\text{el}}/T$  as a function of temperature for Nd-LSCO  $p = 0.24, 0.27, 0.36$  and  $0.40$  (ref. [25](#)). The data are the solid lines and the dashed lines represent extrapolations.

### Extended Data Fig. 3 Fit of the Nd-LSCO $p = 0.24$ data with different scattering-rate models.

**a**, ADMR data on Nd-LSCO  $p = 0.24$  at  $T = 25$  K and  $B = 45$  T. **b, c, e, f**, Best fits for the ADMR data in **a** using the Fermi surface in Fig. [1d](#) and an isotropic scattering-rate model (**b**), and three different anisotropic scattering-rate models: cosine (**c**), tanh (**e**) and polynomial (**f**). **d**, The three different anisotropic scattering rates as a function of the azimuthal angle  $\phi$  at  $T = 25$  K.

### Extended Data Fig. 4 ADMR and quasiparticle scattering rate of Nd-LSCO at $p = 0.24$ for the tanh model.

This figure is the same as Figs. [2a](#), [3a, b](#), except that the ADMR has been fitted using the tanh model instead of the cosine model (Extended Data Fig. [3](#)).

### Extended Data Fig. 5 ADMR and quasiparticle scattering rate of Nd-LSCO at $p = 0.24$ for $B = 35$ T.

**a, b**, This figure is the same as Fig. 2a,c except that the ADMR data are taken at  $B = 35$  T (a). The fit has been carried out using the cosine model. b shows that scattering-rate values are identical to within a percent of those obtained from the fit to the data at  $B = 45$  T, shown in Fig. 2c.

**Extended Data Table 1 Tight-binding parameters from the fit to the ADMR data at  $p = 0.24$**

[Full size table](#)

**Extended Data Table 2 Results of the fit of the Nd-LSCO  $p = 0.24$  data with the cosine scattering-rate model**

[Full size table](#)

## Supplementary information

### [Supplementary Information](#)

This file contains Supplementary Notes and Data, Supplementary Figures 1–2 and Supplementary References.

## Rights and permissions

### [Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Grissonnanche, G., Fang, Y., Legros, A. *et al.* Linear-in temperature resistivity from an isotropic Planckian scattering rate. *Nature* **595**, 667–672 (2021). <https://doi.org/10.1038/s41586-021-03697-8>

## Download citation

- Received: 25 November 2020
- Accepted: 03 June 2021
- Published: 28 July 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03697-8>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03697-8>

- Article
- [Published: 28 July 2021](#)

# Spectroscopic evidence for a gold-coloured metallic water solution

- [Philip E. Mason<sup>1</sup> na1](#),
- [H. Christian Schewe<sup>1,2</sup> na1](#),
- [Tillmann Buttersack](#) [ORCID: orcid.org/0000-0002-4547-2656](#)<sup>1,2,3</sup> na1,
- [Vojtech Kostal<sup>1</sup>](#),
- [Marco Vitek<sup>1</sup>](#),
- [Ryan S. McMullen](#) [ORCID: orcid.org/0000-0002-9644-8461](#)<sup>3</sup>,
- [Hebatallah Ali<sup>2,4</sup>](#),
- [Florian Trinter<sup>2,5,6</sup>](#),
- [Chin Lee<sup>2,7,8</sup>](#),
- [Daniel M. Neumark<sup>7,8</sup>](#),
- [Stephan Thürmer](#) [ORCID: orcid.org/0000-0002-8146-4573](#)<sup>9</sup>,
- [Robert Seidel<sup>10,11</sup>](#),
- [Bernd Winter<sup>2</sup>](#),
- [Stephen E. Bradforth<sup>3</sup>](#) &
- [Pavel Jungwirth](#) [ORCID: orcid.org/0000-0002-6892-3288](#)<sup>1</sup>

[Nature](#) volume 595, pages 673–676 (2021) [Cite this article](#)

- 2677 Accesses
- 1 Citations
- 257 Altmetric
- [Metrics details](#)

## Subjects

- [Chemical physics](#)

## Abstract

Insulating materials can in principle be made metallic by applying pressure. In the case of pure water, this is estimated<sup>1</sup> to require a pressure of 48 megabar, which is beyond current experimental capabilities and may only exist in the interior of large planets or stars<sup>2,3,4</sup>. Indeed, recent estimates and experiments indicate that water at pressures accessible in the laboratory will at best be superionic with high protonic conductivity<sup>5</sup>, but not metallic with conductive electrons<sup>1</sup>. Here we show that a metallic water solution can be prepared by massive doping with electrons upon reacting water with alkali metals. Although analogous metallic solutions of liquid ammonia with high concentrations of solvated electrons have long been known and characterized<sup>6,7,8,9</sup>, the explosive interaction between alkali metals and water<sup>10,11</sup> has so far only permitted the preparation of aqueous solutions with low, submetallic electron concentrations<sup>12,13,14</sup>. We found that the explosive behaviour of the water–alkali metal reaction can be suppressed by adsorbing water vapour at a low pressure of about  $10^{-4}$  millibar onto liquid sodium–potassium alloy drops ejected into a vacuum chamber. This set-up leads to the formation of a transient gold-coloured layer of a metallic water solution covering the metal alloy drops. The metallic character of this layer, doped with around  $5 \times 10^{21}$  electrons per cubic centimetre, is confirmed using optical reflection and synchrotron X-ray photoelectron spectroscopies.

## 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

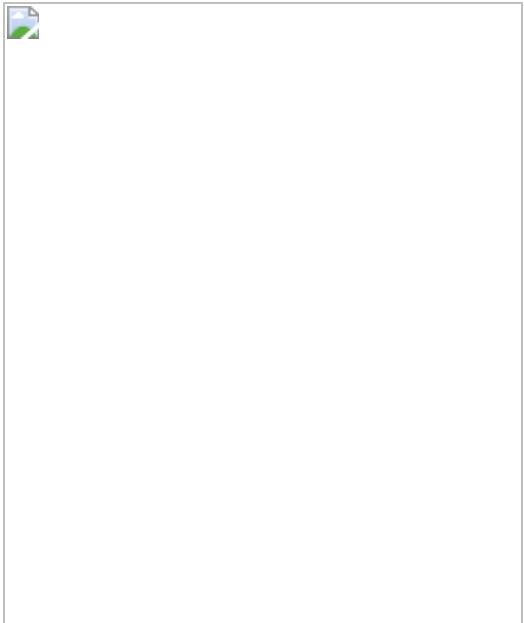
[Rent or Buy](#)

All prices are NET prices.

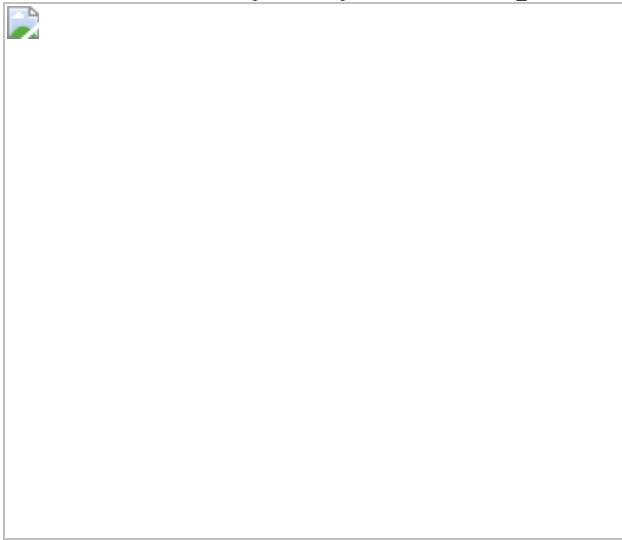
**Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

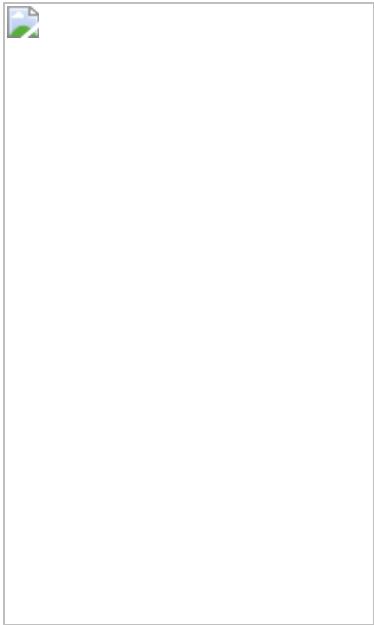
**Fig. 1: A pure NaK drop in vacuum and the time evolution of a NaK drop exposed to water vapour.**



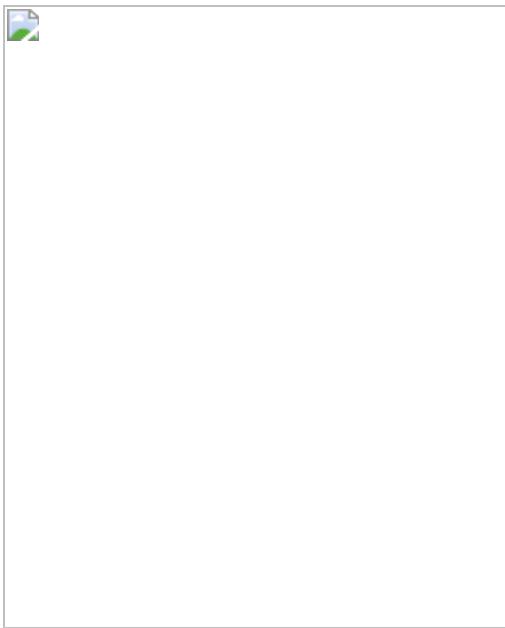
**Fig. 2:** Schematic demonstrating the formation of a thin gold-coloured metallic water layer by water vapour adsorption on a NaK drop.



**Fig. 3:** Spectroscopic signatures of the metallic water solution from optical and X-ray photoelectron spectroscopy.



**Fig. 4: Fits to the experimental data employing a free electron gas model.**



## Data availability

The datasets generated during the current study are available as Source Data or are available from the corresponding author upon reasonable request. [Source data](#) are provided with this paper.

## Code availability

Data-processing and fitting results can be generated using numerical methods described in Methods and [Supplementary Information](#) and developed computer codes that are available from the corresponding author upon reasonable request.

## References

1. 1.

Hermann, A., Ashcroft, N. W. & Hoffmann, R. High pressure ices. *Proc. Natl Acad. Sci. USA* **109**, 745–750 (2012).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

2. 2.

Dubrovinsky, L., Dubrovinskaia, N., Prakapenka, V. B. & Abakumov, A. M. Implementation of micro-ball nanodiamond anvils for high-pressure studies above 6 Mbar. *Nat. Commun.* **3**, 1163 (2012).

[ADS](#) [Article](#) [Google Scholar](#)

3. 3.

Cavazzoni, C. et al. Superionic and metallic states of water and ammonia at giant planet conditions. *Science* **283**, 44–46 (1999).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Mattsson, T. R. & Desjarlais, M. P. Phase diagram and electrical conductivity of high energy-density water from density functional theory. *Phys. Rev. Lett.* **97**, 017801 (2006).

[ADS](#) [Article](#) [Google Scholar](#)

5. 5.

Millot, M. et al. Nanosecond X-ray diffraction of shock-compressed superionic water ice. *Nature* **569**, 251–255 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Zurek, E., Edwards, P. P. & Hoffmann, R. A molecular perspective on lithium-ammonia solutions. *Angew. Chem. Int. Ed.* **48**, 8198–8232 (2009).

[CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Thompson, J. C. *Electrons in Liquid Ammonia* (Clarendon Press, 1976).

8. 8.

Lodge, M. et al. Multielement NMR studies of the liquid-liquid phase separation and the metal-to-nonmetal transition in fluid lithium- and sodium-ammonia solutions. *J. Phys. Chem. B* **117**, 13322–13334 (2013).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Buttersack, T. et al. Photoelectron spectra of alkali metal-ammonia microjets: from blue electrolyte to bronze metal. *Science* **368**, 1086–1091 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Hutton, A. T. Dramatic demonstration for a large audience: the formation of hydroxyl ions in the reaction of sodium with water. *J. Chem. Educ.* **58**, 506 (1981).

[CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Mason, P. E. et al. Coulomb explosion during the early stages of the reaction of alkali metals with water. *Nat. Chem.* **7**, 250–254 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Young, R. M. & Neumark, D. M. Dynamics of solvated electrons in clusters. *Chem. Rev.* **112**, 5553–5577 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Mason, P. E., Buttersack, T., Bauerecker, S. & Jungwirth, P. A non-exploding alkali metal drop on water: from blue solvated electrons to bursting molten hydroxide. *Angew. Chem. Int. Ed.* **55**, 13019–13022 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

14. 14.

Suzuki, T. Ultrafast photoelectron spectroscopy of aqueous solutions. *J. Chem. Phys.* **151**, 090901 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

15. 15.

Alchagirov, B. B. et al. Surface tension and adsorption of components in the sodium–potassium alloy systems: effective liquid metal coolants promising in nuclear and space power engineering. *Inorg. Mater. Appl. Res.* **2**, 461–467 (2011).

[Article](#) [Google Scholar](#)

16. 16.

Addison, C. C. *The Chemistry of Liquid Alkali Metals* (Wiley, 1984).

17. 17.

Citrin, P. H. High-resolution X-ray photoemission from sodium metal and its hydroxide. *Phys. Rev. B* **8**, 5545–5556 (1973).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Lueth, H. *Surfaces and Interfaces of Solid Materials* (Springer, 1997).

19. 19.

Kiskinova, M., Pirug, G. & Bonzel, H. P. Adsorption and decomposition of H<sub>2</sub>O on a K-covered Pt(111) surface. *Surf. Sci.* **150**, 319–338 (1985).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Blass, P. M., Zhou, X. L. & White, J. M. Coadsorption and reaction of water and potassium on Ag(111). *J. Phys. Chem.* **94**, 3054–3062 (1990).

[CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Nachtrieb, N. H. Self-diffusion in liquid metals. *Adv. Phys.* **16**, 309–323 (1967).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Ketteler, G. et al. The nature of water nucleation sites on TiO<sub>2</sub>(110) surfaces revealed by ambient pressure X-ray photoelectron spectroscopy. *J. Phys. Chem. C* **111**, 8278–8282 (2007).

[CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Kittel, C. *Introduction to Solid State Physics* (Wiley, 2005).

24. 24.

Abelès, F., Borensztein, Y., Decrescenzi, M. & Lopezrios, T. Optical evidence for longitudinal-waves in very thin Ag layers. *Surf. Sci.* **101**, 123–130 (1980).

[ADS](#) [Article](#) [Google Scholar](#)

25. 25.

Winter, B., Faubel, M., Vacha, R. & Jungwirth, P. Behavior of hydroxide at the water/vapor interface. *Chem. Phys. Lett.* **474**, 241–247 (2009).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Bonzel, H. P., Pirug, G. & Winkler, A. Adsorption of H<sub>2</sub>O on potassium films. *Surf. Sci.* **175**, 287–312 (1986).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Hart, E. J. & Boag, J. W. Absorption spectrum of hydrated electron in water and in aqueous solutions. *J. Am. Chem. Soc.* **84**, 4090–4095 (1962).

[CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Barzynski, H. & Schulte-Frohlinde, D. On the nature of the electron traps in alkaline ice. *Z. Naturforsch. A* **22**, 2131–2132 (1967).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Kevan, L. Slovated electron structure in glassy matrices. *Acc. Chem. Res.* **14**, 138–145 (1981).

[CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Buttersack, T. et al. Deeply cooled and temperature controlled microjets: liquid ammonia solutions released into vacuum for analysis by photoelectron spectroscopy. *Rev. Sci. Instrum.* **91**, 043101 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Seidel, R., Pohl, M. N., Ali, H., Winter, B. & Aziz, E. F. Advances in liquid phase soft-X-ray photoemission spectroscopy: a new experimental setup at BESSY II. *Rev. Sci. Instrum.* **88**, 073107 (2017).

32. 32.

Sawhney, K. J. S., Senf, F. & Gudat, W. PGM beamline with constant energy resolution mode for U49-2 undulator at BESSY-II. *Nucl. Instrum. Methods Phys. Res. Sect. A* **467**, 466–469 (2001).

[Download references](#)

## Acknowledgements

P.J. acknowledges support from the European Regional Development Fund (project ChemBioDrug number CZ.02.1.01/0.0/0.0/16\_019/0000729). D.M.N. and C.L. acknowledge support from the Director of the Office of Basic Energy Science, Chemical Sciences Division of the US Department of Energy under contract number DE-AC02-05CH11231. D.M.N., C.L. and P.J. thank the Alexander von Humboldt Foundation for support. S.E.B., T.B. and R.S.M. are supported by the US National Science Foundation (CHE-1665532). P.E.M. acknowledges support from the viewers of his YouTube popular science channel. M.V. acknowledges support from the Charles University in Prague and from the International Max Planck Research School in Dresden. S.T. acknowledges support from JSPS KAKENHI grant number JP18K14178. R.S. acknowledges the German Research Foundation (DFG) for an Emmy-Noether Young-Investigator stipend (DFG, project SE 2253/3-1). F.T. and B.W. acknowledge support from the MaxWater initiative of the Max-Planck-Gesellschaft. All authors thank the staff of the Helmholtz-Zentrum Berlin for their support during the beamtime at BESSY II.

## Author information

### Author notes

1. These authors contributed equally: Philip E. Mason, H. Christian Schewe, Tillmann Buttersack

## Affiliations

1. Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague, Czech Republic

Philip E. Mason, H. Christian Schewe, Tillmann Buttersack, Vojtech Kostal, Marco Vitek & Pavel Jungwirth

2. Molecular Physics, Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin, Germany

H. Christian Schewe, Tillmann Buttersack, Hebatallah Ali, Florian Trinter, Chin Lee & Bernd Winter

3. Department of Chemistry, University of Southern California, Los Angeles, CA, USA

Tillmann Buttersack, Ryan S. McMullen & Stephen E. Bradforth

4. Department of Physics, Faculty of Women for Art, Science and Education, Ain Shams University, Cairo, Egypt

Hebatallah Ali

5. Photon Science, Deutsches Elektronen-Synchrotron (DESY), Hamburg, Germany

Florian Trinter

6. Institut für Kernphysik, Goethe-Universität Frankfurt, Frankfurt am Main, Germany

Florian Trinter

7. Department of Chemistry, University of California, Berkeley, CA, USA

Chin Lee & Daniel M. Neumark

8. Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Chin Lee & Daniel M. Neumark

9. Department of Chemistry, Graduate School of Science, Kyoto University, Kyoto, Japan

Stephan Thürmer

10. Helmholtz-Zentrum Berlin für Materialien und Energie, Berlin, Germany

Robert Seidel

11. Department of Chemistry, Humboldt-Universität zu Berlin, Berlin, Germany

Robert Seidel

## Authors

1. Philip E. Mason

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. H. Christian Schewe

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Tillmann Buttersack

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Vojtech Kostal

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Marco Vitek

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Ryan S. McMullen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Hebatallah Ali

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Florian Trinter

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Chin Lee

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Daniel M. Neumark

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Stephan Thürmer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Robert Seidel

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Bernd Winter

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Stephen E. Bradforth

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Pavel Jungwirth

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

P.E.M., H.C.S., T.B., B.W., S.E.B. and P.J. designed the experiments.

P.E.M., H.C.S., T.B., B.W., H.A., V.K., M.V., F.T., C.L., D.M.N., R.S. and P.J. performed the experiments, and P.E.M., H.C.S., T.B., B.W., S.E.B., R.S.M., C.L., R.S., S.T. and P.J. analysed the obtained data. P.J. wrote the main paper, and H.C.S. and P.J. wrote the Supplementary Information, both with critical feedback from all co-authors. P.E.M. produced Supplementary Video 1.

## Corresponding author

Correspondence to [Pavel Jungwirth](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Christoph Salzmann 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 Supplementary Information file contains the following sections: (1) Sample preparation and employment in the experimental setup; (2) Photoelectron spectroscopy measurements; (3) Photoelectron spectra of pure NaK jets versus drops; (4) Photoelectron spectra of water vapour adsorbing on the NaK drop surface; (5) Data-analysis of photoelectron spectra acquired in fixed-mode operation; (6) Data-fitting of photoelectron spectra; (7) Optical reflection spectroscopy; and (8) Supplementary References

### Supplementary Video 1

A detailed description of the experimental setup and conduction of the experiments.

### Peer Review File

## Source data

### Source Data Fig. 3

### Source Data Fig. 4

## Rights and permissions

## [Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Mason, P.E., Schewe, H.C., Buttersack, T. *et al.* Spectroscopic evidence for a gold-coloured metallic water solution. *Nature* **595**, 673–676 (2021).  
<https://doi.org/10.1038/s41586-021-03646-5>

## [Download citation](#)

- Received: 07 October 2020
- Accepted: 14 May 2021
- Published: 28 July 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03646-5>

## Further reading

- [\*\*Water transformed into shiny, golden metal\*\*](#)
  - Davide Castelvecchi

*Nature* (2021)

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03646-5>

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

- Article
- [Published: 20 May 2021](#)

# A radical approach for the selective C–H borylation of azines

- [Ji Hye Kim<sup>1</sup>](#),
- [Timothée Constantin](#) ORCID: [orcid.org/0000-0001-5376-1557<sup>1</sup>](https://orcid.org/0000-0001-5376-1557),
- [Marco Simonetti<sup>1</sup>](#),
- [Josep Llaveria](#) ORCID: [orcid.org/0000-0003-2260-3657<sup>2</sup>](https://orcid.org/0000-0003-2260-3657),
- [Nadeem S. Sheikh<sup>3</sup>](#) &
- [Daniele Leonori](#) ORCID: [orcid.org/0000-0002-7692-4504<sup>1</sup>](https://orcid.org/0000-0002-7692-4504)

[Nature](#) volume 595, pages 677–683 (2021) [Cite this article](#)

- 18k Accesses
- 53 Altmetric
- [Metrics details](#)

## Subjects

- [Synthetic chemistry methodology](#).
- [Photocatalysis](#)

## Abstract

Boron functional groups are often introduced in place of aromatic carbon–hydrogen bonds to expedite small-molecule diversification through coupling of molecular fragments<sup>1,2,3</sup>. Current approaches based on transition-metal-catalysed activation of carbon–hydrogen bonds are

effective for the borylation of many (hetero)aromatic derivatives<sup>4,5</sup> but show narrow applicability to azines (nitrogen-containing aromatic heterocycles), which are key components of many pharmaceutical and agrochemical products<sup>6</sup>. Here we report an azine borylation strategy using stable and inexpensive amine-borane<sup>7</sup> reagents. Photocatalysis converts these low-molecular-weight materials into highly reactive boryl radicals<sup>8</sup> that undergo efficient addition to azine building blocks. This reactivity provides a mechanistically alternative tactic for  $sp^2$  carbon–boron bond assembly, where the elementary steps of transition-metal-mediated carbon–hydrogen bond activation and reductive elimination from azine-organometallic intermediates are replaced by a direct, Minisci<sup>9</sup>-style, radical addition. The strongly nucleophilic character of the amine-boryl radicals enables predictable and site-selective carbon–boron bond formation by targeting the azine’s most activated position, including the challenging sites adjacent to the basic nitrogen atom. This approach enables access to aromatic sites that elude current strategies based on carbon–hydrogen bond activation, and has led to borylated materials that would otherwise be difficult to prepare. We have applied this process to the introduction of amine-borane functionalities to complex and industrially relevant products. The diversification of the borylated azine products by mainstream cross-coupling technologies establishes aromatic amino-boranes as a powerful class of building blocks for chemical synthesis.

## 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

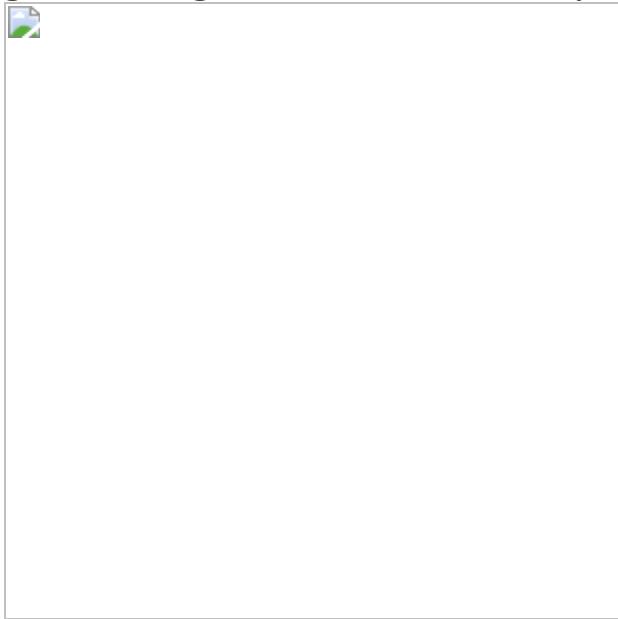
[Rent or Buy](#)

All prices are NET prices.

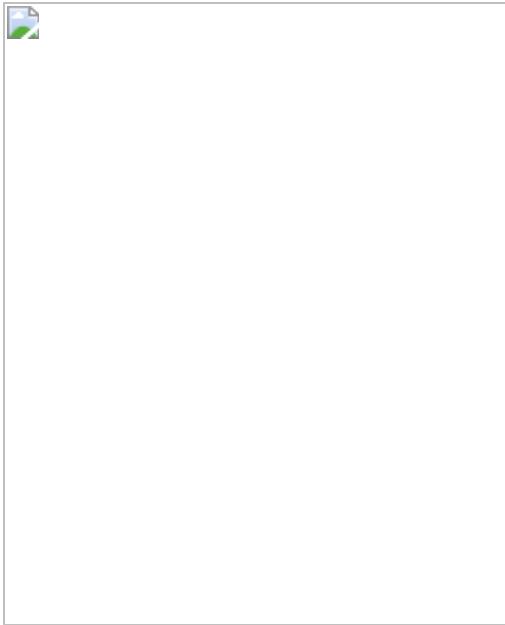
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Strategies for azine C–H borylation.**



**Fig. 2: Substrate scope for the radical C–H borylation of azines.**



**Fig. 3: Diversification of the azine amine-boranes by oxidation, Suzuki cross-coupling, Chan–Lam amination and Chan–Lam etherification.**



## Data availability

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

## References

1. 1.

Roughley, S. D. & Jordan, A. M. The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. *J. Med. Chem.* **54**, 3451–3479 (2011).

[CAS](#) [Google Scholar](#)

2. 2.

Blakemore, D. C. et al. Organic synthesis provides opportunities to transform drug discovery. *Nat. Chem.* **10**, 383–394 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Blakemore, D. in *Synthetic Methods in Drug Discovery* Vol. 1 (eds Blakemore, D. C., Doyle, P. M. & Fobian, Y. M.) 1–69 (Royal Society of Chemistry, 2016).

4. 4.

Mkhalid, I. A. I., Barnard, J. H., Marder, T. B., Murphy, J. M. & Hartwig, J. F. C–H activation for the construction of C–B bonds. *Chem. Rev.* **110**, 890–931 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Larsen, M. A. & Hartwig, J. F. Iridium-catalyzed C–H borylation of heteroarenes: scope, regioselectivity, application to late-stage functionalization, and mechanism. *J. Am. Chem. Soc.* **136**, 4287–4299 (2014).

[CAS](#) [Google Scholar](#)

6. 6.

Taylor, R. D., MacCoss, M. & Lawson, A. D. G. Rings in drugs. *J. Med. Chem.* **57**, 5845–5859 (2014).

[CAS](#) [Google Scholar](#)

7. 7.

Staubitz, A., Robertson, A. P. M., Sloan, M. E. & Manners, I. Amine- and phosphine–borane adducts: new interest in old molecules. *Chem. Rev.* **110**, 4023–4078 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

8. 8.

Baban, J. A., Marti, V. P. J. & Roberts, B. P. Ligated boryl radicals. Part 2. Electron spin resonance studies of trialkylamin–boryl radicals. *J. Chem. Soc. Perkin Trans. 2* 1723–1733 (1985).

9. 9.

Duncton, M. A. J. Minisci reactions: versatile CH-functionalizations for medicinal chemists. *MedChemComm* **2**, 1135–1161 (2011).

[CAS](#) [Google Scholar](#)

10. 10.

Hall, D. G. in *Boronic Acids* (ed. Hall, D. G.) 1–99 (Wiley, 2005).

11. 11.

West, M. J., Fyfe, J. W. B., Vantourout, J. C. & Watson, A. J. B. Mechanistic development and recent applications of the Chan–Lam amination. *Chem. Rev.* **119**, 12491–12523 (2019).

[CAS](#) [Google Scholar](#)

12. 12.

Furuya, T., Kaiser, H. M. & Ritter, T. Palladium-mediated fluorination of arylboronic acids. *Angew. Chem. Int. Ed.* **47**, 5993–5996 (2008).

[CAS](#) [Google Scholar](#)

13. 13.

Bonet, A., Odachowski, M., Leonori, D., Essafi, S. & Aggarwal, V. K. Enantiospecific  $sp^2$ – $sp^3$  coupling of secondary and tertiary boronic esters. *Nat. Chem.* **6**, 584–589 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Ishiyama, T., Murata, M. & Miyaura, N. Palladium(0)-catalyzed cross-coupling reaction of alkoxydiboron with haloarenes: a direct procedure for arylboronic esters. *J. Org. Chem.* **60**, 7508–7510 (1995).

[CAS](#) [Google Scholar](#)

15. 15.

Vitaku, E., Smith, D. T. & Njardarson, J. T. Analysis of the structural diversity, substitution patterns, and frequency of nitrogen heterocycles among U.S. FDA approved pharmaceuticals. *J. Med. Chem.* **57**, 10257–10274 (2014).

[CAS](#) [Google Scholar](#)

16. 16.

Sadler, S. A. et al. Iridium-catalyzed C–H borylation of pyridines. *Org. Biomol. Chem.* **12**, 7318–7327 (2014).

[CAS](#) [Google Scholar](#)

17. 17.

Brown, D. G., Gagnon, M. M. & Boström, J. Understanding our love affair with *p*-chlorophenyl: present day implications from historical biases of reagent selection. *J. Med. Chem.* **58**, 2390–2405 (2015).

[CAS](#) [Google Scholar](#)

18. 18.

Cook, X. A. F., de Gombert, A., McKnight, J., Pantaine, L. R. E. & Willis, M. C. The 2-pyridyl problem: challenging nucleophiles in cross-coupling arylations. *Angew. Chem. Int. Ed.* **60**, 11068–11091 (2020).

[Google Scholar](#)

19. 19.

Li, J. et al. Synthesis of many different types of organic small molecules using one automated process. *Science* **347**, 1221–1226 (2015).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

20. 20.

Dick, G. R., Woerly, E. M. & Burke, M. D. A general solution for the 2-pyridyl problem. *Angew. Chem. Int. Ed.* **51**, 2667–2672 (2012).

[CAS](#) [Google Scholar](#)

21. 21.

Proctor, R. S. J. & Phipps, R. J. Recent advances in Minisci-type reactions. *Angew. Chem. Int. Ed.* **58**, 13666–13699 (2019).

[CAS](#) [Google Scholar](#)

22. 22.

Taniguchi, T. Boryl radical addition to multiple bonds in organic synthesis. *Eur. J. Org. Chem.* **2019**, 6308–6319 (2019).

[CAS](#) [Google Scholar](#)

23. 23.

Curran, D. P. et al. Synthesis and reactions of N-heterocyclic carbene boranes. *Angew. Chem. Int. Ed.* **50**, 10294–10317 (2011).

[MathSciNet](#) [CAS](#) [Google Scholar](#)

24. 24.

Giles, J. R. M. & Roberts, B. P. An electron spin resonance study of the generation and reactions of borane radical anions in solution. *J. Chem. Soc. Perkin Trans. 2* 743–755 (1983).

25. 25.

Mills, H. A., Martin, J. L., Rheingold, A. L. & Spokoyny, A. M. Oxidative generation of boron-centered radicals in carboranes. *J. Am. Chem. Soc.* **142**, 4586–4591 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Yang, L., Uemura, N. & Nakao, Y. Meta-selective C–H borylation of benzamides and pyridines by an iridium–Lewis acid bifunctional catalyst. *J. Am. Chem. Soc.* **141**, 7972–7979 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

27. 27.

Mkhalid, I. A. I. et al. Ir-catalyzed borylation of C–H bonds in N-containing heterocycles: regioselectivity in the synthesis of heteroaryl boronate esters. *Angew. Chem. Int. Ed.* **45**, 489–491 (2006).

[CAS](#) [Google Scholar](#)

28. 28.

Preshlock, S. M. et al. A traceless directing group for C–H borylation. *Angew. Chem. Int. Ed.* **52**, 12915–12919 (2013).

[CAS](#) [Google Scholar](#)

29. 29.

Chotana, G. A., Rak, M. A. & Smith, M. R. Sterically directed functionalization of aromatic C–H bonds: selective borylation ortho to cyano groups in arenes and heterocycles. *J. Am. Chem. Soc.* **127**, 10539–10544 (2005).

[CAS](#) [Google Scholar](#)

30. 30.

Minisci, F. et al. Polar effects in free-radical reactions. Selectivity and reversibility in the homolytic benzylation of protonated heteroaromatic bases. *J. Org. Chem.* **51**, 476–479 (1986).

[CAS](#) [Google Scholar](#)

31. 31.

Ueng, S.-H. et al. Complexes of borane and N-heterocyclic carbenes: a new class of radical hydrogen atom donor. *J. Am. Chem. Soc.* **130**, 10082–10083 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

32. 32.

Dai, W., Geib, S. J. & Curran, D. P. 1,4-Hydroboration reactions of electron-poor aromatic rings by N-heterocyclic carbene boranes. *J. Am. Chem. Soc.* **142**, 6261–6267 (2020).

[Google Scholar](#)

33. 33.

Hioe, J., Karton, A., Martin, J. M. L. & Zipse, H. Borane–Lewis base complexes as homolytic hydrogen atom donors. *Chem. Eur. J.* **16**, 6861–6865 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

34. 34.

Tehfe, M.-A. et al. N-heterocyclic carbene-borane radicals as efficient initiating species of photopolymerization reactions under air. *Polym. Chem.* **2**, 625–631 (2011).

[CAS](#) [Google Scholar](#)

35. 35.

Roberts, P. B. Polarity-reversal catalysis of hydrogen-atom abstraction reactions: concepts and applications in organic chemistry. *Chem. Soc. Rev.* **28**, 25–35 (1999).

[CAS](#) [Google Scholar](#)

36. 36.

Wu, C., Hou, X., Zheng, Y., Li, P. & Lu, D. Electrophilicity and nucleophilicity of boryl radicals. *J. Org. Chem.* **82**, 2898–2905 (2017).

[CAS](#) [Google Scholar](#)

37. 37.

Sheeller, B. & Ingold, K. U. Absolute rate constants for some reactions of the triethylamine-boryl radical and the borane radical anion. *J. Chem. Soc. Perkin Trans. 2* 480–486 (2001).

38. 38.

Lalevée, J., Tehfe, M. A., Allonas, X. & Fouassier, J. P. Boryl radicals as a new photoinitiating species: a way to reduce the oxygen inhibition. *Macromolecules* **41**, 9057–9062 (2008).

[ADS](#) [Google Scholar](#)

39. 39.

Jin, J. & MacMillan, D. W. C. Direct  $\alpha$ -arylation of ethers through the combination of photoredox-mediated C–H functionalization and the Minisci reaction. *Angew. Chem. Int. Ed.* **54**, 1565–1569 (2015).

[CAS](#) [Google Scholar](#)

40. 40.

Bietti, M. Activation and deactivation strategies promoted by medium effects for selective aliphatic C–H bond functionalization. *Angew. Chem. Int. Ed.* **57**, 16618–16637 (2018).

[CAS](#) [Google Scholar](#)

41. 41.

De Vleeschouwer, F., Speybroeck, V. V., Waroquier, M., Geerlings, P. & Proft, F. D. Electrophilicity and nucleophilicity index for radicals. *Org. Lett.* **9**, 2721–2724 (2007).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Baban, J. A. & Roberts, B. P. An electron spin resonance study of phosphine-boryl radicals; their structures and reactions with alkyl halides. *J. Chem. Soc. Perkin Trans. 2* 1717–1722 (1984).

43. 43.

Davis, H. J., Mihai, M. T. & Phipps, R. J. Ion pair-directed regiocontrol in transition-metal catalysis: a meta-selective C–H borylation of aromatic quaternary ammonium salts. *J. Am. Chem. Soc.* **138**, 12759–12762 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

44. 44.

Cox, P. A. et al. Base-catalyzed aryl-B(OH)<sub>2</sub> protodeboronation revisited: from concerted proton transfer to liberation of a transient aryl anion. *J. Am. Chem. Soc.* **139**, 13156–13165 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

45. 45.

Molloy, J. J. et al. Chemoselective oxidation of aryl organoboron systems enabled by boronic acid-selective phase transfer. *Chem. Sci.* **8**, 1551–1559 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Kuninobu, Y., Nagase, M. & Kanai, M. Benzylic C(sp<sup>3</sup>)–H perfluoroalkylation of six-membered heteroaromatic compounds. *Angew. Chem. Int. Ed* **54**, 10263–10266 (2015).

[CAS](#) [Google Scholar](#)

47. 47.

Garattini, E. & Terao, M. Aldehyde oxidase and its importance in novel drug discovery: present and future challenges. *Expert Opin. Drug Discov.* **8**, 641–654 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

Brown, D. G. & Boström, J. Analysis of past and present synthetic methodologies on medicinal chemistry: where have all the new reactions gone? *J. Med. Chem.* **59**, 4443–4458 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Chu, Q. et al. Ionic and organometallic reductions with N-heterocyclic carbene boranes. *Chem. Eur. J.* **15**, 12937–12940 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

D.L. thanks the EPSRC for a fellowship (EP/P004997/1) and the European Research Council for a research grant (758427); J.H.K. thanks Marie Curie Actions for a fellowship (842422). The Chemical Capabilities team of Janssen Research and Development (Toledo) is acknowledged for technical support.

## Author information

### Affiliations

1. Department of Chemistry, University of Manchester, Manchester, UK

Ji Hye Kim, Timothée Constantin, Marco Simonetti & Daniele Leonori

2. Discovery Chemistry, Janssen Research and Development, Toledo, Spain

Josep Llaveria

3. Department of Chemistry, College of Science, King Faisal University, Al-Ahsa, Saudi Arabia

Nadeem S. Sheikh

## Authors

1. Ji Hye Kim

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Timothée Constantin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Marco Simonetti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Josep Llaveria

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Nadeem S. Sheikh

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Daniele Leonori

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

J.H.K. and D.L. designed the project; J.H.K., T.C., M.S. and J.L. performed the experiments; N.S.S. performed the computational studies; and all the authors analysed the results and wrote the paper.

## Corresponding author

Correspondence to [Daniele Leonori](#).

## Ethics declarations

### Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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

## Supplementary information

### [Supplementary Information](#)

This file contains Supplementary Information.

## Rights and permissions

### [Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Kim, J.H., Constantin, T., Simonetti, M. *et al.* A radical approach for the selective C–H borylation of azines. *Nature* **595**, 677–683 (2021).  
<https://doi.org/10.1038/s41586-021-03637-6>

[Download citation](#)

- Received: 02 January 2021
- Accepted: 12 May 2021
- Published: 20 May 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03637-6>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

## Associated Content

Collection

## Synthesis and enabling technologies

### Direct installation of boron groups offers boost to medicinal chemistry

- Christine M. Le

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03637-6>

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

- Article
- [Published: 28 July 2021](#)

# A process-based approach to understanding and managing triggered seismicity

- [Bradford H. Hager](#) ORCID: [orcid.org/0000-0002-5643-1374](https://orcid.org/0000-0002-5643-1374)<sup>1</sup>,
- [James Dieterich](#)<sup>2</sup>,
- [Cliff Frohlich](#)<sup>3</sup>,
- [Ruben Juanes](#) ORCID: [orcid.org/0000-0002-7370-2332](https://orcid.org/0000-0002-7370-2332)<sup>4,5</sup>,
- [Stefano Mantica](#) ORCID: [orcid.org/0000-0002-4137-8355](https://orcid.org/0000-0002-4137-8355)<sup>6</sup>,
- [John H. Shaw](#) ORCID: [orcid.org/0000-0002-3971-3812](https://orcid.org/0000-0002-3971-3812)<sup>7</sup>,
- [Francesca Bottazzi](#) ORCID: [orcid.org/0000-0002-7150-9336](https://orcid.org/0000-0002-7150-9336)<sup>6</sup>,
- [Federica Caresani](#) ORCID: [orcid.org/0000-0002-2126-5409](https://orcid.org/0000-0002-2126-5409)<sup>6</sup>,
- [David Castineira](#)<sup>4</sup>,
- [Alberto Cominelli](#)<sup>6</sup>,
- [Marco Meda](#) ORCID: [orcid.org/0000-0002-1751-1507](https://orcid.org/0000-0002-1751-1507)<sup>6</sup>,
- [Lorenzo Osculati](#) ORCID: [orcid.org/0000-0001-9545-5662](https://orcid.org/0000-0001-9545-5662)<sup>6</sup>,
- [Stefania Petroselli](#) ORCID: [orcid.org/0000-0003-1718-8612](https://orcid.org/0000-0003-1718-8612)<sup>6</sup> &
- [Andreas Plesch](#) ORCID: [orcid.org/0000-0002-3355-9199](https://orcid.org/0000-0002-3355-9199)<sup>7</sup>

[Nature](#) volume **595**, pages 684–689 (2021) [Cite this article](#)

- 721 Accesses
- 82 Altmetric
- [Metrics details](#)

## Subjects

- [Crude oil](#)
- [Geophysics](#)
- [Natural hazards](#)
- [Seismology](#)

## Abstract

There is growing concern about seismicity triggered by human activities, whereby small increases in stress bring tectonically loaded faults to failure. Examples of such activities include mining, impoundment of water, stimulation of geothermal fields, extraction of hydrocarbons and water, and the injection of water, CO<sub>2</sub> and methane into subsurface reservoirs<sup>1</sup>. In the absence of sufficient information to understand and control the processes that trigger earthquakes, authorities have set up empirical regulatory monitoring-based frameworks with varying degrees of success<sup>2,3</sup>. Field experiments in the early 1970s at the Rangely, Colorado (USA) oil field<sup>4</sup> suggested that seismicity might be turned on or off by cycling subsurface fluid pressure above or below a threshold. Here we report the development, testing and implementation of a multidisciplinary methodology for managing triggered seismicity using comprehensive and detailed information about the subsurface to calibrate geomechanical and earthquake source physics models. We then validate these models by comparing their predictions to subsequent observations made after calibration. We use our approach in the Val d'Agri oil field in seismically active southern Italy, demonstrating the successful management of triggered seismicity using a process-based method applied to a producing hydrocarbon field. Applying our approach elsewhere could help to manage and mitigate triggered seismicity.

## 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

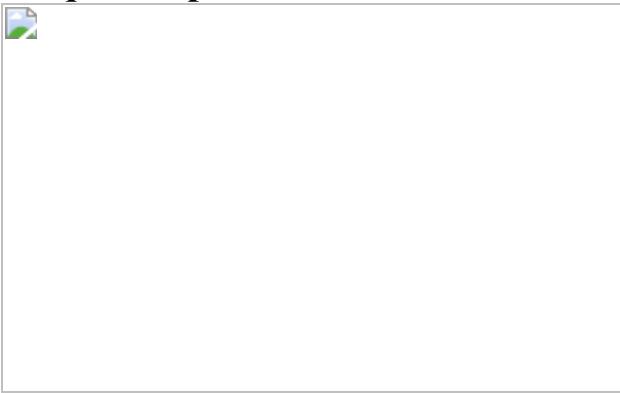
[Rent or Buy](#)

All prices are NET prices.

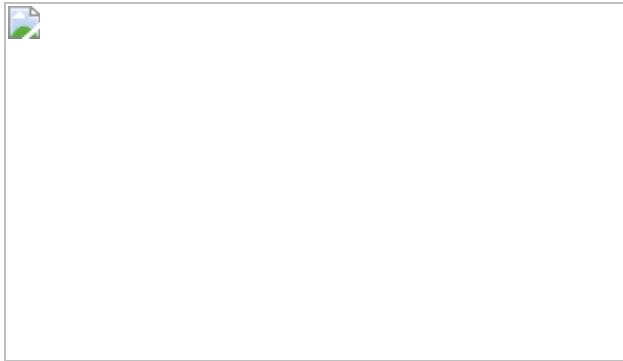
**Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

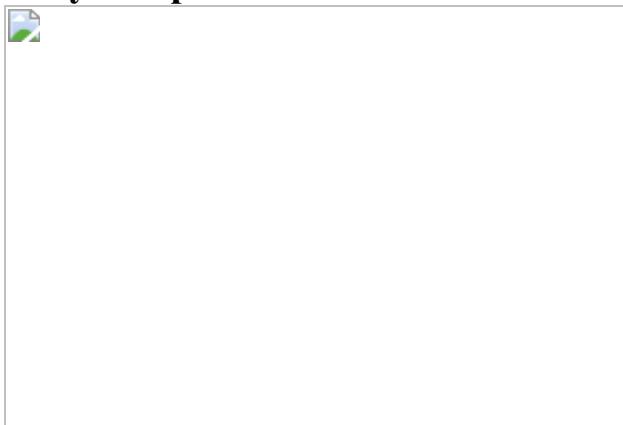
**Fig. 1: Map and subsurface structure of the Val d'Agri region with earthquake epicentres associated with the CM2 injection well.**



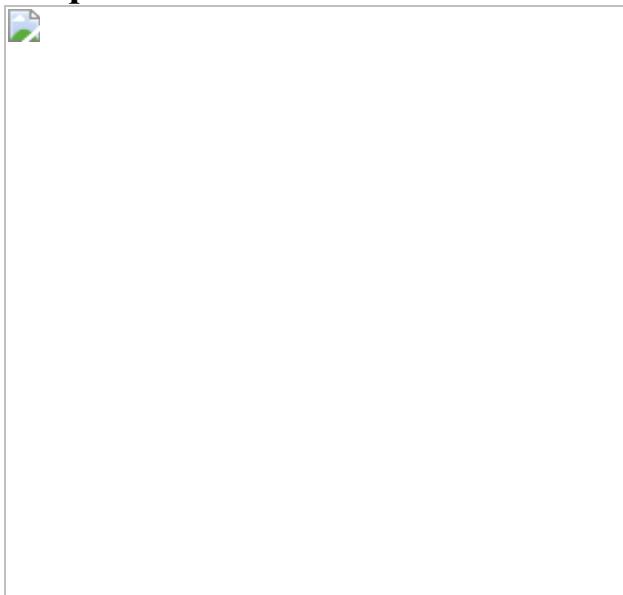
**Fig. 2: Three-dimensional structural model of the Val d'Agri field.**



**Fig. 3: Modelled Coulomb stress changes, fault slip and earthquake density compared with observed seismicity.**



**Fig. 4: Field injection rates, stress changes, moment release and earthquake occurrence over time.**



## **Data availability**

Relevant data are available at <https://doi.org/10.6084/m9.figshare.c.5401509> including the computational meshes with embedded stratigraphic horizons and fault surfaces for the regional and local models, pressure history on the CMF, input file and script for the seismicity rate model, and source data for Figs. 1c and 3. Total field monthly oil and gas production as tabulated by the Italian Ministry of Economic Development are included in the figshare repository. Some input data for the flow–geomechanics models contain proprietary information, made available by Eni for the current study under confidentiality agreement. These data are available from S.M. (stefano.mantica@eni.com) with permission of Eni. [Source data](#) are provided with this paper.

## Code availability

The seismicity rate modelling code is included at <https://doi.org/10.6084/m9.figshare.14519028>.

## References

1. 1.

Foulger, G. R., Wilson, M., Gluyas, J., Julian, B. R. & Davies, R. Global review of human-induced earthquakes. *Earth Sci. Rev.* **178**, 438–514 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

2. 2.

Baisch, S., Koch, C. & Muntendam-Bos, A. Traffic light systems: to what extent can induced seismicity be controlled? *Seismol. Res. Lett.* **90**, 1145–1154 (2019).

[Article](#) [Google Scholar](#)

3. 3.

Kwiatek, G. et al. Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland. *Sci. Adv.* **5**, eaav7224 (2019).

[ADS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

4. 4.

Raleigh, C. B., Healy, J. H. & Bredehoeft, J. D. An experiment in earthquake control at Rangely, Colorado. *Science* **191**, 1230–1237 (1976).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

5. 5.

Improtta, L. et al. Reservoir structure and wastewater induced seismicity at the Val d’Agri oilfield (Italy) shown by three-dimensional  $V_p$  and  $V_p/V_s$  local earthquakes tomography. *J. Geophys. Res. Solid Earth* **122**, 9050–9082 (2017).

[ADS](#) [Article](#) [Google Scholar](#)

6. 6.

Buttinelli, M., Improtta, L., Bagh, S. & Chiarabba, C. Inversion of inherited thrusts by wastewater injection induced seismicity at the Val d’Agri oilfield (Italy). *Sci. Rep.* **6**, 37165 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

7. 7.

Cello, G. & Mazzoli, S. Apennine tectonics in southern Italy: a review. *J. Geodyn.* **27**, 191–211 (1998).

[Article](#) [Google Scholar](#)

8. 8.

D'Argenio, B., Pescatore, T. & Scandone, P. Structural pattern of the Campania-Lucania Apennines. *Quad. Ric. Sci.* **90**, 313–327 (1975).

[Google Scholar](#)

9. 9.

Butler, R. W. H. et al. in *Thrust Tectonics and Hydrocarbon Systems: American Association of Petroleum Geologists Memoir 82* (ed. McClay, K. R.) 647–667 (American Association of Petroleum Geologists, 2004).

10. 10.

Valoroso, L. *Upper Crustal Structure and Seismotectonics of the Val d'Agri Area, Southern Italy, Through Integration of Local Earthquake and Active Seismic Tomographies, and Geological Mapping*. PhD Thesis, Università degli Studi di Napoli Federico II (2007).

11. 11.

Mazzoli, S. et al. Reconstruction of continental margin architecture deformed by the contraction of the Lagonegro Basin, southern Italy. *J. Geol. Soc. Lond.* **158**, 309–319 (2001).

[Article](#) [Google Scholar](#)

12. 12.

Shiner, P., Beccaccini, A. & Mazzoli, S. Thin-skinned versus thick-skinned structural models for Apulian carbonate reservoirs: constrains from the Val d'Agri Fields, S Apennines, Italy. *Mar. Pet. Geol.* **21**, 805–827 (2004).

[CAS Article](#) [Google Scholar](#)

13. 13.

Malinverno, A. & Ryan, W. B. F. Extension in the Tyrrhenian sea and shortening in the Apennines as a result of arc migration driven by sinking of the lithosphere. *Tectonics* **5**, 227–245 (1986).

[ADS](#) [Article](#) [Google Scholar](#)

14. 14.

Dewey, J. F., Helman, M. L., Turco, E., Hutton, D. W. H. & Knott, S. P. in *Alpine Tectonics* (eds Coward, M. P., Dietrich, D. & Park, R. G.) Geol. Soc. Special Publication no. 45, 265–283 (Blackwell Scientific, 1989).

15. 15.

Royden, L. H., Patacca, E. & Scandone, P. Segmentation and configuration of subducted lithosphere in Italy: an important control on thrust-belt and foredeep-basin evolution. *Geology* **15**, 714–717 (1987).

[ADS](#) [Article](#) [Google Scholar](#)

16. 16.

Maschio, L., Ferranti, L. & Burrato, P. Active extension in Val d’Agri area, Southern Apennines, Italy: implications for the geometry of seismogenic belt. *Geophys. J. Int.* **162**, 591–609 (2005).

[ADS](#) [Article](#) [Google Scholar](#)

17. 17.

Devoti, R., Esposito, A., Pietrantonio, G., Pisani, A. R. & Riguzzi, F. Evidence of large scale deformation patterns from GPS data in the Italian subduction boundary. *Earth Planet. Sci. Lett.* **311**, 230–241 (2011).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Silverii, F., D'Agostino, N., Métois, M., Fiorillo, F. & Ventafridda, G. Transient deformation of karst aquifers due to seasonal and multi-year groundwater variations observed by GPS in Southern Apennines (Italy). *J. Geophys. Res. Solid Earth* **121**, 8315–8337 (2016).

[ADS](#) [Article](#) [Google Scholar](#)

19. 19.

Albarello, D., Camassi, R. & Rebez, A. Detection of space and time heterogeneity in the completeness level of a seismic catalogue by a robust statistical approach: an application to the Italian area. *Bull. Seismol. Soc. Am.* **91**, 1694–1703 (2001).

[Article](#) [Google Scholar](#)

20. 20.

Stucci, M., Albini, P., Mirto, C. & Rebez, A. Assessing the completeness of Italian historical earthquake data. *Ann. Geophys.* **47**, 659–673 (2004).

[Google Scholar](#)

21. 21.

Rovida, A. et al. The Italian earthquake catalogue CPTI15. *Bull. Earthquake Eng.* **18**, 2953–2984 (2020).

22. 22.

Valoroso, L. et al. Active faults and induced seismicity in the Val d'Agri area (Southern Apennines, Italy). *Geophys. J. Int.* **178**, 488–502 (2009).

[ADS](#) [Article](#) [Google Scholar](#)

23. 23.

Telesca, L., Giocoli, A., Lapenna, V. & Stabile, T. A. Robust identification of periodic behavior in the time dynamics of short seismic series: the case of seismicity induced by Pertusillo Lake, southern Italy. *Stochastic Environ. Res. Risk Assess.* **29**, 1437–1446 (2015).

[Article](#) [Google Scholar](#)

24. 24.

Plesch, A. et al. Community fault model (CFM) for southern California. *Bull. Seismol. Soc. Am.* **97**, 1793–1802 (2007).

[Article](#) [Google Scholar](#)

25. 25.

Biot, M. A. General theory of three-dimensional consolidation. *J. Appl. Phys.* **12**, 155–164 (1941).

[ADS](#) [MATH](#) [Article](#) [Google Scholar](#)

26. 26.

Coussy, O. *Mechanics of Porous Continua* (Wiley, 1995).

27. 27.

Freed, A. Earthquake triggering by static, dynamic, and postseismic stress transfer. *Annu. Rev. Earth Planet. Sci.* **33**, 335–367 (2005).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Hardebeck, J. L., Nazareth, J. J. & Hauksson, E. The static stress change triggering model: constraints from two southern California aftershock sequences. *J. Geophys. Res. Solid Earth* **103** (B10), 24427–24437 (1998).

[Article](#) [Google Scholar](#)

29. 29.

Aki, K. Generation and propagation of G waves from the Niigata earthquake of June 16, 1964. Part 2. Estimation of earthquake moment, released energy, and stress-strain drop from the G wave spectrum. *Bull. Earthq. Res. Inst.* **44**, 73–78 (1966).

[Google Scholar](#)

30. 30.

Hanks, T. C. & Kanamori, H. A moment magnitude scale. *J. Geophys. Res. Solid Earth* **84** (B5), 2348–2350 (1979).

[Article](#) [Google Scholar](#)

31. 31.

Marone, C. Laboratory-derived friction laws and their application to seismic faulting. *Annu. Rev. Earth Planet. Sci.* **26**, 643–696 (1998).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Linker, M. F. & Dieterich, J. H. Effects of variable normal stress on rock friction: observations and constitutive equations. *J. Geophys. Res. Solid Earth* **97** (B4), 4923–4940 (1992).

[Article](#) [Google Scholar](#)

33. 33.

Dieterich, J. H. Earthquake nucleation on faults with rate- and state-dependent friction. *Tectonophysics* **211**, 115–134 (1992).

[ADS](#) [Article](#) [Google Scholar](#)

34. 34.

Dieterich, J. H. A constitutive law for rate of earthquake production and its application to earthquake clustering. *J. Geophys. Res. Solid Earth* **99**, 2601–2618 (1994).

[Article](#) [Google Scholar](#)

35. 35.

Kaiser, J. Kenntnisse und Folgerungen aus der Messung von Geräuschen bei Zugbeanspruchung von metallischen Werkstoffen. *Arch. Für Isenhütten-Wesen* **24**, 43–45 (1953).

[Article](#) [Google Scholar](#)

36. 36.

Alghannam, M. & Juanes, R. Understanding rate effects in injection-induced earthquakes. *Nat. Commun.* **11**, 3053 (2020).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

37. 37.

DISS Working Group. *Database of Individual Seismogenic Sources (DISS), Version 3.2.1: A compilation of potential sources for earthquakes larger than M 5.5 in Italy and surrounding areas* (Istituto Nazionale di Geofisica e Vulcanologia, accessed 2 January 2019);  
<http://diss.rm.ingv.it/diss/>

38. 38.

Stabile, T. A. et al. Relationship between seismicity and water level of the Pertusillo reservoir (southern Italy). *Boll. Geofis. Teor. Appl.* **56**, 505–517 (2015).

[Google Scholar](#)

39. 39.

Ferranti, L., Maschio, L. & Burrato, P. *Field Trip Guide to Active Tectonics Studies in the High Agri Valley*.  
<http://hdl.handle.net/2122/2749> (Istituto Nazionale di Geofisica e Vulcanologia, 2007).

40. 40.

Stabile, T. A. et al. Evidences of low-magnitude continued reservoir induced seismicity associated with the Pertusillo artificial lake (southern Italy). *Bull. Seismol. Soc. Am.* **104**, 1820–1828 (2014).

[Article](#) [Google Scholar](#)

41. 41.

Aziz, K. & Settari, A. *Petroleum Reservoir Simulation* (Applied Science, 1979).

42. 42.

INTERSECT User Guide/Technical Description v.2016.2 (Schlumberger, 2016).

43. 43.

Aagaard, B. T., Knepley, M. G. & Williams, C. A. A domain decomposition approach to implementing fault slip in finite-element models of quasi-static and dynamic crustal deformation. *J. Geophys. Res. Solid Earth* **118**, 3059–3079 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

44. 44.

ABAQUS User Manual v.2019 (Dassault Systèmes, 2017).

45. 45.

Xu, Y., Cavalcante Filho, J. S., Yu, W. & Sepehrnoori, K. Discrete-fracture modeling of complex hydraulic-fracture geometries in reservoir simulators. *SPE Reservoir Eval. Eng.* **20**, 403–422 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

46. 46.

Cosentino, L. *Integrated Reservoir Studies* (Editions Technip, 2001).

47. 47.

Cucci, L., Pondrelli, S., Frepoli, A., Mariucci, M. T. & Moro, M. Local pattern of stress field and seismogenic sources in Melandro Pergola basin and in Agri valley (Southern Italy). *Geophys. J. Int.* **156**, 575–583 (2004).

[ADS](#) [Article](#) [Google Scholar](#)

48. 48.

Della Vecchia, G., Pandolfi, A., Musso, G. & Capasso, G. An analytical expression for the determination of in situ stress state from borehole data accounting for breakout size. *Int. J. Rock Mech. Min. Sci.* **66**, 64–68 (2014).

[Article](#) [Google Scholar](#)

49. 49.

Chiarabba, C., Jovane, L. & Di Stefano, R. A new view of Italian seismicity using 20 years of instrumental recordings. *Tectonophys.* **395**, 251–268 (2005).

[ADS](#) [Article](#) [Google Scholar](#)

50. 50.

Jha, B. & Juanes, R. Coupled multiphase flow and poromechanics: A computational model of pore pressure effects on fault slip and earthquake triggering. *Wat. Resour. Res.* **50**, 3776–3808 (2014).

[ADS](#) [Article](#) [Google Scholar](#)

51. 51.

Dieterich, J., Cayol, V. & Okubo, P. The use of earthquake rate changes as a stress meter at Kilauea volcano. *Nature* **408**, 457–460 (2000).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

52. 52.

Toda, S., Stein, R. S. & Sagiya, T. Evidence from the ad 2000 Izu islands earthquake swarm that stressing rate governs seismicity. *Nature* **419**, 58–61 (2002).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

53. 53.

Kroll, K. A., Richards-Dinger, K. B., Dieterich, J. H. & Cochran, E. S. Delayed seismicity rate changes controlled by static stress transfer. *J. Geophys. Res. Solid Earth* **122**, 7951–7965 (2017).

[ADS](#) [Article](#) [Google Scholar](#)

54. 54.

Helmstetter, A. & Shaw, B. E. Relation between stress heterogeneity and aftershock rate in the rate-and-state model. *J. Geophys. Res.* **111**, B07304 (2006).

[ADS](#) [Google Scholar](#)

55. 55.

Cochran, E. S., Vidale, J. E. & Tanaka, S. Earth tides can trigger shallow thrust fault earthquakes. *Science* **306**, 1164–1166 (2004).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

56. 56.

Parsons, T., Toda, S., Stein, R. S., Barka, A. & Dieterich, J. H. Heightened odds of large earthquakes near Istanbul: an interaction-based probability calculation. *Science* **288**, 661–665 (2000).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

57. 57.

Segall, P. & Lu, S. Injection-induced seismicity: poroelastic and earthquake nucleation effects. *J. Geophys. Res.* **120**, 5082–5103 (2015).

[ADS](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank L. Impronta for discussions and for providing the high-precision Double Difference events catalogue in Fig. 1; T. Stabile for discussions; M. Miletì for support; G. Roncari, L. Barzaghi, F. Ferulano and A. Orefice for contributions to GPS and seismic data collection and processing; B. Jha for early contributions to the regional geomechanical modelling, including mesh generation; and D. Susanni and L. Magagnini for assistance in obtaining information and organization. We thank Eni and Shell for the authorization to publish this paper and share the data; M. van der Baan and D. Eaton for suggestions that benefitted the manuscript; and A. Puliti for making this study possible.

## Author information

## Affiliations

1. Earth Resources Laboratory, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Bradford H. Hager

2. Department of Earth and Planetary Sciences, University of California, Riverside, Riverside, CA, USA

James Dieterich

3. Institute for Geophysics, University of Texas at Austin, Austin, TX, USA

Cliff Frohlich

4. Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

Ruben Juanes & David Castineira

5. Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Ruben Juanes

6. Development, Operations and Technology, and Exploration, Eni, Milan, Italy

Stefano Mantica, Francesca Bottazzi, Federica Caresani, Alberto Cominelli, Marco Meda, Lorenzo Osculati & Stefania Petroselli

7. Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

John H. Shaw & Andreas Plesch

## Authors

1. Bradford H. Hager

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. James Dieterich

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Cliff Frohlich

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Ruben Juanes

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Stefano Mantica

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. John H. Shaw

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Francesca Bottazzi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Federica Caresani

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. David Castineira

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Alberto Cominelli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Marco Meda

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Lorenzo Osculati

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Stefania Petroselli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Andreas Plesch

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

B.H.H., R.J., D.C., J.H.S., A.P., C.F. and J.D. designed and carried out the initial phase of the project (the regional model) with input from A.C., S.M., M.M. and L.O; S.M., F.B., F.C., A.C. and S.P. improved the project by adding the local flow and geomechanical model, including seismic moment calculations. J.H.S. and A.P. developed the regional and local structural

representation with input from M.M. R.J. and D.C. conducted the regional flow and geomechanical modelling, in collaboration with A.C., S.M., F.B. and F.C. B.H.H. analysed the geodetic data and led the writing. C.F. analysed the seismicity data. J.D. carried out the seismicity rate calculations. J.D., C.F., R.J., S.M., J.H.S., A.C., L.O. and A.P. discussed the results and participated in writing and reviewing the manuscript.

## Corresponding author

Correspondence to [Bradford H. Hager](#).

## Ethics declarations

## Competing interests

Eni initiated a research project to provide an independent assessment of triggered seismicity at the Val d'Agri field based on the most advanced available scientific and technical knowledge. For this purpose, Eni contracted Ramboll Italy S.r.l. to hire the consulting team of J.D., C.F., B.H.H., R.J., A.P. and J.H.S. A.C., S.M., M.M., M. Miletì and L.O. were Eni project references. Eni provided computing resources and technical assistance. The research consulting team submitted a report to Eni addressing activity until the end of 2016. To expand the research scope, Eni's local model—which was developed in parallel—was embedded into the regional model. It was then decided to publish the joint research in the peer-reviewed scientific literature. After their consulting report was completed and presented, the consulting team did not receive further financial support from Eni.

## Additional information

**Peer review information** *Nature* thanks David Eaton, Mirko van der Baan 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 Comparison of injection rates and earthquakes.

Daily injection rate (black solid line) and microseismicity (red circles), 1–30 June 2006. The right *y* axis indicates event magnitude ( $M_L$ )

[Source data.](#)

### Extended Data Fig. 2 Computational mesh for structural model of field and regional faults.

Computational grid of the regional geomechanical model. Top, view of the tetrahedral grid of the computational domain, of approximate dimensions  $80 \times 50 \times 10$  km, with a total of approximately 204,000 cells. Top left, the brown shaded volume indicates the region above the Irpine layers. Top right, a slice through a portion of the grid indicating the grid size within and outside the reservoir. Bottom, view of the gridded faults included in the computational model.

### Extended Data Fig. 3 Representation of faults in the computational mesh.

Top, computational mesh of the local reservoir and geomechanical model. The tetrahedral mesh conforms to the structural model including a complete set of faults, and reservoir and overburden layers. For clarity, a subset of the faults is shown with mesh nodes restricted to a portion of the model. The coloured symbols indicate the pre-production vertical effective stress. Bottom, schematic representation of finite element modelling for the CMF showing node pairs that, starting from an initial condition (left) in which the nodes are superimposed (duplicated), may move separately following the

fault slip (right). The seismic moment can be computed by integrating the nodal slip on the fault.

### **Extended Data Fig. 4 CFF values on Val d'Agri faults.**

Perspective view looking westward of  $\Delta$ CFF (1993–2016) on all model faults. The colour bar is clipped at a  $\Delta$ CFF value of  $\pm 0.1$  MPa to show details of small regions of destabilizing  $\Delta$ CFF at shale–carbonate contacts.

### **Extended Data Fig. 5 Modelled and observed number of earthquakes on the CMF.**

Observed cumulative number of earthquakes (EQs) on the CMF over time (red line) compared with the results of three realizations of the rate–state model (black lines). These models all use the same values for  $\alpha$  and  $\gamma_0$ , but three different pairs of parameters  $\mu$  and  $a$ , providing essentially indistinguishable results. This comparison demonstrates that, although there are large trade-offs among parameters, the resulting forecasts are tightly constrained

[Source data.](#)

### **Extended Data Fig. 6 Well production, reinjection and pressure data.**

Top, historical production and reinjection data for the Val d'Agri field: daily oil (green), gas (red) and water (blue) production data and water reinjection (light blue) in CM2 well. Bottom, shut-in pressure reported at datum-depth (2,400 mTVDssl) for representative wells of Monte Alpi (blue), Monte Enoc (green) and Cerro Falcone (red) culminations.

### **Extended Data Fig. 7 Map of seismic stations in Val d'Agri.**

Locations of the seismic stations operating in the neighbourhood of the Val d'Agri field. For map coordinates, see Fig. 1. The background image is constructed from Copernicus Sentinel data (2017).

## Extended Data Fig. 8 Seismic reflection data near CM2 well defining F10 fault.

Seismic time slice at 2,300 ms from Val d'Agri three-dimensional seismic reflection volume, showing constraints on the F10 fault and the CMF. **a**, Image showing CM2 well and other reservoir faults (RF) included in the regional model. **b**, Interpreted seismic reflections (yellow), and F10 and CMF traces. Note that the primary constraint on the CMF location is from seismicity.

## Extended Data Fig. 9 Modelled and observed reservoir pressures.

Simulated (black curves) and observed (red dots) bottom hole reservoir pressure at representative well locations for the regional model (left) and the local model (right)

[Source data.](#)

## Extended Data Fig. 10 Modelled and observed ground displacements.

Comparison of  $X$ (top row),  $Y$  (middle row), and  $Z$  (bottom row) components of model predictions (lines) and GPS estimates (symbols) of relative displacements between GPS sites SIRI (left column), MTSN (middle column) and MCEL (right column) and reference site PTRP. GPS site locations are shown in Fig. 1. Predictions for our preferred model, with reservoir Biot coefficient 0.1, are given by the black lines; predictions for an alternative model with reservoir Biot coefficient 0.3 are given by the blue lines. Displacements are in the model coordinate system; the lateral displacements are projected along the  $x$  axis (positive to the northeast) and along the  $y$  axis (positive to the northwest)

[Source data.](#)

## Extended Data Table 1 Bulk poromechanical properties [Full size table](#)

# 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. 5](#)

[Source Data Extended Data Fig. 9](#)

[Source Data Extended Data Fig. 10](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Hager, B.H., Dieterich, J., Frohlich, C. *et al.* A process-based approach to understanding and managing triggered seismicity. *Nature* **595**, 684–689 (2021). <https://doi.org/10.1038/s41586-021-03668-z>

## Download citation

- Received: 26 March 2020
- Accepted: 24 May 2021
- Published: 28 July 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03668-z>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

## Associated Content

### [Earthquakes triggered by underground fluid injection modelled for a tectonically active oil field](#)

- Mirko van der Baan

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03668-z>

- Article
- [Published: 14 July 2021](#)

# Dynamical prefrontal population coding during defensive behaviours

- [Daniel Jercog](#) [ORCID: orcid.org/0000-0003-3849-9196](#)<sup>1,2 na1</sup>,
- [Nanci Winke](#)<sup>1,2 na1</sup>,
- [Kibong Sung](#)<sup>1,2 nAff4</sup>,
- [Mario Martin Fernandez](#) [ORCID: orcid.org/0000-0002-4228-5847](#)<sup>1,2</sup>,
- [Claire Francioni](#)<sup>1,2</sup>,
- [Domitille Rajot](#)<sup>1,2</sup>,
- [Julien Courtin](#) [ORCID: orcid.org/0000-0002-6454-3245](#)<sup>1,2 nAff5</sup>,
- [Fabrice Chaudun](#)<sup>1,2 nAff6</sup>,
- [Pablo E. Jercog](#) [ORCID: orcid.org/0000-0001-7852-5884](#)<sup>3</sup>,
- [Stephane Valerio](#)<sup>1,2 nAff7</sup> &
- [Cyril Herry](#) [ORCID: orcid.org/0000-0003-2973-0843](#)<sup>1,2</sup>

[Nature](#) volume 595, pages 690–694 (2021) [Cite this article](#)

- 4345 Accesses
- 55 Altmetric
- [Metrics details](#)

## Subjects

- [Neuroscience](#)
- [Prefrontal cortex](#)

## Abstract

Coping with threatening situations requires both identifying stimuli that predict danger and selecting adaptive behavioural responses to survive<sup>1</sup>. The dorsomedial prefrontal cortex (dmPFC) is a critical structure that is involved in the regulation of threat-related behaviour<sup>2,3,4</sup>. However, it is unclear how threat-predicting stimuli and defensive behaviours are associated within prefrontal networks to successfully drive adaptive responses. Here we used a combination of extracellular recordings, neuronal decoding approaches, pharmacological and optogenetic manipulations to show that, in mice, threat representations and the initiation of avoidance behaviour are dynamically encoded in the overall population activity of dmPFC neurons. Our data indicate that although dmPFC population activity at stimulus onset encodes sustained threat representations driven by the amygdala, it does not predict action outcome. By contrast, transient dmPFC population activity before the initiation of action reliably predicts avoided from non-avoided trials. Accordingly, optogenetic inhibition of prefrontal activity constrained the selection of adaptive defensive responses in a time-dependent manner. These results reveal that the adaptive selection of defensive responses relies on a dynamic process of information linking threats with defensive actions, unfolding within prefrontal networks.

## 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

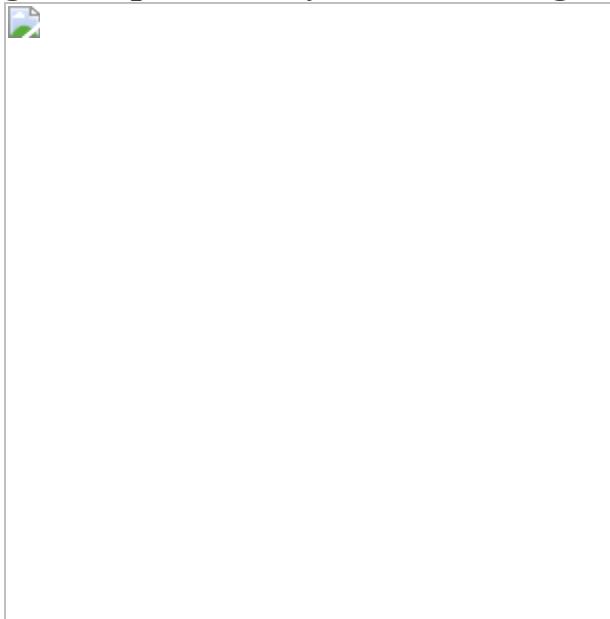
[Rent or Buy](#)

All prices are NET prices.

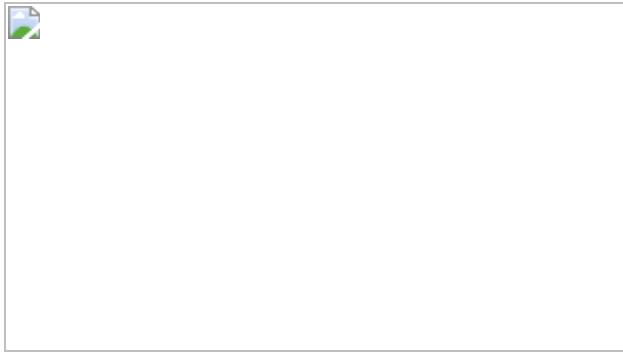
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

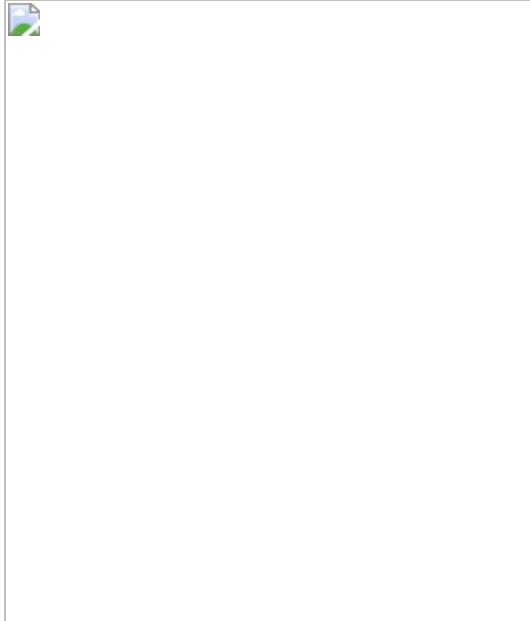
**Fig. 1: Population dynamics during active avoidance behaviour.**



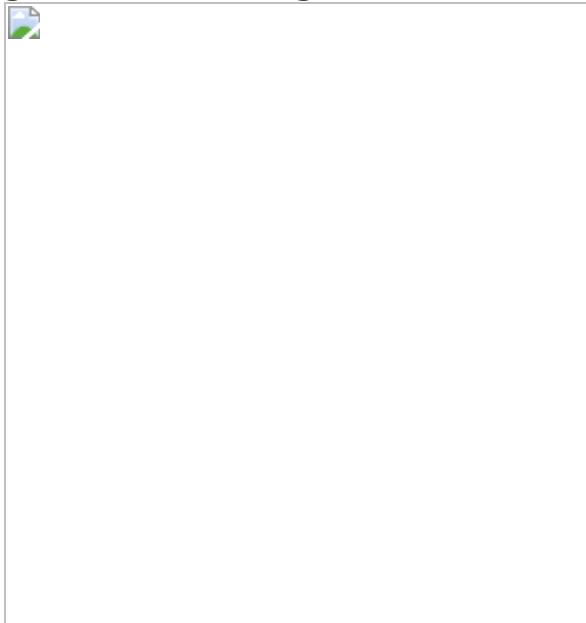
**Fig. 2: dmPFC threat representations are amygdala-dependent.**



**Fig. 3: dmPFC encodes upcoming avoidance actions.**



**Fig. 4: dmPFC integrates threat information to drive defensive actions.**



## Data availability

The data that support the findings of this study are available from the corresponding authors on request.

## Code availability

The custom-written codes used to analyse data from this study are available from the corresponding authors on request.

## References

1. 1.

Pearce, J. M. & Bouton, M. E. Theories of associative learning in animals. *Annu. Rev. Psychol.* **52**, 111–139 (2001).

[CAS](#) [Article](#) [Google Scholar](#)

2. 2.

Sotres-Bayon, F. & Quirk, G. J. Prefrontal control of fear: more than just extinction. *Curr. Opin. Neurobiol.* **20**, 231–235 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Herry, C. & Johansen, J. P. Encoding of fear learning and memory in distributed neuronal circuits. *Nat. Neurosci.* **17**, 1644–1654 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Klavir, O., Prigge, M., Sarel, A., Paz, R. & Yizhar, O. Manipulating fear associations via optogenetic modulation of amygdala inputs to

prefrontal cortex. *Nat. Neurosci.* **20**, 836–844 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Laubach, M., Wessberg, J. & Nicolelis, M. A. Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature* **405**, 567–571 (2000).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Ährlund-Richter, S. et al. A whole-brain atlas of monosynaptic input targeting four different cell types in the medial prefrontal cortex of the mouse. *Nat. Neurosci.* **22**, 657–668 (2019).

[Article](#) [Google Scholar](#)

7. 7.

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J. & Rushworth, M. F. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).

[CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Saez, A., Rigotti, M., Ostojic, S., Fusi, S. & Salzman, C. D. Abstract context representations in primate amygdala and prefrontal cortex. *Neuron* **87**, 869–881 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Salay, L. D., Ishiko, N. & Huberman, A. D. A midline thalamic circuit determines reactions to visual threat. *Nature* **557**, 183–189 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Hayden, B. Y. & Platt, M. L. Neurons in anterior cingulate cortex multiplex information about reward and action. *J. Neurosci.* **30**, 3339–3346 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Do-Monte, F. H., Quiñones-Laracuente, K. & Quirk, G. J. A temporal shift in the circuits mediating retrieval of fear memory. *Nature* **519**, 460–463 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Taub, A. H., Perets, R., Kahana, E. & Paz, R. Oscillations synchronize amygdala-to-prefrontal primate circuits during aversive learning. *Neuron* **97**, 291–298 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Courtin, J. et al. Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* **505**, 92–96 (2014).

[ADS](#) [Article](#) [Google Scholar](#)

14. 14.

Dejean, C. et al. Prefrontal neuronal assemblies temporally control fear behaviour. *Nature* **535**, 420–424 (2016).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Karalis, N. et al. 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. *Nat. Neurosci.* **19**, 605–612 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Likhtik, E., Stujenske, J. M., Topiwala, M. A., Harris, A. Z. & Gordon, J. A. Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. *Nat. Neurosci.* **17**, 106–113 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Burgos-Robles, A. et al. Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. *Nat. Neurosci.* **20**, 824–835 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Diehl, M. M. et al. Active avoidance requires inhibitory signaling in the rodent prelimbic prefrontal cortex. *eLife* **7**, e34657 (2018).

[Article](#) [Google Scholar](#)

19. 19.

Bravo-Rivera, C., Roman-Ortiz, C., Brignoni-Perez, E., Sotres-Bayon, F. & Quirk, G. J. Neural structures mediating expression and extinction

of platform-mediated avoidance. *J. Neurosci.* **34**, 9736–9742 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Capuzzo, G. & Floresco, S. B. Prelimbic and infralimbic prefrontal regulation of active and inhibitory avoidance and reward-seeking. *J. Neurosci.* **40**, 4773–4787 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Yu, B. M. et al. Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *J. Neurophysiol.* **102**, 614–635 (2009).

[Article](#) [Google Scholar](#)

22. 22.

Rozeske, R. R. et al. Prefrontal-periaqueductal gray-projecting neurons mediate context fear discrimination. *Neuron* **97**, 898–910.e6 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Legendre, P. & Legendre, L. *Numerical Ecology (Developments in Environmental Modelling)* (Elsevier, 1998).

24. 24.

Rigotti, M. et al. The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Choi, J. S., Cain, C. K. & LeDoux, J. E. The role of amygdala nuclei in the expression of auditory signaled two-way active avoidance in rats. *Learn. Mem.* **17**, 139–147 (2010).

[Article](#) [Google Scholar](#)

26. 26.

Quirk, G. J., Repa, C. & LeDoux, J. E. Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. *Neuron* **15**, 1029–1039 (1995).

[CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Duvarci, S. & Pare, D. Amygdala microcircuits controlling learned fear. *Neuron* **82**, 966–980 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Pape, H. C. & Pare, D. Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. *Physiol. Rev.* **90**, 419–463 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Goosens, K. A., Hobin, J. A. & Maren, S. Auditory-evoked spike firing in the lateral amygdala and Pavlovian fear conditioning: mnemonic code or fear bias? *Neuron* **40**, 1013–1022 (2003).

[CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Senn, V. et al. Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron* **81**, 428–437 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Sotres-Bayon, F., Sierra-Mercado, D., Pardilla-Delgado, E. & Quirk, G. J. Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. *Neuron* **76**, 804–812 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Poremba, A. & Gabriel, M. Amygdalar lesions block discriminative avoidance learning and cingulothalamic training-induced neuronal plasticity in rabbits. *J. Neurosci.* **17**, 5237–5244 (1997).

[CAS](#) [Article](#) [Google Scholar](#)

33. 33.

Sierra-Mercado, D., Padilla-Coreano, N. & Quirk, G. J. Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. *Neuropsychopharmacology* **36**, 529–538 (2011).

[Article](#) [Google Scholar](#)

34. 34.

Ramirez, F., Moscarello, J. M., LeDoux, J. E. & Sears, R. M. Active avoidance requires a serial basal amygdala to nucleus accumbens shell circuit. *J. Neurosci.* **35**, 3470–3477 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

35. 35.

LeDoux, J. E., Moscarello, J., Sears, R. & Campese, V. The birth, death and resurrection of avoidance: a reconceptualization of a troubled paradigm. *Mol. Psychiatry* **22**, 24–36 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

36. 36.

Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

37. 37.

Franklin, K. B. J. & Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates* (Academic, 2001).

38. 38.

Meyers, E. M., Qi, X. L. & Constantinidis, C. Incorporation of new information into prefrontal cortical activity after learning working memory tasks. *Proc. Natl Acad. Sci. USA* **109**, 4651–4656 (2012).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank D. Salzman, A. Luthi, K. Benchenane, S. Bagur and R. Rozeske for discussions and reading of the manuscript; A. Sant Ana and S. Khoder for preliminary experiments; T. Stefanelli for electrode building assistance; K. Deisseroth and E. Boyden for sharing material; and S. Laumond, J. Tessaire and the technical staff of the housing and experimental animal

facility of the Neurocentre Magendie. Microscopy was performed in the Bordeaux Imaging Center of the CNRS-INSERM and Bordeaux University, a member of France BioImaging. This work was supported by grants from the French National Research Agency (ANR-DOPAFEAR, ANR-PRELONGIN, ANR-10-EQPX-08 OPTOPATH), the Conseil Regional d'Aquitaine, and the Fondation pour la Recherche Médicale (FRM-Equipes FRM 2017).

## Author information

### Author notes

#### 1. Kibong Sung

Present address: Division of Integrative Biosciences and Biotechnology (IBB), Pohang University of Science and Technology (POSTECH), Pohang, Korea

#### 2. Julien Courtin

Present address: Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

#### 3. Fabrice Chaudun

Present address: Department of Basic Neurosciences, Faculty of Medicine, University of Geneva, Geneva, Switzerland

#### 4. Stephane Valerio

Present address: Aquineuro SAS, Bordeaux, France

#### 5. These authors contributed equally: Daniel Jercog, Nanci Winke

## Affiliations

#### 1. Univ. Bordeaux, Neurocentre Magendie, U1215, Bordeaux, France

Daniel Jercog, Nanci Winke, Kibong Sung, Mario Martin Fernandez, Claire Francioni, Domitille Rajot, Julien Courtin, Fabrice Chaudun, Stephane Valerio & Cyril Herry

2. INSERM, Neurocentre Magendie, U1215, Bordeaux, France

Daniel Jercog, Nanci Winke, Kibong Sung, Mario Martin Fernandez, Claire Francioni, Domitille Rajot, Julien Courtin, Fabrice Chaudun, Stephane Valerio & Cyril Herry

3. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Hospital Clínic, Universitat de Barcelona, Barcelona, Spain

Pablo E. Jercog

## Authors

1. Daniel Jercog

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Nanci Winke

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Kibong Sung

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Mario Martin Fernandez

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Claire Francioni

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Domitille Rajot

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Julien Courtin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Fabrice Chaudun

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Pablo E. Jercog

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Stephane Valerio

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Cyril Herry

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

D.J., N.W. and K.S. performed electrophysiological recordings. D.J. and N.W. performed optogenetics and pharmacological experiments. D.J., N.W. and C.F. performed histology. D.J., N.W., C.F. and D.R. performed behavioural experiments on nonimplanted mice. M.M.F. performed preliminary viral injections. J.C. and F.C. provided tone fear conditioning

data. D.J., S.V. and C.H. designed the experiments. D.J. and P.E.J. designed analyses. D.J. wrote software codes. D.J. and C.H. analysed the data and wrote the paper.

## Corresponding authors

Correspondence to [Daniel Jercog](#) or [Cyril Herry](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Andrew Huberman, Joshua Johansen 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 Discriminative two-way active avoidance behaviour.](#)

**a**, The number of trials until first CS<sup>+</sup> NAV trial significantly increases until day 3 (\*\*P < 0.001 for CS<sup>+</sup>, two-way repeated measures (RM) ANOVA) (left). Avoidance latencies (centre) or non-avoidance (right) latencies were unaltered after day 3 (\*P < 0.05 for CS<sup>+</sup>, two-way RM ANOVA; n = 34 mice). **b**, Distribution of avoidance latencies across days for CS<sup>-</sup> (left) and CS<sup>+</sup> (right) trials. Vertical dashed lines show distribution means per day. **c**, Freezing was rapidly expressed after training started (left). Mean freezing

across days significantly differed between CS<sup>-</sup> and CS<sup>+</sup> on day 3 and 4 (right) (\* $P = 0.040$  \*\* $P = 0.0075$ , two-way RM ANOVA;  $n = 34$  mice). **d**, Detection sensitivity ( $d'$ , top) and response bias ( $d'$  criterion, bottom) showing that high discrimination between CS<sup>-</sup> and CS<sup>+</sup> achieved at day 2; response bias is only observable at day 1 ( $d'$  criterion  $>0$  bias towards NAV). **e**, Extinction training (nonimplanted mice) for short (4 days) (top) ( $P < 0.001$  for 1st block of day 5, two-way RM ANOVA;  $n = 8$  mice) and long training (11 days) (bottom) ( $P < 0.001$  for 1st–3rd and 7th block, two-way RM ANOVA;  $n = 8$  mice). **f**, Contingency degradation, US (1-s maximum) delivered during inter-trial intervals (\* $P < 0.05$  \*\* $P < 0.01$  \*\*\* $P < 0.001$ , two-way RM ANOVA;  $n = 8$  mice). **g**, Permanence map during the entire sessions across days for an example mouse, showing the time spent in the different regions of the shuttlebox (colour axis in log scale). **h**, Probability distributions along the maze across days. **i**, Trajectories for different trial types and response outcomes for the example mouse shown in **g** at day 4. White circles indicate the location of the mouse at CS onset, grey circles indicate the location after 7 s (AV (left) and NAV (right); CS<sup>-</sup> (top), CS<sup>+</sup> (bottom)). **j**, Mouse spatial location at CS onset for AV (top) and NAV (bottom) trials does not predict avoidance behaviour ( $n = 68$  sessions;  $P = 0.548$ , two-sampled Kolmogorov–Smirnov test of both cumulative distributions for  $x$ -axis position; day 3 and day 4 sessions). Learning curves show five-trial blocks. Error bars showing s.e.m. Error shaded areas are 95% confidence interval (CI).

## **Extended Data Fig. 2 dmPFC pharmacological inactivation and speed-related analyses.**

**a**, Bilateral cannula tip position ( $n = 12$  mice). **b**, Protocol for pharmacological inactivation experiments. **c**, Avoidance probability for CS<sup>-</sup> (left) and CS<sup>+</sup> (right) during training for muscimol ( $n = 7$  mice) and vehicle ( $n = 5$  mice) groups. **d**, Number of infrared beam breaks 2 min before session starts during pre-infusion, infusion and post-infusion sessions for muscimol and vehicle groups, showing that muscimol infusion did not affect basal levels of locomotion ( $P > 0.9999$ , RM mixed-effects model). **e**, CS<sup>+</sup> active avoidance was impaired upon dmPFC inactivation (\*\* $P = 0.0285$ , two-way RM ANOVA; outliers indicated as dots) whereas

CS<sup>-</sup> avoidance probability did not change ( $P = 0.0761$ , two-way RM ANOVA). **f**, Mean freezing probability dynamics during CS<sup>-</sup>, CS<sup>+</sup> AV and CS<sup>+</sup> NAV trials. **g**, Percentage freezing during sound presentation for CS<sup>-</sup>, CS<sup>+</sup> AV and CS<sup>+</sup> NAV trials on day 3 and 4 (violin plot median and 25th–75th percentiles; \*\*\* $P < 0.001$  one-way RM ANOVA;  $n = 68$  sessions;  $n = 34$  mice). **h**, CS-onset-triggered median speed for CS<sup>-</sup>, CS<sup>+</sup> AV and CS<sup>+</sup> NAV trials during the habituation session (left) and during training day 3 and 4 (right). For day-3–4 sessions, trials were only included in the analyses if the AV shuttle time was a minimum of 2.5 s. **i**, CS-onset-triggered median speed for CS<sup>+</sup> AV trials grouped by different AV latencies. In all cases a drop in speed at CS onset was observed, preceding the sharp increase in speed associated with avoidance responses. **j**, Schematics of AV-run alignment definition. Instantaneous acceleration is computed for CS<sup>+</sup> AV individual trials and AV-run initiation is defined by the first event in which acceleration surpasses a threshold (50 cm per s<sup>-2</sup>; minimum AV-run time of 2.5 s were included in the analyses). The same alignment was used for CS<sup>+</sup> NAV (and CS<sup>-</sup>) trials, by randomly assigning AV-run times from CS<sup>+</sup> AV trials. **k**, Individual AV-run-aligned speed traces from CS<sup>+</sup> AV and CS<sup>+</sup> NAV trials (top). AV-run and avoidance (hurdle crossing) time distributions for CS<sup>+</sup> AV trials (bottom-left). Time difference between AV-run initiation and hurdle crossing distribution (bottom right). **l**, Median speed at CS-onset-aligned (left) and AV-run-aligned (right) conditions. AV-run alignment for both CS<sup>+</sup> NAV and CS<sup>-</sup> defined by sampling CS<sup>+</sup> AV-run times (Methods). **m**, Learning curves during training in which the tones used for CS<sup>-</sup> and CS<sup>+</sup> were switched (left; \*\*\* $P < 0.0001$  for 7th block forward, two-way RM ANOVA;  $n = 9$  mice). CS-onset-triggered median speed for different trials type at day 3 and day 4 for switched tones (right). Error bars are s.e.m. Error shaded areas are 95% CI.

### **Extended Data Fig. 3 In vivo single-unit recordings in the dmPFC.**

**a**, Representative electrolytic lesion (left) and electrode tip locations for recorded mice (centre,  $n = 34$  mice). Example of single unit recorded in the dmPFC showing the average waveforms, isolation in the principal

component (PC) space and autocorrelogram (right). **b**, Area under the curve (AUC) and trough-to-peak distance were used to separate units into putative excitatory (exc.) and inhibitory (inh.), using *k*-means clustering ( $k = 2$ ). Individual classified waveforms and percentages shown. **c**, Violin plots displaying average firing rates across the entire session for individual units during habituation (day 0, 872 units) and at high levels of avoidance (day 4, 711 units) do not show significant differences ( $P = 0.3519$ , unpaired *t*-test). **d**, Z-scored neuronal activity for CS<sup>+</sup> (left) and CS<sup>-</sup> (right) trials ordered by the magnitude of CS<sup>-</sup> responses (same data as in Fig. [1d](#); day 3 and 4 data pooled, 68 sessions,  $n = 1,261$  units). **e**, Heat maps displaying *z*-scores of the population of recorded units for different trial types ( $n = 34$  sessions, 643 units displayed; units with undefined *z*-scores ( $n = 22$  units) were removed). **f**, Percentage of total units ( $n = 643$  units) significantly responding to CS<sup>+</sup> or CS<sup>-</sup> during the first 7 s after CS onset. For each category, the percentage of units showing excitatory and inhibitory responses (these are not mutually exclusive categories, as a unit can display both excited and inhibited responses). **h**, Average spectrograms of dmPFC local field potentials (LFP) during spontaneous freezing and nonfreezing periods ( $n = 14$  mice). **i**, Power spectra during nonfreezing and freezing periods. Inset displays ‘4 Hz’ mean signal-to-noise ( $P = 0.0015$ , Wilcoxon test; error bar are s.e.m.). **j**, During the active avoidance task, average spectrograms during different trial types and responses. **k**, Corresponding power spectra for different trial types and responses together with pre-CS periods. **l**, Average spectrogram differences highlighting the changes for different trial types. **m**, Power spectra for different trial types. Shaded error areas are 95% CI.

## Extended Data Fig. 4 Neuronal activity at AV-run alignment.

**a**, Examples of different unit response patterns observed at CS-onset- and AV-run-aligned activity (for each case: raster on top, average firing rate and speed on bottom). Each column shows examples of different observed patterns. **b**, Z-scored activity for different unit responses during CS<sup>+</sup> AV (left), CS<sup>+</sup> NAV (centre) and CS<sup>-</sup> (right) trials, ordered by the magnitude of CS<sup>+</sup> AV responses (here we considered sessions with at least 6 trials of each type;  $n = 33$  sessions, 611 units displayed; units with undefined *z*-scores

( $n = 23$  units) were removed. **c**, Average firing rates across units for different trial type. **d**, Average  $z$ -score across units for different trial types. **e**, Mean prediction error of GPFA as a function of the number of latent dimensions considered (following a previous publication<sup>21</sup>) for CS-onset- (top) and AV-run- (bottom) aligned conditions (error bars are s.e.m.). Shaded error areas are 95% CI.

## Extended Data Fig. 5 CS decoding from dmPFC neuronal populations.

**a**, Tone fear conditioning protocol. **b**, During habituation (day 0), mice ( $n = 38$  mice) exhibited low mean freezing during CS<sup>−</sup> and CS<sup>+</sup>. Following fear conditioning (FC) (day 1, mice  $n = 38$  mice), CS<sup>+</sup> (CS<sup>+</sup> 1–12 grouped in blocks of 4) induced high mean freezing in comparison to CS<sup>−</sup> ( $P < 0.0001$ , one-way RM ANOVA;  $n = 38$  mice). **c**, Freezing probability as a function of time during 4 CS<sup>−</sup> and 4 (first) CS<sup>+</sup>. **d**, Decoding accuracy of CS<sup>+</sup> (top) and CS<sup>−</sup> (bottom) from baseline population activity for tone-fear-conditioned mice on habituation (day 0,  $n = 553$  units;  $n = 38$  sessions). **e**, As in **d**, but post-FC (day 1,  $n = 582$  units;  $n = 38$  sessions). **f**, Mean avoidance probability during extinction learning ( $n = 14$  mice). Although CS<sup>+</sup> avoidance behaviour (top) was extinguished after the first extinction session (\*\* $P < 0.0110$  for 1st and 3rd block, two-way RM ANOVA), CS<sup>+</sup> from baseline decoding accuracy (bottom) was reduced only on the subsequent extinction session. **g**, Extinction training avoidance and shuttling mean probability during 12 s (maximal CS duration, in dark colours) or 7 s (as in **f** top, in light colours) after CS onset. After the first half of the first extinction session, CS<sup>+</sup> shuttling decreases and CS<sup>−</sup> increases reaching similar values to those observed during the habituation session ( $n = 14$  mice). **h**, Average CS<sup>+</sup>-evoked freezing during the first (day 6) (top) and second (day 7) (bottom) extinction session, for the first (light red) and last (dark red) half of the session. **i**, Mean avoidance probability during reversal learning ( $n = 11$  mice). Previous CS<sup>+</sup> tone became neutral (reversed CS<sup>−</sup> (rCS<sup>−</sup>)) and the previous CS<sup>−</sup> tone was associated with the US (rCS<sup>+</sup>). Avoidance behaviour was progressively reversed during training (top) (\*\* $P < 0.01$  \* $P < 0.05$  difference between rCS<sup>−</sup> and rCS<sup>+</sup>, two-way RM ANOVA) and decoding accuracies changed accordingly (bottom). **j**,

Assessing safety signal properties on CS<sup>-</sup>. Two groups of mice ( $n = 8$  mice each) were trained in normal protocol until day 4 (top). Retardation test was performed during the 3 subsequent days in one of the groups (group A) (bottom left) by pairing CS<sup>-</sup> with the US, the transfer of active avoidance behaviour matched the initial AV probability to the CS<sup>+</sup> only at day 7. ns,  $P = 0.0691$  for CS<sup>-</sup> versus CS<sup>+</sup> at day 3, two-way RM ANOVA.

Additionally, to test the effect of a new association in the learning rate in another group of mice (group B) we paired a novel tone with the US (bottom right), which led to a rapid transfer of active avoidance behaviour (ns,  $P = 0.2133$  for CS<sup>x</sup> versus CS<sup>+</sup> day 3, two-way RM ANOVA). Our data suggest that CS<sup>-</sup> acquired safety properties ('CS<sup>-</sup> with no US' association) during the initial training. Significant decoding accuracy periods over shuffle accuracies were represented by thick lines ( $P < 0.05$ , permutation test; Methods). Error bars are s.e.m. Accuracy data are mean  $\pm 1$  s.d.

### **Extended Data Fig. 6 Confined task and threat information decoding at CS onset.**

**a**, After training, mice were confined to one of the compartments where unreinforced conditioned stimuli were presented (confined task;  $n = 21$  mice). Median speed dynamics in the confined task showed decreased speed selective to CS<sup>+</sup> trials. **b**, Mean freezing probability dynamics during CS<sup>-</sup> and CS<sup>+</sup> trials in the confined task. **c**, Freezing to CS<sup>+</sup> progressively increases during the first 3 CS<sup>+</sup> trials (5th trial) and reaches stable levels from trial 8 to 12 (\*\* $P < 0.001$ , one-way RM ANOVA;  $n = 21$  mice). **d**, Average spectrograms for different trial types during the confined task ( $n = 11$  mice). **e**, Spectrogram difference between CS<sup>+</sup> and CS<sup>-</sup>. **f**, Power spectra for CS<sup>+</sup> and CS<sup>-</sup> in the confined task. Inset displays mean '4 Hz' signal-to-noise ( $P = 0.02$ , Wilcoxon test). **g**, Freezing probabilities when mice exhibit freezing (CS<sup>+</sup> FzOns) or nonfreezing (CS<sup>+</sup> NFzOns) at CS onset, light red area indicates the period in which mice showed pure freezing or nonfreezing bouts, respectively (left) ( $n = 13$  sessions) (Methods). Decoding CS<sup>+</sup> FzOns and CS<sup>+</sup> NFzOns by dmPFC activity (right). **h**, Decoding spontaneous freezing (outside CS) from nonfreezing episodes (minimum episode durations of 4 s) (left). Decoding spontaneous freezing (outside CS) and CS<sup>+</sup> (right). **i**, Decoding accuracy of CS<sup>+</sup> (top)

and CS<sup>-</sup> (bottom) from baseline population activity during the confined task. **j**, Trained decoders to classify CS<sup>+</sup> trials from baseline in the active avoidance task (CS<sup>+</sup><sub>AA</sub>) displayed high accuracy when decoding CS<sup>+</sup> in the confined task (CS<sup>+</sup><sub>CNF</sub>). **k**, Trained decoders to classify CS<sup>+</sup> trials from baseline in the confined task (CS<sup>+</sup><sub>CNF</sub>) displayed high accuracy when decoding CS<sup>+</sup> in the active avoidance task (CS<sup>+</sup><sub>AA</sub>). To compare the information across panels, decoding was performed by randomly sampling 250 units in each case. Significant decoding accuracy periods over shuffle accuracies are represented by thick lines ( $P < 0.05$ , permutation test) (Methods). Error bars are s.e.m. Error shaded areas are 95% CI. Accuracy data are mean  $\pm$  1 s.d.

### **Extended Data Fig. 7 Influence of unit amount and population versus single-mouse CS decoding.**

**a**, Decoding accuracy of CS<sup>+</sup> from baseline activity for individual units from day 3 and day 4. Average accuracy across single units displayed in light grey. **b**, Cumulative distribution of mean single-unit decoding accuracies during 4 s after CS onset. **c**, Effect of the number of units in the amount of decoding accuracy considering single-unit, multi-unit (as the summation of single-units) and the multidimensional population vector. We refer to the population vector as the pseudopopulation vector that is composed of the firing rate from individual units (pooled from different sessions). **d**, Mean decoding accuracy during 4 s after CS onset across days on which the population vector was formed by pooling units from different mice (250 units randomly sampled for each of the bootstrap repetition; bootstrap 200 repetitions, shuffled label 1,000 repetitions, fivefold cross-validation; \* $P < 0.05$ , permutation test; Error bars represent 1 s.d.). **e**, Mean decoding accuracies in which population vector is formed by units from individual mice ( $n = 14$  mice; for each mouse, bootstrap 200 repetitions, shuffled label 1,000 repetitions, fivefold cross-validation; real and shuffle accuracies were pooled across mice to assess statistical significance, \* $P < 0.05$ , permutation test; error bars represent 1 s.e.m.). Dashed horizontal lines indicate accuracy chance level.

## Extended Data Fig. 8 dmPFC threat representations depend on amygdala processing.

**a**, Cannula tip placement for behaviour-only experiments targeting the BLA ( $n = 11$  mice). **b**, Mice received training until day 3. On day 4, 1 h before the session started, mice were infused with vehicle ( $n = 5$  mice) or muscimol ( $n = 6$  mice). **c**, Avoidance probability for CS<sup>-</sup> (top) and CS<sup>+</sup> (bottom) during habituation and training for both muscimol and vehicle groups (ns,  $P > 0.9999$ , RM mixed-effects model). **d**, Number of infrared beam breaks 2 min before session starts during pre-infusion (Pre), infusion (Infus.) and post-infusion (Post) sessions, showing that muscimol did not affect basal levels of locomotion on infusion session (but changed the post-infusion session; \*\* $P = 0.0083$ , RM mixed-effects model). **e**, CS<sup>+</sup> active avoidance was impaired upon BLA muscimol inactivation (\*\* $P < 0.0001$ , RM mixed-effects model; outliers indicated as dots) whereas CS<sup>-</sup> avoidance probability did not change (ns,  $P = 0.9424$ , RM mixed-effects model) **f**, Cannula tip location targeting the BLA and electrode implantation site in the dmPFC ( $n = 11$  mice). **g**, Muscimol infusion (Mus) in the BLA impairs active avoidance behaviour to CS<sup>+</sup> (\*\* $P < 0.0001$ , \*\* $P < 0.001$ , one-way RM ANOVA) but not CS<sup>-</sup> (ns,  $P = 0.2929$ , one-way RM ANOVA). **h**, CS decoding accuracies from dmPFC activity, obtained during the pre-session (day 3). **i**, CS decoding accuracy from dmPFC activity during BLA muscimol inactivation. Despite the fact that decoding accuracies in between pips are markedly reduced, we still decode the auditory sensory stimuli at pip times. **j**, CS decoding accuracy from dmPFC activity the post-session after BLA inactivation returns to previous levels of decoding accuracy. **k**, Neuronal population trajectories for CS<sup>+</sup> and CS<sup>-</sup> presentations during Pre (day 3) (left) and Mus (day 4) (centre) sessions projected onto the first 3 latent dimensions (30 trials for each trial type shown). Circles indicate time zero. Arrows schematize trajectory flow. Mean  $\pm$  s.d. of normalized distance ( $d$ ) between CS<sup>+</sup> and CS<sup>-</sup> for Pre (light blue) and Mus (dark blue) conditions. Thick lines indicate significantly increased normalized distance ( $P < 0.05$ , permutation test). Horizontal coloured lines indicate periods in which mean accuracy is significantly different from those observed from the shuffle accuracy distribution ( $P < 0.05$ , permutation test) (Methods). To compare

the information across panels, decoding was performed by randomly sampling 250 units for each bootstrap repetition. Error bars are s.e.m. Accuracy data are mean  $\pm$  1 s.d.

## **Extended Data Fig. 9 Avoidance action coding by dmPFC populations.**

**a**, To preclude the possibility that a single session inordinately contributed to the observed surge in AV versus NAV decoding accuracy before action initiation, a jackknife procedure was used by repeatedly removing an individual session ( $n = 33$  sessions in total) to compute the decoding accuracy at AV-run-aligned condition. **b**, Distributions of the number of CS<sup>+</sup> NAV and CS<sup>+</sup> AV trials for different sessions (Methods). **c**, Number of sessions and units as a function of the minimum number of trials available for both the CS<sup>+</sup> NAV and CS<sup>+</sup> AV trial types. **d**, Decoding CS<sup>+</sup> NAV versus CS<sup>+</sup> AV using a different number of minimum trials for session inclusion criteria. **e**, Mean (top) and maximum (bottom) accuracy from  $-1$  to 0 s to AV-run for the different number of trials for session inclusion criteria (error bars display s.e.). **f**, AV-run-triggered snout, centre-of-mass and tail average speed for CS<sup>+</sup> AV (left) and CS<sup>+</sup> NAV (right) trials. **g**, CS<sup>+</sup> AV and CS<sup>+</sup> NAV speed discrimination index  $d'$ . **h**, dmPFC neural activity and speed-based decoding accuracy for different body parts (as in Fig. 3). **i**, Transition from freezing to nonfreezing is not preceded by a long time-scale surge in decoding accuracy in dmPFC populations. Spontaneous freezing episodes (longer than 4 s) were both aligned to freezing-to-nonfreezing transition times ( $F_z > NF_z$ ) or 2 s before this freezing-to-freezing transition ( $F_z > F_z$ ) (left) (Methods). Median speed for  $F_z > NF_z$  and  $F_z > F_z$  trials (centre top) and discrimination index  $d'$  (centre bottom). dmPFC neural activity based  $F_z > NF_z$  vs  $F_z > F_z$  decoding accuracy (dark grey) is compared with AV versus NAV decoding accuracy (light grey) in AV-run-aligned condition (right). **j**, Heat map displaying z-score of CS<sup>+</sup> evoked activity for single units ordered by the magnitude of the response (left). For each unit and in the same order, the corresponding mean feature weight (2 s after CS onset) in the linear classifiers for the CS<sup>+</sup>/baseline activity. **k**, Mean feature weight for CS<sup>+</sup>/baseline versus CS<sup>+</sup> AV/NAV before action initiation for individual units were not correlated ( $r = 0.0069$ ,  $P = 0.09$ ,

Pearson correlation). **I**, Decoder accuracies from dmPFC activity (top) and multidimensional (snout, centre and tail) speed signals (bottom), when activity is aligned at CS onset (left) or AV-run (right). **m**, Comparison between dmPFC activity-based and multidimensional speed-based decoding accuracies across the different transition to movement: spontaneous freezing to nonfreezing transitions (left), spontaneous snout-movement without running (run detection based on snout speed, but no detection on centre speed) (centre) and spontaneous run (right). Decoding based on dmPFC activity was computed using the full number of units (in blue: 1,276 units (left), 1,089 units (centre) and 1,475 units (right)) or subsampling the number of units used for decoding CS<sup>+</sup> AV from CS<sup>+</sup> NAV activity (black: 634 units). Horizontal coloured lines indicate periods in which mean accuracy is significantly different from those observed from the shuffle accuracy distribution ( $P < 0.05$ , permutation test; Methods). Accuracy data are mean  $\pm$  1 s.d.

### Extended Data Fig. 10 Photoinactivation of dmPFC and BLA-to-dmPFC projections in the active avoidance task.

**a**, Optic fibre tip location for dmPFC inactivation experiments. Shaded areas are showing virus spreading. **b**, Median speed for different trial types of the ArchT groups for laser off and on conditions. **c**, Body length (measured as the distance between snout and tail) during the time course of CS<sup>-</sup> (top) and CS<sup>+</sup> NAV (bottom) trials for off and on conditions displaying overall stable values. **d**, Snout median speed during the time course of CS<sup>-</sup> (top) and CS<sup>+</sup> NAV (bottom) trials for off and on conditions. **e**, Injection of retro-ArchT and optrode implantation in the dmPFC. Light delivery induced changes in neuronal activity in the dmPFC (13 units recorded,  $n = 1$  mouse). **f**, Micrograph example of dmPFC projecting cells in the BLA labelled by the retro-ArchT strategy. Scale bars, 100  $\mu\text{m}$ . **g**, Comparison of light diffusion in 0.6% agar of standard 200  $\mu\text{m}$  core 0.39 NA with 400  $\mu\text{m}$  core 0.66 NA fibres (scale bar, 1 mm; 10 mW measured at the tip of each optic fibre). We used the 400  $\mu\text{m}$  core 0.66 NA optic fibres to maximize the illuminated area. **h**, Location of fibre tips in the amygdala (retro-ArchT  $n = 11$  mice, GFP  $n = 6$  mice). **i**, Avoidance probability at day 3 compared to day 4, in which all trials were stimulated, did not significantly differ (ns,

$P = 0.9771$  for  $\text{CS}^+$  and  $P = 0.6671$  for  $\text{CS}^-$ , RM mixed-effects model). **j**, CS decoding accuracies on day 3 and day 4, both using 85 units ( $n = 5$  mice). Thick horizontal lines indicate significant decoding accuracy values compared to the shuffle condition ( $P < 0.05$ , permutation test) (Methods). **k**, On day 5, we performed a stimulation protocol of 10 first trials with laser off followed by 10 trials of laser on. Decoding  $\text{CS}^+$  in laser on from laser off conditions (ns,  $P > 0.065$ , permutation test). **l**, Schematics of an alternative model for threat and action processing in the dmPFC. The dmPFC represents threat information inputs coming from upstream structures in a sustained manner. A downstream structure integrates that information from the dmPFC, implementing a decision mechanism such as an integrator-to-bound mechanism (top). Once the decision variable reaches a certain threshold (avoidance decision), broadcast information about the impending movement is reflected in the dmPFC (bottom). Speed error bars are 95% CI. Accuracy data are mean  $\pm 1$  s.d.

## Supplementary information

### Supplementary Discussion

This file contains Supplementary text and supplementary references.

### Reporting Summary

### Supplementary Table 1

Summary of statistical results associated with the Figures and Extended Data Figures.

### Peer Review File

### Supplementary Video 2

Supplementary Video 2 . **dmPFC population activity at AV-run.** Temporal evolution of neural trajectories assessed by GPFA showed for 30  $\text{CS}^+\text{AV}$ ,

CS<sup>+</sup>NAV and CS<sup>-</sup> trials for AV-run aligned activity. Time indicated on top (red). Related to **Figure 1i**.

## **Supplementary Video 1**

Supplementary Video 1 . ***dmPFC population activity at CS onset***. Temporal evolution of neural trajectories assessed by GPFA showed for 30 CS<sup>+</sup>AV, CS<sup>+</sup>NAV and CS<sup>-</sup> trials for CS onset aligned activity. Time indicated on top (red). Related to **Figure 1i**.

## **Rights and permissions**

### Reprints and Permissions

## **About this article**



Check for  
updates

## **Cite this article**

Jercog, D., Winke, N., Sung, K. *et al.* Dynamical prefrontal population coding during defensive behaviours. *Nature* **595**, 690–694 (2021).  
<https://doi.org/10.1038/s41586-021-03726-6>

### Download citation

- Received: 16 April 2020
- Accepted: 14 June 2021
- Published: 14 July 2021
- Issue Date: 29 July 2021

- DOI: <https://doi.org/10.1038/s41586-021-03726-6>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03726-6>

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

Food cue regulation of AGRP hunger neurons guides learning

[Download PDF](#)

- Article
- [Published: 14 July 2021](#)

# Food cue regulation of AGRP hunger neurons guides learning

- [Janet Berrios](#)<sup>1</sup>,
- [Chia Li](#)<sup>2,3</sup>,
- [Joseph C. Madara](#)<sup>1</sup>,
- [Alastair S. Garfield](#)<sup>1</sup> nAff5,
- [Jennifer S. Steger](#)<sup>1</sup> nAff6,
- [Michael J. Krashes](#) ORCID: orcid.org/0000-0003-0966-3401<sup>2,3</sup> &
- [Bradford B. Lowell](#) <sup>1,4</sup>

*Nature* volume 595, pages 695–700 (2021) [Cite this article](#)

- 6529 Accesses
- 107 Altmetric
- [Metrics details](#)

## Subjects

- [Hypothalamus](#)
- [Neural circuits](#)
- [Reward](#)

## Abstract

Agouti-related peptide (AGRP)-expressing neurons are activated by fasting—this causes hunger<sup>1,2,3,4</sup>, an aversive state that motivates the seeking and consumption of food<sup>5,6</sup>. Eating returns AGRP neuron activity towards baseline on three distinct timescales: rapidly and transiently following sensory detection of food cues<sup>6,7,8</sup>, slowly and longer-lasting in response to nutrients in the gut<sup>9,10</sup>, and even more slowly

and permanently with restoration of energy balance<sup>9,11</sup>. The rapid regulation by food cues is of particular interest as its neurobiological basis and purpose are unknown. Given that AGRP neuron activity is aversive<sup>6</sup>, the sensory cue-linked reductions in activity could function to guide behaviour. To evaluate this, we first identified the circuit mediating sensory cue inhibition and then selectively perturbed it to determine function. Here, we show that a lateral hypothalamic glutamatergic → dorsomedial hypothalamic GABAergic ( $\gamma$ -aminobutyric acid-producing)<sup>12</sup> → AGRP neuron circuit mediates this regulation. Interference with this circuit impairs food cue inhibition of AGRP neurons and, notably, greatly impairs learning of a sensory cue-initiated food-acquisition task. This is specific for food, as learning of an identical water-acquisition task is unaffected. We propose that decreases in aversive AGRP neuron activity<sup>6</sup> mediated by this food-specific circuit increases the incentive salience<sup>13</sup> of food cues, and thus facilitates the learning of food-acquisition tasks.

[Download PDF](#)

## Main

We hypothesized that the rapid regulation of AGRP neurons by food cues is likely to arise from an afferent neural pathway that is distinct from those that cause sustained inhibition. One such afferent to AGRP neurons is the GABAergic ventral dorsomedial hypothalamic (vDMH) projection that expresses leptin receptor (LEPR) and prodynorphin (PDYN)<sup>12</sup>. These neurons (hereafter referred to as DMH<sup>LEPR</sup> neurons) are activated by the sensory detection of food and their activation is proportional to caloric value and palatability<sup>12</sup>. To precisely determine the relationship between food cue detection and changes in neuronal activity, we trained mice to associate an arbitrary cue (light) with access to food. Food-restricted mice were trained on a two-alternative forced-choice task (2AFC) in which the illumination of a light source was randomly presented at one of two active ports, indicating food availability in the illuminated port<sup>14</sup>. With training, mice learned to poke the correct port associated with this initially arbitrary visual cue (light) to receive the food reward (Ensure) (Fig. 1a). Using this paradigm in combination with fibre photometry, we found that in trained mice, DMH<sup>LEPR</sup> neurons are robustly activated and their activation follows visual cue presentation (green line) as opposed to nose poking for ‘ensure’ (black line) (Fig. 1b, Extended Data Fig. 1a). As anticipated, this neural response to the otherwise arbitrary light cue was absent in naive mice—thus it developed over time as the mouse learned to associate the light cue with food availability (Fig. 1c,d, Extended Data Fig. 1a,b).

**Fig. 1: Regulation of DMH<sup>LEPR</sup> neurons by learned food cues.**

---

 **figure1**

**a**, Illustration of the 2AFC behavioural task. Randomized cue lights are associated with the appearance of food (beige). **b**, Representative heat map of a well-trained mouse using DMH<sup>LEPR</sup> neuron fibre photometry to observe cue responses during a single behavioural session. Single trials are sorted by response time. The green line indicates light onset (time = 0), the black line indicates time when the mouse poked correctly. **c**, Example traces in a mouse when completely naive to the 2AFC task (top) and after it has become proficient at the task (bottom). **d**, Mean peak amplitude of DMH<sup>LEPR</sup> neuron fibre photometry responses in naive and trained mice in the 2AFC task. Peak amplitude was taken within a 2-s window following cue-light presentation.  $n = 4$  mice; unpaired two-tailed  $t$ -test,  $*P = 0.0241$ . Data represent mean  $\pm$  s.e.m.

[Source data](#)

[Full size image](#)

DMH<sup>LEPR</sup> neurons are clearly rapidly activated by the sensory detection of food cues<sup>12</sup> (Fig. 1b, Extended Data Fig. 1a). However, it is unknown whether DMH<sup>LEPR</sup> neurons are also capable of responding to nutrients in the gastrointestinal tract, in a similar way to AGRP neurons<sup>9,10</sup>. To address this, we infused both caloric (Ensure and glucose) and non-caloric (saline and saccharin) substances directly into the stomach. DMH<sup>LEPR</sup> neurons were activated by infusions of caloric substances, but not by non-caloric substances (Extended Data Fig. 1c–g). However, compared with sensory detection of food, the DMH<sup>LEPR</sup> neuron response to gut infusions was slower and smaller in magnitude (approximately 5 min and 5%  $\Delta F_n$ ) compared to the pellet drop response (about 1 s, about 20%  $\Delta F_n$ ) (Extended Data Fig. 1d–f, h). Of note, this regulation appears to not be consequential for satiation, as inhibition of DMH<sup>LEPR</sup> neurons did not increase the amount of food eaten during a meal (Extended Data Fig. 1i, Supplementary Information).

As DMH<sup>LEPR</sup> neurons are not necessary for regulating processes that occur over longer timescales—that is, satiation (Extended Data Fig. 1i) or satiety<sup>12</sup>—we next investigated the basis for and function of the rapid regulation of DMH<sup>LEPR</sup> neurons by food cues. To find the monosynaptic afferents to DMH<sup>LEPR</sup> neurons, we used EnvA pseudotyped, G-deleted rabies tracing<sup>15,16</sup> and channelrhodopsin (ChR2)-assisted

circuit mapping. Notably, whereas many sites provide GABAergic input, only one site provides glutamatergic input—the lateral hypothalamus (LH) (Extended Data Fig. 2, Supplementary Information). As neurons in the LH expressing vesicular glutamate transporter 2 (VGLUT2) ( $LH^{VGLUT2}$  neurons) are excitatory and would activate  $DMH^{LEPR}$  neurons, and as the general LH<sup>VGLUT2</sup> population has been shown to be involved in sensory cue detection<sup>17</sup>, we characterized the  $LH^{VGLUT2} \rightarrow DMH^{LEPR}$  circuit in greater depth.

The LH is a well-known but poorly understood central regulator of motivational drive, reward learning, and feeding behaviour<sup>18,19</sup>. Previous studies have demonstrated that  $LH^{VGLUT2}$  neurons, en masse, potently suppress food intake and are aversive<sup>20,21,22,23</sup>. However, as LH<sup>VGLUT2</sup> neurons are extremely diverse, both genetically and in their projection targets<sup>24,25</sup>, we sought to test the effects of activating the DMH-projecting subset of LH<sup>VGLUT2</sup> neurons on feeding and valence (that is, positive–rewarding versus negative–aversive). As these LH<sup>VGLUT2</sup> neurons activate  $DMH^{LEPR}$  neurons, which in turn inhibit AGRP neurons, we hypothesized that their activation would produce effects consistent with AGRP neuron inhibition—namely, suppress food intake and be appetitive during times of caloric deficiency (when AGRP neurons are active). To test this, we bilaterally injected adeno-associated virus (AAV) expressing Cre-dependent-ChR2 into the LH of *Vglut2-IRES-Cre* mice and placed a single optic fibre within the midline above the DMH to selectively stimulate  $LH^{VGLUT2} \rightarrow DMH$  terminals (Fig. 2a). We then measured food intake while optically stimulating  $LH^{VGLUT2} \rightarrow DMH$  terminals, beginning at the onset of the dark cycle (when AGRP neuron activity is high and mice are inclined to eat). Consistent with our proposed circuit, we found that  $LH^{VGLUT2} \rightarrow DMH$  terminal stimulation significantly decreased food intake (Fig. 2b). To evaluate the valence of the  $LH^{VGLUT2} \rightarrow DMH$  circuit, we used a real-time place-preference assay (RTPP) in which mice were optically stimulated in one side of a behavioural arena and could roam between the stimulation and non-stimulation side. When ad libitum-fed mice were placed in the RTPP arena at the onset of the light cycle (that is, when AGRP activity is low<sup>8</sup>), the mice were agnostic to  $LH^{VGLUT2} \rightarrow DMH$  stimulation (Fig. 2c, Extended Data Fig. 3a). We then repeated the RTPP assay at the onset of the dark cycle (that is, when AGRP activity is high<sup>8</sup>). Under these conditions, ad libitum-fed mice showed a robust preference (about 80%) for  $LH^{Vglut2} \rightarrow DMH$  stimulation, which is expected to inhibit AGRP neuron activity (Fig. 2d, Extended Data Fig. 3a). Finally, to drive AGRP activity to a maximal level and to control for the time of day, we food-restricted the mice overnight and repeated the RTPP assay at the onset of the light cycle. Remarkably, the mice displayed a robust preference (approximately 80%) for  $LH^{VGLUT2} \rightarrow DMH$  stimulation (Fig. 2e, Extended Data Fig. 3a). Notably, total locomotor activity was not significantly different between X-coloured fluorescent protein (XFP)- and ChR2-expressing groups (Extended Data Fig. 3b–d). Consistent with these

$\text{LH}^{\text{VGLUT2}} \rightarrow \text{DMH}$  neurons not promoting aversion, we found that, unlike other  $\text{LH}^{\text{VGLUT2}}$  neurons, they do not send collaterals to sites known to promote aversive behaviours<sup>20,21,22,23</sup> (Extended Data Fig. 3e–g). In total, these findings suggest that the DMH-projecting  $\text{LH}^{\text{VGLUT2}}$  neurons are functionally and anatomically distinct from previously studied  $\text{LH}^{\text{VGLUT2}}$  neurons<sup>20,21,22,23,26</sup>. Furthermore, as AGRP neurons promote hunger and aversion<sup>1,2,3,4,6,27,28</sup>, the observed effects (decreased hunger and being appetitive) are consistent with these  $\text{LH}^{\text{VGLUT2}}$  neurons being upstream of the  $\text{DMH}^{\text{LEPR}}$  (GABAergic) → AGRP circuit.

**Fig. 2: Functional characterization of a  $\text{LH}^{\text{VGLUT2}} \rightarrow \text{DMH}^{\text{LEPR}}$  neuronal circuit.**



**a**, Experimental design schematic and representative image. AAV-DIO-ChR2 was bilaterally injected into  $\text{LH}^{\text{VGLUT2}}$  neurons and a single optic fibre was placed in the midline above the DMH. Scale bar, 500  $\mu\text{m}$ .  $n = 8$  mice. **b**, Optogenetic stimulation of  $\text{LH}^{\text{VGLUT2}} \rightarrow \text{DMH}$  terminals decreases nighttime food intake in ad libitum-fed mice.  $n = 8$  mice; repeated-measures two-way ANOVA; Sidak's multiple comparison test; \*\* $P = 0.0016$ , \*\*\* $P = 0.0002$ . **c–e**, Real-time place preference: optogenetic stimulation (stim.) of  $\text{LH}^{\text{VGLUT2}} \rightarrow \text{DMH}$  terminals has no effect when mice are calorically replete ( $P = 0.6531$ ) (**c**) and is appetitive when mice are calorically deficient at night ( $P = 0.008$ ) (**d**) and during fasting ( $P = 0.0129$ ) (**e**).  $n = 7$  (ChR2) and 7 (XFP). Two-tailed unpaired  $t$ -test, \* $P \leq 0.05$ , \*\* $P \leq 0.01$ ; NS, not significant. **f**, Experimental design (top) and representative image (bottom). AAV-DIO-GCaMP6s

was injected into LH<sup>VGLUT2</sup> neurons. Optic fibre was placed over the DMH to perform fibre photometry in LH<sup>VGLUT2</sup> terminals. Coloured arrows indicate light wavelengths delivered and collected. Scale bar, 500 µm.  $n = 9$  (mice). **g–j**, LH<sup>VGLUT2</sup> → vDMH axons are rapidly activated upon the sensory detection of food (**g**) and are scalable by caloric value (**h, i**) and palatability (**j**). Heat map (**g**) represents the trial-by-trial response in a representative mouse from 15 mg pellet presentation (time = 0). Line in **h–j** represents mean ± s.e.m. **k**, Mean peak response within the first 5 s following pellet presentation. One-way ANOVA; Friedman test, \*\*\* $P = 0.0007$ .  $n = 9$  mice. Data represent mean ± s.e.m.

[Source data](#)

[Full size image](#)

The LH is known to participate in many aspects of feeding behaviour<sup>18,19,24</sup> and LH<sup>VGLUT2</sup> → nucleus accumbens (NAc) neurons have been found to be necessary for signalling cue–outcome associations<sup>17</sup>. With this in mind, we tested whether the DMH-projecting LH<sup>VGLUT2</sup> neurons were responsive to food pellet drops. We used axon-fibre photometry to record LH<sup>VGLUT2</sup> → DMH axonal Ca<sup>2+</sup> activity in freely moving mice (Fig. 2f) while presenting food pellets of increasing caloric value and palatability<sup>7,12</sup> (Fig. 2g–k). LH<sup>VGLUT2</sup> → DMH axons were rapidly activated upon the presentation of a food item, were scalable to caloric content, and the responses were transient (returning to baseline within 10–20 s) (Fig. 2h–k, Extended Data Fig. 3i, j). In addition, we found that the response was mostly specific for food in that a much smaller and much more delayed response was seen when a non-edible object was presented (Extended Data Fig. 3h, i) or when water was presented to water-deprived mice (Extended Data Fig. 3k–m). In sum, our findings suggest that LH<sup>VGLUT2</sup> → DMH neurons are activated by sensory cues preceding food ingestion, and that they probably provide this information to DMH<sup>LEPR</sup> neurons.

If the LH does indeed provide food cue information to the DMH<sup>LEPR</sup> → AGRP pathway, then selectively inhibiting LH<sup>VGLUT2</sup> → DMH afferents should attenuate the food cue response in DMH<sup>LEPR</sup> neurons. We used two distinct methods to determine whether LH → DMH afferents were necessary for the cue-evoked response in DMH<sup>LEPR</sup> neurons. First, we used a retrograde AAV-FlpO virus to selectively express FlpO in LH neurons that project to the DMH. To record Ca<sup>2+</sup> activity while simultaneously inhibiting LH → DMH afferents, we injected Flp-dependent AAV-hM4Di into the LH and Cre-dependent AAV-GCaMP6s into the DMH of *Lepr-IRES-Cre* mice (Fig. 3a). Flp-dependent AAV-hM4Di efficiently hyperpolarized LH neurons (Extended Data Fig. 4a, b). Notably, inhibiting LH → DMH afferents attenuated the food cue-evoked response in DMH<sup>LEPR</sup> neurons by about 70% (Fig. 3b, c). Clozapine-

*N*-oxide (CNO) injections in GFP-expressing control mice had no effect (Extended Data Fig. 4c, d). To specifically investigate the role of LH<sup>VGLUT2</sup> neurons in mediating DMH<sup>LEPR</sup> neuron food cue responses, we used *Vglut2-IRES2-FlpO::Lepr-IRES-Cre* mice to express Flp-dependent AAV-hM4Di in LH<sup>VGLUT2</sup> neurons and Cre-dependent AAV-GCaMP6s in DMH<sup>LEPR</sup> neurons (Fig. 3d). Selectively inhibiting LH<sup>VGLUT2</sup> neurons also attenuated the rapid DMH<sup>LEPR</sup> cue-evoked response by approximately 68% (Fig. 3e, f). Collectively, these studies show that LH<sup>VGLUT2</sup> neurons drive food cue-evoked responses in DMH<sup>LEPR</sup> neurons.

**Fig. 3: LH afferents to the DMH are necessary for rapid food cue-evoked responses in DMH<sup>LEPR</sup> and AGRP neurons.**

---

 **figure3**

**a**, Top, retrograde AAV-FlpO and AAV-DIO-GCaMP6s were injected into the DMH of *Lepr-IRES-Cre* mice and AAV-fDIO-hM4Di was injected into the LH to inhibit LH → DMH neurons. Bottom, sample images of hM4Di expression (left) and GCaMP6s expression and fibre placement (right). Scale bar, 200 μm.  $n = 4$  mice. **b**, Example traces of vehicle and CNO recording session within the same animal. Vehicle and CNO recording sessions were split so that treatment comparisons were done within the same day. Vehicle 1.1 represents the first vehicle recording session and vehicle 1.2 represents the second recording session. Vehicle 2.1 represents vehicle recording session on the second day and CNO 2.2 represents the second recording

session with CNO injection. The black vertical line represents cue presentation; the purple box represents the analysed 2-s period; the blue horizontal line represents peak of vehicle control from the first recording session. **c**, LH → DMH inhibition attenuates the cue response in DMH<sup>LEPR</sup> neurons. Normalized responses are calculated as  $\Delta F/F_0$  for the second recording session (R2) divided by  $\Delta F/F_0$  for the first recording session (R1).  $n = 4$  mice. Two-tailed *t*-test, \*\**P* = 0.0035. **d**, Top, AAV-fDIO-hM4Di was injected into the LH and AAV-DIO-GCaMP6s was injected into the DMH of *Vglut2-IRES2-FlpO::Lepr-IRES-Cre* mice. Bottom, sample images of AAV-fDIO-hM4Di expression (left) and AAV-DIO-GCaMP6s expression and fibre placement (right). Scale bar, 500 μM.  $n = 4$  mice. **e**, Example traces from one mouse for experiment described in **d** (labelled as in **b**). **f**, LH<sup>VGLUT2</sup> inhibition attenuates the cue response in DMH<sup>LEPR</sup> neurons.  $n = 4$  mice; two-tailed *t*-test, \**P* = 0.0009. **g**, Top, AAV-fDIO-hM4Di was injected into the LH and AAV-DIO-GCaMP6s was injected into the ARC of *Vglut2-IRES2-FlpO::Agrp-IRES-Cre* mice. Bottom, sample images of expression and fibre placement: AAV-fDIO-hM4Di (right) and AAV-DIO-GCaMP6s (left). Scale bar, 500 μm.  $n = 4$  mice. **h**, Example traces in one mouse for vehicle and CNO (labelled as in **b**). **i**, LH<sup>VGLUT2</sup> inhibition decreases the cue response in AGRP neurons.  $n = 4$  mice; two-tailed *t*-test, \**P* = 0.027. Data represent mean ± s.e.m.

### [Source data](#)

### [Full size image](#)

To determine whether LH<sup>VGLUT2</sup> neurons do indeed cause food cue activation of AGRP neurons, we used *Vglut2-IRES2-FlpO::AGRP-IRES-Cre* mice to restrict expression of Flp-dependent AAV-hM4Di in LH<sup>VGLUT2</sup> neurons and Cre-dependent AAV-GCaMP6s in AGRP neurons (Fig. 3g). Selective inhibition of LH<sup>VGLUT2</sup> neurons robustly attenuated the rapid, cue-evoked response in AGRP neurons (approximately 76% decrease in response magnitude) but did not affect sustained AGRP inhibition during food consumption of large food pellets (Fig. 3h,i, Extended Data Fig. 4f). Thus, LH<sup>VGLUT2</sup> neurons have an essential role in food cue-evoked inhibition of AGRP neurons but appear to not be involved in gut nutrient-mediated regulation of AGRP neurons.

We next explored the function of this food cue-mediated regulation of the LH<sup>VGLUT2</sup> → DMH<sup>LEPR</sup> → AGRP circuit. In mice that learned the task, CNO–hM4Di inhibition of LH → DMH neurons had no effect on task performance (Extended Data Fig. 4e,g–j). This lack of effect may be owing to habit-like performance in such highly trained mice<sup>29</sup>. To determine whether this food cue-regulated circuit has a role in the learning of this task, we expressed Cre-dependent AAV-tetanus toxin (TeNT) in DMH<sup>LEPR</sup> neurons to eliminate evoked synaptic release within the circuit (Fig. 4a).

We then trained naive mice on the 2AFC task (Fig. 1a) until they mastered the task (or a maximum of 21 training days). We calculated the number of correct responses in addition to the errors made, including misses (failing to poke within the response window) and false alarms (poking in the incorrect port). We also assessed the response time (how long after cue presentation the poke was performed). TeNT-mediated DMH<sup>LEPR</sup> silencing caused a significant delay in task acquisition (Fig. 4b,c). In addition, the TeNT-mediated DMH<sup>LEPR</sup>-silenced mice took many more days of training to reduce their misses, false alarms and response times (Extended Data Fig. 5a–c). The higher number of misses and false alarms and, in particular, the longer response times suggest that in DMH<sup>LEPR</sup>-silenced mice, the cue light is less effective in motivating food-seeking behaviour. Consistent with the view that the DMH<sup>LEPR</sup> → AGRP neuron circuit is not involved in either satiety<sup>12</sup> or satiation (Extended Data Fig. 1i), body weight (Extended Data Fig. 5d) and post-fast refeeding (Extended Data Fig. 5e) were not affected in TeNT-mediated DMH<sup>LEPR</sup>-silenced mice.

**Fig. 4: Afferent modulation of AGRP neurons is required for learning a cue-initiated food acquisition task.**

 figure4

**a**, Top, AAV-DIO-TeNT injection into DMH<sup>LEPR</sup> neurons. Bottom, example images of TeNT expression. Scale bar, 500 µm.  $n = 5$  (mice). **b**, DMH<sup>LEPR</sup> neuron silencing attenuates correct responses.  $n = 5$  (GFP) and 5 (TeNT) mice. Red line is the line of best fit ( $***P < 0.0001$ , plateau = 97.49 (GFP), 132.4 (TeNT);  $\tau = 6.928$  (GFP),

20.80 (TeNT)). Two-way repeated measures ANOVA; main effect of days, \*\*\* $P < 0.0001$ ; main effect of group, \*\*\* $P < 0.0001$ . **c**, Silencing DMH<sup>LEPR</sup> neurons increases time to reach learning criterion (>80% correct across three consecutive days); 5 out of 5 GFP and 4 out of 5 TeNT mice reached the learning criterion.  $n = 5$  (GFP) and 5 (TeNT) mice. Two-tailed, unpaired *t*-test; \* $P = 0.0238$ . **d**, Top, AAV-DIO-TeNT injection into DMH<sup>LEPR</sup> neurons. Bottom, example images of TeNT expression. Scale bar, 500  $\mu\text{m}$ .  $n = 6$  mice. **e**, Silencing DMH<sup>LEPR</sup> neurons does not affect behavioural performance in learning to obtain water.  $n = 4$  (GFP) and 6 (TeNT) mice. Red line is the line of best fit (plateau = 129.3, tau = 11.53). Two-way repeated measures ANOVA; main effect of days, \*\*\* $P < 0.0001$ ; main effect of group,  $P = 0.7203$ . **f**, Silencing DMH<sup>LEPR</sup> neurons does not affect learning (>80% correct across three consecutive days) when water-deprived mice are trained to receive water rewards.  $n = 4$  (GFP) and 6 (TeNT) mice. Two-tailed, unpaired *t*-test;  $P = 0.4375$ . Data represent mean  $\pm$  s.e.m.

[Source data](#)

[Full size image](#)

Although AGRP neuron-regulating DMH<sup>LEPR</sup> neurons project only to the arcuate nucleus<sup>12</sup> (ARC), the DMH contains other LEPR-expressing neurons that project elsewhere<sup>12</sup>. Thus, the impairment in learning caused by TeNT silencing of DMH<sup>LEPR</sup> neurons could be owing to the silencing of DMH<sup>LEPR</sup> neurons that project to other parts of the brain (Extended Data Fig. 5f). To address this, we used *Pdyn-IRES-Cre* mice to express TeNT selectively in DMH neurons expressing PDYN (DMH<sup>PDYN</sup> neurons). It was previously established that PDYN is expressed by a major subset of DMH<sup>LEPR</sup> neurons and that DMH<sup>PDYN</sup> neurons send long-range projections exclusively to the ARC, where they inhibit AGRP neurons<sup>12</sup> (Extended Data Fig. 6). These mice were then trained on the 2AFC task. Of interest, silencing DMH<sup>PDYN</sup> neurons modestly increased body weight in TeNT-expressing mice (Extended Data Fig. 5l) but did not affect post-fast refeeding (Extended Data Fig. 5m). As body weight was not increased when DMH<sup>LEPR</sup> neurons were silenced (Extended Data Fig. 5d), this body weight effect may be owing to the silencing of LEPR-negative DMH<sup>PDYN</sup> neurons. Of note, silencing DMH<sup>PDYN</sup> neurons, similar to silencing DMH<sup>LEPR</sup> neurons, significantly altered task performance in that it delayed task acquisition, increased mistakes made and increased response time (Extended Data Fig. 5g–k). Given that DMH<sup>PDYN</sup> neurons project selectively to AGRP neurons, these findings, along with those from the DMH<sup>LEPR</sup>-silencing study (Fig. 4a–c, Extended Data Fig. 5a–c), strongly support the view that the DMH → AGRP neuronal circuit has a key role in the learning of a sensory cue–food acquisition task.

To specifically test the role of the LH → DMH segment of the circuit in learning the 2AFC task, we selectively silenced DMH-projecting LH neurons by bilaterally injecting a retrograde AAV-Cre virus into the DMH and a Cre-dependent TeNT AAV into the LH (Extended Data Fig. 5n). Indeed, silencing these neurons significantly impaired the rate of task acquisition and significantly increased misses and response times (Extended Data Fig. 5o–s). Of note, TeNT-mediated silencing of LH → DMH neurons did not affect body weight or post-fast refeeding (Extended Data Fig. 5t,u), suggesting that this LH → DMH projection is not necessary for satiety or satiation as seen when silencing DMH<sup>LEPR</sup> neurons (Extended Data Fig. 5d,e). Collectively, these data (Fig. 4a–c, Extended Data Fig. 5a–u) show that the LH<sup>VGLUT2</sup> → DMH<sup>GABAergic</sup> → AGRP neuron circuit has an important role in promoting the mastery of this caloric-deficiency-driven, sensory cue-initiated food-acquisition task.

To assess the possibility that TeNT expression in DMH<sup>LEPR</sup> neurons somehow interferes non-specifically with deprivation state-motivated learning of an operant task, we generated a cohort of TeNT-expressing water-deprived mice and determined their ability to learn the identical task with water, instead of food, as the reward (Fig. 4d). We hypothesized that the principles of learning in these two tasks would be identical, with the sole exception that water cue regulation of aversive thirst neurons<sup>6,30,31,32</sup>, instead of food cue regulation of hunger neurons, would promote learning. As the DMH → AGRP circuit is involved in food cue regulation of hunger neurons, and not in water cue regulation of thirst neurons, this experiment controls for nonspecific effects. As shown in Fig. 4e,f and Extended Data Fig. 5v–x, TeNT expression had no effect on learning in the water-oriented version of the task. Thus, the DMH<sup>LEPR</sup> → AGRP neuronal circuit is specific for learning food cue-initiated food-acquisition tasks, which is congruous with the very specific role of AGRP neurons in regulating food intake but not water intake.

## Discussion

These findings lead us to propose the following model. Caloric deficiency activates AGRP neurons (reviewed in ref. 11) and this causes the aversive feeling of hunger<sup>6</sup>. Environmental cues instructive for food acquisition engage the LH<sup>VGLUT2</sup> → DMH<sup>LEPR</sup> → AGRP neuron circuit and this transiently reduces AGRP neuron activity. As AGRP neurons have been proposed to transmit a negative-valence teaching signal<sup>6</sup>, these ‘appetitive’ falls in aversive AGRP neuron activity, over time, increase the incentive salience<sup>13</sup> of food cues, thereby facilitating the learning of food acquisition tasks. If it holds, this model implies the existence of a very important general neurobiological mechanism for how homeostatic deficiency states promote the learning of tasks directed at acquiring the cognate goals (such as caloric deficiency

promoting food acquisition tasks or dehydration promoting water acquisition tasks)—by providing the ‘substrate’—the deficiency-activated aversive drive neurons<sup>6,30,31,32</sup>—on which the inhibitory sensory cue-regulated afferent neurons can operate.

## Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

### Experimental subjects

*Vglut2-IRES-Cre* (JAX 016963)<sup>33</sup>, *Vglut2-IRES2-FlpO* (JAX 030212) (unpublished, donating investigator H. Zeng (Allen Institute for Brain Science)), *AGRP-IRES-Cre* (JAX 012899)<sup>34</sup>, *Pdyn-GFP*<sup>35</sup> and wild-type mice (JAX101045) were obtained through Jackson Laboratories or in house. *Lepr-IRES-Cre*<sup>36</sup> mice were maintained as previously described<sup>12</sup>. All mice were maintained on a mixed genetic background unless otherwise noted. *Pdyn-IRES-Cre*<sup>27</sup> mice were maintained as previously described and kept on a congenic C57 BL/6J background. The National Institute of Health and Beth Israel Deaconess Medical Center Institutional Animal Care and Use Committee approved all animal care and experimental procedures. Mice were housed at 22–24 °C, 20–30% humidity with a 12:12 light:dark cycle with standard mouse chow (Teklad F6 Rodent Diet 8664) and water was provided ad libitum, unless otherwise stated. All diets were provided as pellets. For all behavioural studies, we used male mice between 8 and 20 weeks of age. For electrophysiological recordings, we used male mice between 8 and 12 weeks of age.

### Brain tissue preparation

Mice were terminally anaesthetized with chloral hydrate (Sigma-Aldrich C8383) and transcardially perfused with phosphate-buffered saline (PBS) followed by 10% neutral buffered formalin (Fisher Scientific SF100). Brains were extracted then cryoprotected in 20% sucrose. All brains were sectioned coronally on a freezing sliding microtome (Leica Biosystems) at 40 µm and collected in four equal series.

### Immunohistochemistry

Tissue sections were washed with 0.1 M phosphate-buffered saline (pH 7.4) then blocked in 5% normal donkey serum/0.2% Triton X-100 in PBS for 1 h at room temperature. Sections were then incubated overnight at room temperature in blocking solution containing: rat anti-mCherry (1:3,000, Invitrogen M11217) and chicken anti-

GFP (1:1,000, Invitrogen A10262). Secondary detection was performed with Alexa Fluor 488 or 594 conjugated donkey anti-chicken or donkey anti-rat (1:1,000, Invitrogen) for 1 h at room temperature. After secondary incubation, sections were washed and mounted onto gelatin-coated slides and fluorescent images were obtained with an Olympus VS120 slide scanner microscope. Our inferred Bregma coordinates on all histological images were adopted from a stereotaxic atlas<sup>37</sup>.

## Stereotaxic surgeries and viral injections

For viral injections, six-to-eight week old male mice were anaesthetized with a ketamine (100 mg kg<sup>-1</sup>) and xylazine (10 mg kg<sup>-1</sup>) cocktail diluted in 0.9% saline and placed into a stereotaxic apparatus (Kopf model 940). Subcutaneous injection of sustained-release meloxicam (4 mg kg<sup>-1</sup>) was provided as postoperative care. A pulled glass micropipette (20–40 µm diameter tip) was used for stereotaxic injections of AAV. For electrophysiological experiments, bilateral injections (25 nl) of purified AAV ( $6.24 \times 10^{12}$  viral genomes ml<sup>-1</sup>) were injected into the NAc (from bregma: +1.3 AP, ± 0.5 ML, -4.25 DV), BNST (from bregma: +0.14 AP, ± 0.75ML, -4.9 DV), MPO (from bregma: +0.4 AP, ± 0.25 ML, -4.75 DV), LH (from bregma: -1.3 AP, ± 0.9 ML, -5.3 DV), VTA (from bregma: -3.15 AP, ± 0.6 ML, -4.74 DV), IPN (from bregma: -3.4 AP, ± 0.0 ML, -4.8 DV), and PAG (from bregma: -4.16 AP, ± 0.3 ML, -2.5 DV). For optogenetic experiments, bilateral injections (15 nl) of AAV9-EF1α-DIO-ChR2(H134R)-eYFP purchased from the University of Pennsylvania School of Medicine Vector Core (donating investigator, K. Deisseroth; AV-9-20298P; Addgene: 20298;  $2.7 \times 10^{13}$  viral genomes ml<sup>-1</sup>) were injected into the LH (coordinates as above). For in vivo fibre photometry experiments, AAV1-hSyn-DIO-GCaMP6s (University of Pennsylvania Vector core; Addgene 100845-AAV1;  $1.6 \times 10^{13}$  viral genomes ml<sup>-1</sup>) was injected unilaterally into either the LH (25 nl, coordinates as above), DMH (25 nl, from bregma: -1.80 AP, ± 0.3 ML, -5.2 DV), or ARC (150 nl, from bregma: -1.45 AP, ± 0.25 ML, -5.85 DV). For retrograde chemogenetic silencing studies, AAV6-CAG-FlpO (Boston Children's Hospital Viral Vector Core, modified from Addgene 67829;  $2.88 \times 10^{14}$  viral genomes ml<sup>-1</sup>) was bilaterally injected into the DMH (15 nl, coordinates as above) and AAV8-nEF-fDIO-hM4Di-mCherry (Boston Children's Hospital Viral Vector Core, modified from Addgene 44362 and 55644;  $9.98 \times 10^{13}$  viral genomes ml<sup>-1</sup>) was bilaterally injected into the LH (25 nl, coordinates as above). For Flp-dependent chemogenetic studies in combination with fibre photometry, AAV8-nEF-fDIO-hM4Di-mCherry was bilaterally injected into the LH (25 nl, coordinates as above) or, the DMH (25 nl, coordinates as above). For tetanus toxin mediated silencing studies, AAVDJ-CMV-DIO-eGFP-2A-TeNT (Stanford; GVVC-AAV-71;  $3.6 \times 10^{12}$  viral genomes ml<sup>-1</sup>) was injected bilaterally into the DMH (15 nl, coordinates as above). Finally, for projection specific tetanus toxin mediated silencing studies, AAVDJ-CMV-DIO-eGFP-2A-TeNT was

injected bilaterally into the LH (20 nl, coordinates as above) and rAAV2-hSyn-Cre (Addgene 105553;  $1.2 \times 10^{13}$  viral genomes ml $^{-1}$ ) was injected bilaterally into the DMH (15 nl, coordinates as above). Mice were allowed to recover for a minimum of three weeks before the initiation of any experiments. Following each experimental procedure, accuracy of AAV injections were confirmed via post hoc histological analysis of mCherry, YFP, or GFP fluorescent protein reporters, viral expression of each individual surgery is catalogued in detail within Extended Data Figs. 7–11. All subjects determined to be surgical ‘misses’ were those with absent or low reporter expression and were removed from the experimental dataset. In addition, mice were excluded from the data set if reporter expression was primarily outside of the area of interest. Anatomical boundaries were drawn by using the DAPI signal so that we could clearly discern landmarks within a histological section. On the basis of the landmarks present, we then inferred the A/P coordinate in the atlas<sup>37</sup> and traced the outline of different nuclei and superimposed the outline on our histological images Extended Data Fig. 12.

## Optic fibre implantation

Optic fibre implantations were performed during the same surgery as viral injections (above). For optogenetic photostimulation of LH → DMH terminals, ceramic ferrule (Precision Fibre Products) optical fibres (200 μm diameter core, 0.39 NA, multimode; Thorlabs) were implanted within the midline over the DMH (from bregma: −1.8 AP, ± 0 ML, −4.7 DV). For LH<sup>VGLUT2</sup> → DMH axon fibre photometry, a metal ferrule (Precision Fibre Products) optic fibre (400 μm diameter core, 0.5 NA, multimode; Thorlabs) was implanted unilaterally over the DMH (from bregma: −1.8 AP, ± 0.3 ML, −5.1 DV). For DMH<sup>LEPR</sup> and ARC<sup>AGRP</sup> cell body fibre photometry, a stainless steel ferrule optic fibre (same as above) was implanted unilaterally over the DMH (coordinates as above) or the ARC (from bregma: −1.45 AP, ± 0.25 ML, −5.8 DV). Fibres were fixed to the skull using dental acrylic and mice were allowed to recover for three weeks before the start of acclimation to behavioural testing.

## Monosynaptic rabies mapping

*Lepr-IRES-Cre* mice were unilaterally injected with a 1:1 mixture of AAV8-EF1α-DIO-TVA-mCherry (University of North Carolina Vector Core, donating investigator N. Uchida) and AAV8-CAG-FLEX-Rabies G (Stanford; GVVC-AAV-59) into the DMH (15 nl, coordinates as above) (Extended Data Fig. 2a). Mice recovered for three weeks after TVA and RG transduction to ensure adequate levels of TVA and RG viral expression. Following recovery, mice underwent a second surgery in which mice were injected with SADΔG-EGFP (EnvA) rabies (Salk Gene Transfer Targeting and Therapeutics Core) into the DMH (15 nl, coordinates as above). Mice recovered for 7 days to allow for the retrograde transport of rabies virus and EGFP expression before

perfusion and histological processing. Sites of afferent input to DMH<sup>LEPR</sup> neurons were assessed by the presence of EnvA-eGFP positive neurons.

## Rabies collateral mapping

*Vglut2-IRES-Cre* mice were unilaterally injected with AAV8-EF1 $\alpha$ -DIO-TVA-mCherry into the LH (15 nl) and allowed to recover for three weeks (Extended Data Fig. 3e–g). Then, SAD $\Delta$ G-EGFP (EnvA) rabies was unilaterally injected into the DMH (15 nl). Mice were allowed to recover for seven days to allow for the retrograde transport of rabies virus and EGFP transgene expression before perfusion and histological processing. Comprehensive examination of SAD $\Delta$ G-EGFP (EnvA) axonal and retrograde transduction was assayed using immunohistochemistry followed by imaging the entire brain for the presence of EGFP expression.

## Electrophysiology

Mice were deeply anaesthetized, and intracardially perfused with ice-cold dissection buffer (in mM: 2.5 KCl, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 20 HEPES, 10 MgSO<sub>4</sub> $\cdot$ 7H<sub>2</sub>O, 0.5 CaCl<sub>2</sub> $\cdot$ 2H<sub>2</sub>O, 92 choline chloride, 25 glucose, 2 thiourea, 5 sodium ascorbate, 3 sodium pyruvate, and 20 NaHCO<sub>3</sub>) bubbled with 95% O<sub>2</sub>, 5% CO<sub>2</sub>. Brains were then rapidly removed and immersed in ice-cold dissection buffer. DMH sections were dissected and 300  $\mu$ m thick coronal slices were prepared using a vibrating microtome (Campden 7000smz 2). Slices recovered for 10 min in a 35 °C submersion chamber filled with oxygenated dissection buffer. Slices were then transferred to a secondary 35 °C submersion chamber filled with oxygenated artificial cerebrospinal fluid (ACSF; in mM: 125 NaCl, 2.5 KCl, 25 NaHCO<sub>3</sub>, 2 CaCl<sub>2</sub>, 1 MgCl<sub>2</sub>, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 25 glucose) and allowed to recover for an additional 15 min. Slices were then kept at room temperature in oxygenated ACSF for  $\geq$  30 min until use (Extended Data Fig. 2b–m).

## Channelrhodopsin-2 assisted circuit mapping

To isolate optically evoked inhibitory post-synaptic currents (oIPSCs) and excitatory post-synaptic currents (oEPSCs), slices were placed in a submersion chamber and perfused at 4 ml min<sup>-1</sup> with oxygenated ACSF. Cells were visualized with a Scientifica SliceScope Pro 1000 microscope equipped with infrared differential interference contrast optics. DMH<sup>PDYN</sup> GABAergic neurons were identified by GFP fluorescence ventral to the DMC (dorsomedial hypothalamic nucleus compacta). Open-tip resistances for patch pipettes were between 2–4 M $\Omega$  and were backfilled with a Cs-based internal solution containing (in mM): 135 CsMeSO<sub>3</sub>, 10 HEPES, 1

EGTA, 3.3 QX-314 ( $\text{Cl}^-$  salt), 4  $\text{Mg}_2$ -ATP, 0.3  $\text{Na}_2$ -GTP, and 8  $\text{Na}_2$ -phosphocreatine with pH adjusted to 7.3 with CsOH and osmolarity adjusted to about 295 mOsM by the addition of sucrose. oEPSCs were isolated with membrane potential clamped at  $V_h = -70$  mV and oIPSCs were isolated with membrane potential clamped at  $V_h = 0$  mV. Bath solutions for pharmacological isolation of excitatory or inhibitory currents in whole-cell voltage clamp recordings contained SR95531 (10  $\mu\text{M}$ , gabazine), kynurenic acid (1 mM), tetrodotoxin (TTX, 1  $\mu\text{M}$ ), and 4-Aminopyridine (4-AP, 500  $\mu\text{M}$ ). To photostimulate ChR2-positive fibres, an LED light source was used (470 nM, Cool LED pE-100). The blue light was focused onto the back aperture of the microscope objective (40 $\times$ ) producing wide-field exposure around the recorded cell of 10–15 mW per  $\text{mm}^2$  as measured using an optical power meter (PM100D, Thorlabs). A programmable pulse stimulator, Master-8 (A.M.P.I.) and pClamp 10.2 and 10.6 software (Molecular Devices, Axon Instruments) controlled the photostimulation output. The oIPSC/oEPSC detection protocol consisted of one blue-light pulse (5 ms pulse length) at 30 s intervals for at least 6 consecutive sweeps. Changes in series and input resistance were monitored throughout the experiment by giving a test pulse every 30 s and measuring the amplitude of the capacitive current. Cells were discarded if series resistance rose above 25  $\text{M}\Omega$  (Extended Data Fig. [2b–m](#)). All electrophysiology data was analysed using Clampfit 10.2 and 10.6.

## Optogenetic behavioural experiments

In vivo photostimulation of  $\text{LH}^{\text{VGLUT2}} \rightarrow \text{DMH}$  terminals was conducted by firmly attaching a fibre optic cable (1.25 m long, 200  $\mu\text{m}$  core diameter, 0.63 NA; Doric Lenses) with ceramic split sleeves (Precision Fibre Products) (Fig. [2](#), Extended Data Fig. [3a–d](#)). Mice were acclimated by connecting them to a ‘dummy’ fibre optic cable three days before the initiation of the experiment. Mice were stimulated with blue light (465 nM LED; Plexon) at 20 Hz, 5 ms pulses for 1 s with a 3 s recovery period (LED off) during stimulation trains to avoid ChR2 desensitization, neuronal transmitter depletion, and tissue heating. Light pulse trains were programmed using a waveform generator (National Instruments) that provided TTL input to the blue light LED. The light power exiting the fibre optic cable measured by an optical power meter (Thorlabs) was 7–8 mW in all experiments. After completion of photostimulation experiments, mice were perfused for assessment of surgical accuracy of both ChR2-expression and optic fibre tip location via histological analysis as described in ‘Stereotaxic surgeries and viral injections’.

## Food intake studies

To test the sufficiency of  $\text{LH}^{\text{VGLUT2}} \rightarrow \text{DMH}$  neurons for satiety, mice were tested under conditions of physiological hunger at the onset of the dark cycle (Fig. [2b](#)). For

dark cycle feeding, mice with ad libitum access to food were photostimulated for 5 min before the onset of the dark cycle (a time when mice often eat) and photostimulation continued throughout the duration of the study (three hours). For post-fast refeeding assays (Extended Data Figs. 1, 5), mice were food restricted for 24 h then, were given ad libitum access to food. Food was then weighed each hour to determine the amount consumed during the experimental manipulation.

## Real-time place preference assays

Mice were placed in a custom-made behavioural arena (transparent acrylic, 50 × 50 × 25 cm) for 20 min (Fig. 2, Extended Data Fig. 3a–d). One counterbalanced side of the arena was designated as the photostimulation side. The mouse was placed in the stimulation side at the onset of the experiment and each time the mouse crossed to the non-stimulation side of the arena, the photostimulation immediately stopped until the animal crossed back into the stimulation side. Behavioural data was recorded with Ethovision software (Noldus Information Technologies). To test photostimulation preference during different hunger states, ad libitum-fed mice were placed within the arena immediately before the onset of the dark cycle. Following a one-week rest period, ad libitum fed mice were then placed in the same arena and tested for photostimulation preference at the onset of the light cycle. Following a one-week rest period, mice were then fasted overnight then, placed in the arena and tested for photostimulation preference at the onset of the light cycle. Each RTPP assay was counterbalanced within animal and within day.

## In vivo fibre photometry

Fibre photometry was performed on a rig constructed as follows: A 465-nm LED (PlexBright LED, Plexon) was used as the excitation source which was passed through a fluorescence mini cube (excitation: 460–490 nm, detection: 500–550 nm; Doric Lenses), and transmitted onto the sample via a fibre optic cable (1 m long, 400 µm diameter, 0.48 NA; Doric Lenses). The optic fibre was coupled to the implanted optic fibre with a ceramic mating sleeve (Precision Fibre Products). Light intensity was measured as 100–200 µW at the end of the patch cord and was kept constant across sessions for each mouse. Emitted light was collected by a photodetector (2151; Newport). The signal was digitized at 1 kHz with a data acquisition card (National Instruments) and collected with a custom MATLAB (MATLAB2016a; MathWorks) script (Figs. 1–3, Extended Data Figs. 1, 3, 4).

For LH<sup>VGLUT2</sup>→DMH axon fibre photometry recordings (Fig. 2, Extended Data Fig. 3h–j), mice underwent a 25-min recording session within their home cage that consisted of: eight ‘small’ chow (15 mg; Bio-Serv) trials, two ‘large’ chow (500 mg; Bio-Serv) trials, and one peanut butter (Reese’s peanut butter chips; Hershey) trial

which were dropped into a Pyrex Petri dish. Each mouse underwent only one session per day and was food restricted (85% of ad libitum body weight) for at least one week before beginning the experiment in which it was habituated to the pellet drops within the Petri dish. All trials were pooled to calculate mean peak response (0–5 s following food presentation) to each food presentation. For object drop experiments, mice underwent a 12-min recording session within their home cage that consisted of 10 total trials of non-edible object drops. Objects consisted of uniform, white plastic marbles (BC Precision, Hungry Hungry Hippos Marbles). After an object drop session, a 500 mg chow pellet was dropped as a positive control. Each mouse underwent a single session per day. All trials were pooled to calculate mean z-score response to each object drop.

For water presentation experiments (Extended Data Fig. 3k–m), mice were water restricted (85% of original body weight) for at least one week before beginning the experiment. Mice were then habituated to receiving water in a ceramic dish within their home cage and were given free access for 5 min. All sessions were pooled to calculate the mean peak response (0–30 s) and the time to peak response.

Data were analysed using a custom MATLAB (MATLAB2016b; MathWorks) script (Figs. 1–3, Extended Data Figs. 1, 3, 4). Fluorescence traces were down-sampled from 1 kHz to 100 Hz and smoothed using a 1-s running average. The fractional change in fluorescence was calculated as  $\Delta F/F = (F - F_0)/F_0$ , in which  $F_0$  was the mean of all data points from the baseline before each trial. In home cage pellet drop or water presentation experiments,  $F_0$  was the average for 5 s before the food drop/water presentation. In operant chamber silencing experiments,  $F_0$  was the average for 1 s before cue presentation. All trials in a single session were averaged then the mean peak amplitude was taken for quantification.

## Two-alternative forced-choice task

To determine if LH<sup>VGLUT2</sup>→DMH neurons and DMH<sup>LEPR</sup> neurons are necessary for food cue responses (Figs. 1–3, Extended Data Figs. 1, 4), mice underwent training in a three-nosepoke operant chamber (Bpod r2; Sanworks) controlled by a custom MATLAB (MATLAB2016a; Mathworks) script. In brief, food restricted mice (85% of ad libitum body weight) were trained to associate a light presentation with Ensure (10 µl; Ensure PLUS, vanilla) and were required to nose poke and hold their snouts within the port for 200 ms before Ensure was delivered. Light delivery was randomized between the left and right nose poke and, mice had a 10 s response window with a 10 s inter-trial interval. The required learning criterion was a success rate of  $\geq 80\%$  across three consecutive days. After mice learned the task, mice were then attached to a fibre optic cable as described above. To determine the necessity of LH afferents on food-cue responses (Fig. 3, Extended Data Fig. 4), mice underwent two, 11-min recording

sessions on the same day separated by a 10-min ‘break’ period. The first session was always a saline run and mice were injected with saline 10 min before the onset of the recording session. The second session was a clozapine-*n*-oxide (CNO; 1 mg kg<sup>-1</sup>; 0.5% body weight volume) or saline injection 10 min before the onset of each recording session. Comparisons between vehicle and CNO recordings were made within day; therefore, a mouse received two saline injections or, one saline and one CNO injection in a single recording session. A TTL pulse triggered at the onset of each trial determined cue onset. All trials were pooled to calculate the mean peak response (0–2 s following cue presentation) and were normalized to the first recording session within the same day.

## Intragastric catheter surgery

Mice with DMH<sup>LEPR</sup> photometry neural activity signal larger than 10%  $\Delta F/F$  to chow during fast re-feed were implanted with intragastric catheters<sup>38</sup> (Extended Data Fig. 1). During surgery, mice were anaesthetized with isoflurane (1.5–3%) and treated post-operation with buprenorphine (1 mg kg<sup>-1</sup> subcutaneously) analgesia. A midline incision on the abdomen was made through skin and muscle layers. Micro-Renathane catheter tubing 6–7 cm in length (Braintree Scientific, MRE-033, 0.033 × 0.014 in) was anchored with epoxy spheres on each end (Devcon Clear Epoxy Adhesive, 92926, Lowes). The catheter was inserted into the fundus of the stomach through a puncture hole and secured with surgical mesh (5-mm diameter piece, Bard, 0112660). The other end of the catheter was directed out of an intrascapular incision. A metal cap made out of 27G blunt needle was placed in the exposed end for seal. The gastric catheter was flushed with sterile water immediately, and daily, after surgery to prevent blockage. Mice were fed with gel chow diet and given at least 1–1.5 week for recovery before experimentation. Daily body weight was monitored until stable pre-surgical weight was regained.

## Gastric infusion

Upon recovery, gastric infusions of liquid substances listed below were performed in a counterbalanced experimental design, under both overnight fasted and sated conditions (Extended Data Fig. 1). The intrascapularly exposed end of the gastric catheters were connected to tubing and a syringe driven by an infusion pump (Harvard Apparatus, 70-3007). At a rate of 0.1 ml min<sup>-1</sup>, 1-ml infusions were performed<sup>38</sup> over the course of 10 min. Fibre photometry recordings were collected via a lock-in amplifier (TDT) and the software Synapse (TDT). Each trial was approximately 27 min (>7 min baseline recording, followed by a 10 min gastric infusion and a 10 min chow refeed). Each mouse underwent infusions of the following infusates: 0.9% isotonic saline, Ensure Original Nutrition Shake (vanilla), 1% saccharine, 25% d-glucose (equal caloric content as Ensure).

Each photometry recording data point was normalized against the average of the last 5 min of baseline period to produce the normalized traces. For normalized and delta ( $\Delta F_n/F_n$ , %) quantified comparisons, we use 1-min averages at the end of baseline period ( $t = -1$  to 0 min), 5 min into infusion ( $t = 4$  to 5 min), and 10 min into infusion ( $t = 9$  to 10 min). Owing to the transient nature of neural response to food, comparison of delta ( $\Delta F_n/F_n$ , %) maximum magnitude between ensure infusion and food presentation were made using 30 s averages at the end of ensure infusion ( $t = 9.5$ –10 min) and at the beginning of food presentation after saline infusion ( $t = 10$ –10.5 min).

## TeNT-mediated silencing

AAV-DIO-TeNT was injected into the DMH of *Lepr-IRES-Cre* mice, *Pdyn-IRES-Cre* mice, or wild-type mice (Fig. 4, Extended Data Fig. 5). Littermate controls were used for AAV-DIO-TeNT and AAV-DIO-GFP behavioural groups. Following three weeks, mice were then placed within the Bpod arena and trained everyday as described above. Mice were either food or water restricted and maintained at  $\geq 85\%$  ad libitum body weight. Food and water restriction was performed in separate cohorts of mice. For food-learning assays, mice were trained for a total of 21 days. For water-learning assays, mice were trained for a total of 14 days. Mice were excluded if they were non-learners meaning that they did not increase their performance rate for five consecutive days (one GFP-expressing mouse (in water-deprived group) and one TeNT-expressing mouse (in food-deprived group) were removed based on this criterion). Mice were placed in the Bpod for a total of 20 min and allowed to perform as many trials as possible with an intertrial interval of 5 s and a response window of 10 s. Mice were given either 10  $\mu$ l of Ensure or 5  $\mu$ l of water for food and water-learning assays, respectively. Each day, performance in the task was quantified with a custom MATLAB script (Fig. 4, Extended Data Fig. 5). It should be noted that *Pdyn-IRES-Cre* and wild-type behavioural cohorts (Extended Data Fig. 5–u) were different strains than our *Lepr-IRES-Cre* cohort; this was owing to the limited availability of mice during the COVID-19 pandemic. As such, they performed slightly differently on the 2AFC task. Therefore, their learning criterion was lowered to  $>70\%$  correct responses across three consecutive days.

## Quantification and statistical analysis

Statistical analyses were performed using Prism 5 and Prism 8 (GraphPad) software and are described in the figure legends in all cases. No statistical method was used to predetermine sample size, nor were randomization and blinding methods used.

Statistical significance was defined as  $P < 0.05$ . All data presented met the assumptions of the statistical test employed. As mentioned in sections above, experimental mice were excluded if histological validation revealed poor or absent reporter expression or poor fibre optic placement in the region of interest. These

criteria were established before data collection.  $n$  values reflect the final number of validated mice per group.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

The datasets generated and/or analysed during the current study are available from the corresponding author upon request. [Source data](#) are provided with this paper.

## Code availability

MATLAB scripts to run and analyse 2AFC behaviour are freely available on Github ([https://github.com/jbwallace123/2AFC\\_BPOD\\_ANALYSIS](https://github.com/jbwallace123/2AFC_BPOD_ANALYSIS)). For more details on installation please visit: <https://sites.google.com/site/bpodddocumentation/installing-bpod>.

## References

1. 1.

Aponte, Y., Atasoy, D. & Sternson, S. M. AgRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nat. Neurosci.* **14**, 351–355 (2011).

[CAS Article](#) [Google Scholar](#)

2. 2.

Gropp, E. et al. Agouti-related peptide-expressing neurons are mandatory for feeding. *Nat. Neurosci.* **8**, 1289–1291 (2005).

[CAS Article](#) [Google Scholar](#)

3. 3.

Krashes, M. J. et al. Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J. Clin. Invest.* **121**, 1424–1428 (2011).

[CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Luquet, S., Perez, F. A., Hnasko, T. S. & Palmiter, R. D. NPY/AgRP neurons are essential for feeding in adult mice but can be ablated in neonates. *Science* **310**, 683–685 (2005).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Hull, C. L. *Principles of Behavior* (D. Appleton-Century, 1943).

6. 6.

Betley, J. N. et al. Neurons for hunger and thirst transmit a negative-valence teaching signal. *Nature* **521**, 180–185 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Chen, Y., Lin, Y. C., Kuo, T. W. & Knight, Z. A. Sensory detection of food rapidly modulates arcuate feeding circuits. *Cell* **160**, 829–841 (2015).

[Article](#) [Google Scholar](#)

8. 8.

Mandelblat-Cerf, Y. et al. Arcuate hypothalamic AgRP and putative POMC neurons show opposite changes in spiking across multiple timescales. *eLife* **4**, e07122 (2015).

[Article](#) [Google Scholar](#)

9. 9.

Beutler, L. R. et al. Dynamics of gut–brain communication underlying hunger. *Neuron* **96**, 461–475 (2017).

[MathSciNet](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Su, Z., Alhadeff, A. L. & Betley, J. N. Nutritive, post-ingestive signals are the primary regulators of AgRP neuron activity. *Cell Rep.* **21**, 2724–2736 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Andermann, M. L. & Lowell, B. B. Toward a wiring diagram understanding of appetite control. *Neuron* **95**, 757–778 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Garfield, A. S. et al. Dynamic GABAergic afferent modulation of AgRP neurons. *Nat. Neurosci.* **19**, 1628–1635 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Berridge, K. C. From prediction error to incentive salience: mesolimbic computation of reward motivation. *Eur. J. Neurosci.* **35**, 1124–1143 (2012).

[Article](#) [Google Scholar](#)

14. 14.

Shimp, C. P. Probabilistically reinforced choice behavior in pigeons. *J. Exp. Anal. Behav.* **9**, 443–455 (1966).

[CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Wall, N. R., Wickersham, I. R., Cetin, A., De La Parra, M. & Callaway, E. M. Monosynaptic circuit tracing in vivo through Cre-dependent targeting and complementation of modified rabies virus. *Proc. Natl Acad. Sci. USA* **107**, 21848–21853 (2010).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Wickersham, I. R., Finke, S., Conzelmann, K.-K. & Callaway, E. M. Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nat. Methods* **4**, 47–49 (2007).

[CAS](#) [Article](#) [Google Scholar](#)

17. 17.

de Jong, J. W. et al. A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron* **101**, 133–151 (2019).

[Article](#) [Google Scholar](#)

18. 18.

Petrovich, G. D. Lateral hypothalamus as a motivation–cognition interface in the control of feeding behavior. *Front. Syst. Neurosci.* **12**, 14 (2018).

[Article](#) [Google Scholar](#)

19. 19.

Sternson, S. M. & Eiselt, A.-K. Three pillars for the neural control of appetite. *Annu. Rev. Physiol.* **79**, 401–423 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Stamatakis, A. M. et al. Lateral hypothalamic area glutamatergic neurons and their projections to the lateral habenula regulate feeding and reward. *J. Neurosci.* **36**, 302–311 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Jennings, J. H., Rizzi, G., Stamatakis, A. M., Ung, R. L. & Stuber, G. D. The inhibitory circuit architecture of the lateral hypothalamus orchestrates feeding. *Science* **341**, 1517–1521 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Nieh, E. H. et al. Inhibitory input from the lateral hypothalamus to the ventral tegmental area disinhibits dopamine neurons and promotes behavioral activation. *Neuron* **90**, 1286–1298 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Li, Y. et al. Hypothalamic circuits for predation and evasion. *Neuron* **97**, 911–924.e5 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Bonnavion, P., Mickelsen, L. E., Fujita, A., de Lecea, L. & Jackson, A. C. Hubs and spokes of the lateral hypothalamus: cell types, circuits and behaviour. *J. Physiol.* **594**, 6443–6462 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Mickelsen, L. E. et al. Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. *Nat. Neurosci.* **22**, 642–656 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Jennings, J. H. et al. Visualizing hypothalamic network dynamics for appetitive and consummatory behaviors. *Cell* **160**, 516–527 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Krashes, M. J. et al. An excitatory paraventricular nucleus to AgRP neuron circuit that drives hunger. *Nature* **507**, 238–242 (2014).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Atasoy, D., Betley, J. N., Su, H. H. & Sternson, S. M. Deconstruction of a neural circuit for hunger. *Nature* **488**, 172–177 (2012).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Balleine, B. W. The meaning of behavior: discriminating reflex and volition in the brain. *Neuron* **104**, 47–62 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Allen, W. E. et al. Thirst-associated preoptic neurons encode an aversive motivational drive *Science* **357**, 1149–1155 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Leib, D. E. et al. The forebrain thirst circuit drives drinking through negative reinforcement. *Neuron* **96**, 1272–1281.e4 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Zimmerman, C. A., Leib, D. E. & Knight, Z. A. Neural circuits underlying thirst and fluid homeostasis. *Nat. Rev. Neurosci.* **18**, 459–469 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

33. 33.

Vong, L. et al. Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71**, 142–154 (2011).

[CAS](#) [Article](#) [Google Scholar](#)

34. 34.

Tong, Q., Ye, C.-P., Jones, J. E., Elmquist, J. K. & Lowell, B. B. Synaptic release of GABA by AgRP neurons is required for normal regulation of energy balance. *Nat. Neurosci.* **11**, 998–1000 (2008).

[CAS](#) [Article](#) [Google Scholar](#)

35. 35.

Gong, S. et al. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**, 917–925 (2003).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

36. 36.

Leshan, R. L., Björnholm, M., Münzberg, H. & Myers, M. G. Jr. Leptin receptor signaling and action in the central nervous system. *Obesity* **14** (Suppl 5), 208S–212S (2006).

[CAS](#) [Article](#) [Google Scholar](#)

37. 37.

Paxinos, G. & Franklin, K. *The Mouse Brain In Stereotaxic Coordinates* (Academic, 1997).

38. 38.

Han, W. et al. Striatal dopamine links gastrointestinal rerouting to altered sweet appetite. *Cell Metab.* **23**, 103–112 (2016).

[Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank M. Andermann, Y. Livneh and J. Resch for advice regarding photometry, behavioural experiments and analyses; the Lowell laboratory for helpful discussions; H. Kucukdereli and Y. Livneh for coding assistance; Z. Yang for technical assistance; C. Wu for assistance with creating viral plasmids; and the Boston Children's Viral Core for packaging/preparation of viral plasmids. This research was funded by the following NIH grants to B.B.L.: NIH P30 DK046200, NIH P30 DK057521, R01 DK075632, R01 DK089044, R01 DK096010; to J.B.: F32DK118807; and to M.J.K.: 1ZIADK075087-07.

## Author information

## Author notes

### 1. Alastair S. Garfield

Present address: Rhythm Pharmaceuticals, Boston, MA, USA

### 2. Jennifer S. Steger

Present address: Nanodropper, Rochester, MN, USA

## Affiliations

### 1. Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

Janet Berrios, Joseph C. Madara, Alastair S. Garfield, Jennifer S.  
Steger & Bradford B. Lowell

### 2. Diabetes, Endocrinology and Obesity Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA

Chia Li & Michael J. Krashes

### 3. National Institute on Drug Abuse, National Institutes of Health, Baltimore, MD, USA

Chia Li & Michael J. Krashes

### 4. Program in Neuroscience, Harvard Medical School, Boston, MA, USA

Bradford B. Lowell

## Authors

### 1. Janet Berrios

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

### 2. Chia Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Joseph C. Madara

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Alastair S. Garfield

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Jennifer S. Steger

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Michael J. Krashes

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Bradford B. Lowell

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

Conceptualization: J.B. and B.B.L.; methodology: J.B., C.L., J.C.M., A.S.G., J.S.S., M.J.K. and B.B.L.; validation: J.B., C.L. and J.C.M.; formal analysis: J.B. and C.L.; resources: B.B.L. and M.J.K.; writing, original draft: J.B., C.L. and B.B.L.; writing, review and editing: J.B., C.L., J.C.M., AS.G., J.S.S., M.J.K. and B.B.L.; funding acquisition: J.B. and B.B.L.

## Corresponding author

Correspondence to [Bradford B. Lowell](#).

# Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

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

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

## Extended data figures and tables

[Extended Data Fig. 1 DMH<sup>LEPR</sup> neuron photometry responses are greater with pellet drops than with gut-infusions of caloric substances and inhibiting DMH<sup>LEPR</sup> neurons does not affect food intake following an overnight fast.](#)

**a**, Heat maps of additional well-trained mice included in Fig. 1d of DMH<sup>LEPR</sup> neuron fibre photometry responses to cue (light) onset. Each heat map represents a single behavioural session. Single trials have been sorted by response time. Green line indicates light onset (time = 0), black line indicates time when the mouse poked within the correct port. **b**, Heat maps of all naive mice included in Fig. 1d of DMH<sup>LEPR</sup> neuron fibre photometry responses to cue (light) onset. Each heat map represents a single behavioural session. Single trials have been sorted by response time. Green line indicates light onset (time = 0), black line indicates time when the mouse poked within the correct port. It should be noted that naive mice perform significantly fewer trials than experienced mice. **c**, Gastric infusion experimental design. **d–f**, Plot of fibre photometry responses from

$\text{DMH}^{\text{LEPR}}$  neurons aligned to gut infusions of Ensure (**d**, **e**), saline (**d**, **e**), or caloric and non-caloric sweeteners (**f**). These were also aligned to food pellet drops post-infusion. Mice were either food deprived or fed ad libitum for sated Ensure infusions. Light colour tones indicate S.E.M. **g**, Quantification of fluorescence changes ten minutes post-gut infusions of saline, Ensure, and Ensure when animals were fed ad libitum (sated Ensure).  $n = 9$  (mice); Repeated measures one-way ANOVA, \*\*\* $P = 0.0007$ . **h**, Comparison between peak fibre photometry responses during the last 30 s of Ensure infusion (red) and first 30 s of pellet drop response (beige) when animals were given saline infusions. Peak response is significantly larger with pellet drops vs. Ensure infusions.  $n = 9$  (animals); Two-tailed, paired *t*-test, \* $P = 0.0241$ . **i**, hM4Di-mediated inhibition of  $\text{DMH}^{\text{LEPR}}$  neurons does not affect food intake in a post-fast refeeding assay.  $n = 12$  (mice). Two-way ANOVA,  $P = 0.9992$ . All data represents the mean  $\pm$  S.E.M.

### [Source data](#)

### [Extended Data Fig. 2 LH<sup>VGLTU2</sup> → DMH neurons preferentially synapse onto DMH<sup>PDYN-GFP</sup> neurons.](#)

**a**, Sample image of rabies labelling of afferents to  $\text{DMH}^{\text{LEPR}}$  neurons. (Top left) Sample of injection site of AAV-DIO-TVA, AAV-DIO-RG, and EnvA-GFP within the DMH. (Right and bottom row) Sample images of rabies labelled afferents. Scale bar = 200  $\mu\text{M}$ . A/P levels are inferred based on histological landmarks.  $n = 3$  (animals). **b**, Monosynaptic connection probability of optically evoked IPSCs and EPSCs to vDMH<sup>PDYN-GFP</sup> neurons. X-axis = afferent source. (**c**, **d**) Quantification of oIPSC (a) and oEPSC (b) amplitude from all candidate afferents to  $\text{DMH}^{\text{LEPR}}$  neurons in the presence and absence of TTX and 4AP. Two-tailed, paired *t*-test, \* $P \leq 0.05$ . (b)  $P = 1.664$  (NAc),  $P = 0.0285$  (BNST),  $P = 0.6413$  (LH),  $P = 0.0165$  (VTA) (c)  $P = 0.9366$  (LH). **e**, Experimental design schematic. AAV-DIO-ChR2 injected in LH<sup>VGLUT2</sup> neurons and oEPSCs were measured in vDMH<sup>PDYN-GFP</sup> neurons. Bottom: Example current-clamp recordings in vDMH<sup>PDYN-GFP</sup> neuron during optogenetic activation of LH<sup>VGLUT2</sup> neurons. **f**, LH<sup>VGLUT2</sup> neurons preferentially connect to

about 80% of DMH<sup>PDYN-GFP</sup> neurons and have sparse connectivity to DMH<sup>PDYN-GFP</sup> negative neurons.  $n = 15$  (cells) GFP+, 17 (cells) GFP-. **g**, oEPSC amplitude is unchanged when TTX and 4AP are added.  $n = 9$  (ACSF) and 11 (TTX+4AP). Two-tailed unpaired *t*-test,  $P = 0.1666$ . **h**, AAV-DIO-ChR2 was injected into the LH of *Vglut2-IRES-Cre* mice. Whole-cell recordings were performed in vDMH<sup>PDYN-GFP</sup> negative neurons while stimulating LH<sup>VGLUT2</sup> terminals. Bottom: representative ChR2-assisted circuit mapping traces from vDMH<sup>PDYN-GFP</sup> negative neurons in the presence and absence of TTX and 4AP. **i**, Summarized postsynaptic current amplitude in DMH<sup>PDYN-GFP</sup> negative neurons in the presence and absence of TTX and 4AP.  $n = 9$  (ACSF) and 8 (TTX, 4AP). Two-tailed, unpaired *t*-test,  $P = 0.26$ . **j**, Summarized input resistance in both PDYN-GFP positive and negative neurons.  $n = 20$  (GFP+) and 17 (GFP-). Two-tailed, unpaired *t*-test,  $P = 0.14$ . **k**, Summarized series resistance in both PDYN-GFP positive and negative neurons.  $n = 20$  (GFP+) and 17 (GFP-). Two-tailed, unpaired *t*-test,  $P = 0.24$ . **l, m**, Averaged current onset latency (**l**) and jitter (**m**) in PDYN-GFP positive neurons in the absence of TTX and 4AP.  $n = 8$  (cells). All data represents the mean  $\pm$  S.E.M.

### [Source data](#)

### [Extended Data Fig. 3 LH<sup>VGLUT2</sup> → DMH neurons are appetitive, do not send collaterals to the LHb or PAG, have negligible responses to non-edible object presentations, and respond differently to water when compared to food.](#)

**a**, Representative wire plots of an XFP and ChR2 expressing animals throughout the entire 20-min RTPP session across three different conditions: lights-on ad libitum, lights-off ad libitum, and lights-on fasted. Blue line represents stimulation side of the RTPP chamber and black represents the non-stimulation side. **b–d**, Total distance travelled is unchanged between XFP and ChR2 expressing animals throughout the RTPP session across all three conditions (lights-on ad libitum, lights-off ad libitum, and lights-on fasted).  $n = 7$  (animals) for both XFP and ChR2-expressing groups. Unpaired *t*-test,  $P = 0.7498$  (**b**),  $P = 0.5277$  (**c**),  $P = 0.3482$  (**d**). **e**, DIO-TVA-mCherry was injected into the LH of *Vglut2*-

*IRES-Cre* mice and pseudotyped EnvA-GFP was injected into the DMH to map collaterals of LH<sup>VGLUT2</sup> → DMH neurons. **f**, Injection sites within the LH (left) and DMH (right). Scale bar = 500 µM (LH) and 200 µM (DMH). **g**, Collaterals were found within the BNST, MPO/LPO, and AHA and were absent in LH targets known to promote aversion such as the LHb and PAG.  $n = 3$  (animals). Scale bar = 500 µM. **h**, Averaged axonal fibre-photometry response in LH<sup>VGLUT2</sup>→vDMH terminals across all animals when presented with a non-edible object. Vertical line represents object presentation.  $n = 7$  (animals). **i**, Summarized Z-score of LH<sup>VGLUT2</sup>→vDMH axonal fibre photometry response to pellet ( $n = 9$ ) and object ( $n = 7$ ) drops in fasted mice. One-way ANOVA; Kruskal–Wallis, \*\*\* $P < 0.0001$ . **j**, Summarized mean peak response in LH<sup>VGLUT2</sup>→vDMH axonal fibre photometry response to pellet drops when animals were fed ad libitum.  $n = 9$ ; Two-tailed, unpaired *t*-test, \* $P = 0.0129$ . **k**, Averaged axonal fibre-photometry response in LH<sup>VGLUT2</sup>→vDMH terminals across all animals when given a 500 mg food pellet when food restricted (green) or, a bowl of water when water restricted (blue). Vertical line represents food/water presentation. **l**, **m**, Quantification of mean peak response of food (in food-deprived mice) and water presentations (in water-deprived mice) (**l**,  $P = 0.0206$ ).  $n = 9$  (food) and 8 (water). Peak water response occurred at a much later time point than food responses suggesting that the neurons respond differently to water cues when compared to food (**m**,  $P = 0.0391$ ).  $n = 9$  (food) and 8 (water); Two-tailed, unpaired *t*-test, \* $P < 0.05$ . A/P levels are inferred based on histological landmarks from the referenced mouse atlas. All data represents the mean ± S.E.M.

### Source data

### Extended Data Fig. 4 AAV-fDIO-hM4Di efficiently inhibits LH<sup>VGLUT2</sup> neurons, CNO does not affect DMH<sup>LEPR</sup> activation, and hM4Di inhibition does not affect sustained AGRP inhibition.

**a**, Validation of AAV-fDIO-hM4Di construct using whole-cell patch clamp electrophysiology. Upon CNO wash-on (red), the rheobase increase is indicative of a successfully hyperpolarized LH<sup>VGLUT2</sup> neuron given that a

higher injection of current is needed to fire an action potential during CNO-wash on. **b**, CNO wash-on also eliminated spontaneous action potential firing in LH<sup>VGLUT2</sup> neurons in fDIO-hM4Di expressing neurons. **c, d**, XFP controls were used to determine if CNO would affect DMH<sup>LEPR</sup> activation in the absence of hM4Di. CNO did not affect DMH<sup>LEPR</sup> neuron photometry responses (c,  $P = 0.7616$ ) or behavioural performance (d).  $n = 3$  (mice). **e**, Inhibiting LH→DMH afferents does not affect behavioural response time in well-trained mice.  $n = 4$  (mice). Relates to Fig. [3a–c](#). Two-tailed *t*-test,  $P = 0.7068$ . **f**, CNO treatment does not affect the later consummatory response in AGRP neurons when mice are given a large (250 mg) food pellet.  $n = 4$  (mice); Two-tailed *t*-test,  $P = 0.61$ . **g**, Inhibiting LH→DMH afferents in well-trained mice does not affect behavioural performance.  $n = 4$  (mice), Two-tailed *t*-test,  $P = 0.3819$ . Relates to Fig. [3a–c](#). **h**, Inhibition of LH<sup>VGLUT2</sup> neurons does not affect behavioural performance.  $n = 4$  (mice), Two-tailed *t*-test  $P = 0.3661$ . Relates to Fig. [3d–f](#). **i, j**, Inhibition of LH<sup>VGLUT2</sup> neurons does not affect behavioural accuracy (**i**) or response time (**j**).  $n = 4$  (mice), Two-tailed *t*-test,  $P = 0.7107$ . Relates to Fig. [3g–i](#). All data represents the mean ± S.E.M.

### Source data

### Extended Data Fig. 5 TeNT-mediated silencing of DMH<sup>LEPR</sup>, DMH<sup>PDYN</sup>, and LH→DMH neurons significantly attenuates learning a cue-initiated food acquisition task by increasing mistakes made during the task but TeNT-mediated silencing of DMH<sup>LEPR</sup> neurons does not affect learning a cue-initiated water acquisition task.

**a, b**, DMH<sup>LEPR</sup> neuron silencing leads to an increase in response errors in the 2AFC task. TeNT-expressing mice largely miss the response window (**a**) in the initial phases of training and have an increase in false alarms (**b**).  $n = 5$  (GFP), 5 (TeNT). **a**, Red dotted line indicates line of best fit ( $***P < 0.0001$ , plateau = 5.735 (GFP), -2.910 (TeNT); tau = 2.923 (GFP), 12.64 (TeNT)). Two-way repeated measures ANOVA, main effect of days =  $****P < 0.0001$ , main effect of group = 0.05. **b**, Red dotted line

indicates line of best fit ( $***P < 0.0001$ , plateau = 0.1324 (GFP), -162.8 (TeNT); tau = 6.679 (GFP), 117.6 (TeNT)). Two-way repeated measures ANOVA, main effect of days =  $***P < 0.0001$ , main effect of group = 0.6. **c**, TeNT-expressing *Lepr-IRES-Cre* mice take significantly longer to nose poke following cue-light presentation.  $n = 5$  (GFP), 5 (TeNT). Red dotted line indicates line of best fit ( $***P < 0.0001$ , plateau = 1.319 (GFP), 0.633 (TeNT); tau = 4.272 (GFP), 9.860 (TeNT)). Two-way repeated measures ANOVA, main effect of days =  $**P < 0.01$ , main effect of group = 0.08. **d**, Body weight is unaffected upon DMH<sup>LEPR</sup> silencing with TeNT (7 weeks post-surgery).  $n = 5$  (GFP), 5 (TeNT). Two-tailed, unpaired *t*-test,  $P = 0.7043$ . **e**, Food-intake does not significantly differ with DMH<sup>LEPR</sup> silencing following an overnight fast.  $n = 5$  (GFP), 6 (TeNT). Two-tailed, unpaired *t*-test,  $P = 0.9495$ . **f**, AAV-DIO-TeNT injection into DMH<sup>PDYN</sup> neurons. Bottom: Example images of TeNT expression. Scale bar = 500  $\mu\text{M}$ .  $n = 5$  (mice). **g, h**, Similar to (Fig. 2b) in DMH<sup>PDYN</sup> neurons.  $n = 7$  (GFP), 5 (TeNT). Red line indicates = line of best fit ( $***P < 0.0001$ , plateau = 82.83 (GFP), 65.52 (TeNT); tau = 5.619 (GFP), 7.337 (TeNT)). Two-way repeated measures ANOVA, main effect of days =  $***P < 0.0001$ , main effect of group =  $**P < 0.01$ . **h**, Similar to (Fig. 2c) in DMH<sup>PDYN</sup> neurons (learning criterion  $>70\%$  correct across three consecutive days).  $>20$  on the *y*-axis is to indicate that by day 20, mice had not learned. 7/7 GFP and 1/6 TeNT mice learned.  $n = 7$  (GFP), 5 (TeNT). Two-tailed, unpaired *t*-test,  $**P = 0.0094$ . **i, j**, DMH<sup>PDYN</sup> neuron silencing leads to an increase in response errors in the 2AFC task. TeNT-expressing mice largely miss the response window (**i**) in the initial phases of training and do not have a difference in false alarm rate (**j**).  $n = 7$  (GFP), 5 (TeNT). **(i)** Red dotted line indicates line of best fit ( $***P < 0.0001$ , plateau = 7.142 (GFP), 20.89 (TeNT); tau = 0.8597 (GFP), 1.96 (TeNT)). Two-way repeated measures ANOVA, main effect of days =  $***P < 0.001$ , main effect of group =  $*P < 0.05$ . **j**, Line of best fit could not be calculated. Two-way repeated measures ANOVA, main effect of days =  $***P < 0.001$ , main effect of group = 0.8267. **k**, TeNT-expressing Pdyn-Cre mice take significantly longer to nose poke following cue-light presentation.  $n = 7$  (GFP), 5 (TeNT). Red dotted line indicates line of best fit ( $***P < 0.0001$ , plateau = 1.185 (GFP), 2.748 (TeNT); tau = 2.122 (GFP), 3.939 (TeNT)). Two-way repeated measures ANOVA, main effect of days =  $***P < 0.0001$ , main effect of group =  $***P < 0.0001$ . **l**, Body

weight increases upon long term DMH<sup>PDYN</sup> silencing with TeNT (20 weeks post-surgery).  $n = 7$  (GFP), 5 (TeNT). Two-tailed, unpaired *t*-test, \*\* $P = 0.0086$ . **m**, Food-intake does not significantly differ with DMH<sup>PDYN</sup> silencing following an overnight fast.  $n = 7$  (GFP), 5 (TeNT). Two-tailed, unpaired *t*-test,  $P = 0.2668$ . **n**, AAV-DIO-TeNT injection into the LH and rAAV-Cre injection into the DMH in wild-type mice. Bottom: Example images of TeNT expression. Scale bar = 500  $\mu\text{M}$ .  $n = 6$  (mice). **o, p**, Similar to (Fig. 2b) in LH $\rightarrow$ DMH neurons.  $n = 5$  (GFP), 6 (TeNT). Red line = line of best fit (\*\* $P < 0.0001$ , slope = 2.033 (GFP), 1.488 (TeNT); Y-intercept = 44.50 (GFP), 30.02 (TeNT)). Two-way repeated measures ANOVA, main effect of days = \*\*\* $P < 0.0001$ , main effect of group =  $P = 0.0548$ . **p**, Similar to (Fig. 2c) in LH $\rightarrow$ DMH neurons (learning criterion  $>70\%$  correct across three consecutive days). 5/5 GFP and 2/6 TeNT-expressing mice learned.  $n = 5$  (GFP), 6 (TeNT). Two-tailed, unpaired *t*-test,  $P = 0.0789$ . **q, r**, LH $\rightarrow$ DMH neuron silencing leads to an increase in response errors in the 2AFC task. TeNT-expressing mice largely miss the response window (**q**) in the initial phases of training and do not have a difference in false alarm rate (**r**).  $n = 5$  (GFP), 6 (TeNT). **q**, Red dotted line indicates line of best fit (\*\* $P < 0.0001$ , slope = -0.299 (GFP), -0.4169 (TeNT); Y-intercept = 16.75 (GFP), 39.91 (TeNT)). Two-way repeated measures ANOVA, main effect of days = \* $P = 0.0416$ , main effect of group =  $P = 0.1896$ . **r**, Line of best fit is not significantly different for TeNT and GFP-expressing animals ( $P = 0.1724$ , plateau = -22.34 (GFP), -74.42 (TeNT); Tau = 23.91 (GFP), 83.89 (TeNT)). Two-way repeated measures ANOVA, main effect of days = \* $P = 0.0101$ , main effect of group = 0.9094. **s**, TeNT-expression in LH $\rightarrow$ DMH neurons significantly increases nose poke response following cue-light presentation.  $n = 5$  (GFP), 6 (TeNT). Red dotted line indicates line of best fit (\*\* $P < 0.0001$ , plateau = 1.790 (GFP), 3.314 (TeNT); tau = 1.818 (GFP), 8.042 (TeNT)). Two-way repeated measures ANOVA, main effect of days = \*\*\* $P = 0.0004$ , main effect of group = \*\*\* $P < 0.0001$ . **t**, Body weight is unaffected upon LH $\rightarrow$ DMH silencing with TeNT (8 weeks post-surgery).  $n = 5$  (GFP), 6 (TeNT). Two-tailed, unpaired *t*-test,  $P = 0.2820$ . **u**, Food-intake does not significantly differ with LH $\rightarrow$ DMH silencing following an overnight fast.  $n = 5$  (GFP), 6 (TeNT). Two-tailed, unpaired *t*-test,  $P = 0.8666$ . **v–x**, DMH<sup>LEPR</sup> neuron silencing does not affect response errors or response

times in the 2AFC task when dehydrated mice receive water rewards.  $n = 4$  (GFP), 6 (TeNT). **v**, Red dotted line indicates line of best fit (plateau = 57.74, tau = 4.477). Two-way repeated measures ANOVA, main effect of days = \*\*\* $P < 0.0001$ , main effect of group = 0.942. **w**, Red dotted line indicates line of best fit (plateau = -15.32, tau = 11.67). Two-way repeated measures ANOVA, main effect of days = \* $P < 0.05$ , main effect of group = 0.13. **x**, Red dotted line indicates line of best fit (plateau = 0.7882, tau = 2.929). Two-way repeated measures ANOVA, main effect of days = \*\*\* $P < 0.0001$ , main effect of group = 0.82. All data represents the mean  $\pm$  S.E.M.

### Source data

### Extended Data Fig. 6 PDYN-expressing DMH neurons exclusively project to the ARC.

AAV-DIO-ChR2-mCherry was injected into the DMH of *Pdyn-Cre* mice to search for long range projections of DMH<sup>PDYN</sup> neurons. The ARC was the only area of the brain that received DMH<sup>PDYN</sup> innervation. A portion of this data was previously published in Garfield et al. 2016<sup>12</sup>.

### Extended Data Fig. 7 Schematic representation of viral spread and fibre placement.

**a–c**, Schematic representing the extent of viral spread (transparent shaded regions) and fibre placement (solid ovals) for every animal related to studies presented in Fig. 1b-d, Extended Data Fig. 1a (a), Fig. 2a-e, Extended Data Fig. 3a-d (b) and Fig. 2f-k, Extended Data Fig. 3h-m (c). Please note that different animals are represented by different colours and that “hits” and “misses” are represented on different hemispheres (miss is on the left). GENERAL CRITERION FOR HITS AND MISSES: Animals were deemed a “hit” if viral expression was within the area of interest and fibre placement was either within or immediately dorsal to the area of interest (that is, the DMH or the LH). Excluded animals are specified by the “miss” label. Reason each animal was considered a “miss”: Animal 5; fibre placement dorso-lateral to DMH (a); Animal 10; unilateral viral expression

and fibre placement lateral to midline/DMH (**b**); Animal 10; viral expression ventro-medial to LH, fibre placement lateral to the DMH (**c**).

### Extended Data Fig. 8 Schematic representation of viral spread and fibre placement within *Lepr-Cre* animals.

Schematic representing the extent of viral spread (transparent shaded regions) and fibre placement (solid ovals) for every animal related to studies presented in Fig. 3a–c and Extended Data Fig. 4e–g. Please note that individual animals are represented by different colours. Bilateral hM4Di injections were targeted to the LH and bilateral AAV6-FlpO injections were targeted to the DMH. GCaMP6s injections into the DMH were unilateral. CRITERION FOR HITS AND MISSES: Animals were deemed a “hit” if viral expression was within the area of interest (that is, hM4Di bilateral in the LH, GCaMP unilateral in the DMH) and fibre placement was either within or immediately dorsal to the area of interest (that is, the DMH). For clarity, hits are shown on the top and misses on the bottom. Reason each animal was considered a “miss”: Animal 1; viral expression (hM4Di) was medial to DMH, fibre placement ventral to DMH. Animal 2; viral expression (hM4Di) was unilateral, fibre placement ventral to DMH. Animal 3; viral expression (hM4Di) was unilateral, fibre placement dorsal to DMH. Animal 4; hM4Di expression undetected, fibre placement ventral to DMH.

### Extended Data Fig. 9 Schematic representation of viral spread and fibre placement within *Vglut2-FlpO::Lepr-Cre* animals.

Schematic representing the extent of viral spread (transparent shaded regions) and fibre placement (solid ovals) for every animal related to studies presented in Fig. 3d–f and Extended Data Fig. 4h. Please note that individual animals are represented by different colours. Bilateral hM4Di injections were targeted to the LH and GCaMP6s injections were unilateral into the DMH. GENERAL CRITERION FOR HITS AND MISSES: Animals were deemed a “hit” if viral expression was within the area of interest (that is, hM4Di bilateral in the LH, GCaMP unilateral in the DMH) and fibre placement was either within or immediately dorsal to the area of

interest (that is, the DMH). For clarity, hits are shown on the top and misses on the bottom. Reason each animal was considered a “miss”: Animal 1; viral expression of hM4Di was dorsal to LH. Animal 2; hM4Di viral expression was unilateral and dorsal to LH and minimal GCaMP6s expression was detected lateral to DMH.

### **Extended Data Fig. 10 Schematic representation of viral spread and fibre placement within *Vglut2-FlpO::AGRPs-Cre* animals.**

Schematic representing the extent of viral spread (transparent shaded regions) and fibre placement (solid ovals) for every animal related to studies presented in Fig. 3g–i and Extended Data Fig. 4i,j. Please note that different animals are represented by different colours. Bilateral hM4Di injections were targeted to the LH and GCaMP6s injections were unilateral into the ARC. Purple regions indicate areas with no viral expression. GENERAL CRITERION FOR HITS AND MISSES: Animals were deemed a “hit” if viral expression was within the area of interest (that is, hM4Di bilateral in the LH, GCaMP unilateral in the ARC) and fibre placement was either within or immediately dorsal to the area of interest (that is, the ARC). For clarity, hits are shown on the top and misses on the bottom. Reason each animal was considered a “miss”: Animal 1; viral expression of hM4Di was unilateral and dorsal to LH and GCaMP6s expression was undetectable. Animal 2; viral expression of hM4Di was unilateral and dorsal to LH and GCaMP6s expression was undetectable.

### **Extended Data Fig. 11 Schematic representation of viral spread.**

**a–d**, Schematic representing the extent of viral spread (transparent shaded regions) for every animal related to studies presented in Fig. 4a–c, Extended Data Fig. 5a–e (a), Extended Data Fig. 5f–m (b), Extended Data Fig. 5n–u (c) and Fig. 4d–f and Extended Data Fig. 5v–x (d). Please note that different animals are represented by different colours. GENERAL CRITERION FOR HITS AND MISSES: Animals were deemed a “hit” if viral expression was within the area of interest (that is, the DMH or the LH). Excluded animals are specified by the “miss” label. Reason each

animal was considered a “miss”: Animal 7, viral expression had spread to the DMH and VMH (**a**); Animal 6, viral expression was dorsal to DMH (**b**); Animal 7, viral expression was barely detected (**b**); Animal 7, viral expression was dorsal to LH (**c**); Animal 8, diffuse, widespread viral expression dorsal to LH (**c**); Animal 9, viral expression was dorsal to LH and within zona incerta (**c**); Animal 10, viral expression was dorsal to LH and within zona incerta (**c**); Animal 7; viral expression was dorsal to DMH (**d**); Animal 8, viral expression was barely detectable in a small region within the anterior DMH (**d**).

### [Extended Data Fig. 12 Example of method used to draw anatomical boundaries in histological images.](#)

DAPI-stained images were taken and contrast was enhanced (left). Images were matched with the appropriate A/P coordinate within the histological atlas (middle) then, boundaries were drawn using the axis as a template (right). All anatomical boundaries represented within the figures were drawn using this method. The atlas used for these studies was *The Mouse Brain in Stereotaxic Coordinates* by Paxinos and Franklin, Second Edition.

## **Supplementary information**

### [Supplementary Information](#)

This file contains supplementary discussion.

### [Reporting Summary](#)

## **Source data**

### [Source Data Fig. 1](#)

### [Source Data Fig. 2](#)

### [Source Data Fig. 3](#)

[\*\*Source Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 1\*\*](#)

[\*\*Source Data Extended Data Fig. 2\*\*](#)

[\*\*Source Data Extended Data Fig. 3\*\*](#)

[\*\*Source Data Extended Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 5\*\*](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Berrios, J., Li, C., Madara, J.C. *et al.* Food cue regulation of AGRP hunger neurons guides learning. *Nature* **595**, 695–700 (2021).  
<https://doi.org/10.1038/s41586-021-03729-3>

[Download citation](#)

- Received: 13 November 2019
- Accepted: 15 June 2021
- Published: 14 July 2021

- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03729-3>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03729-3>

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

- Article
- [Published: 14 July 2021](#)

# Astrocytic interleukin-3 programs microglia and limits Alzheimer's disease

- [Cameron S. McAlpine<sup>1,2,3</sup>](#),
- [Joseph Park<sup>4</sup>](#),
- [Ana Graciuc<sup>4</sup>](#),
- [Eunhee Kim<sup>4</sup>](#),
- [Se Hoon Choi<sup>4</sup>](#),
- [Yoshiko Iwamoto<sup>1</sup>](#),
- [Máté G. Kiss](#) [ORCID: orcid.org/0000-0002-9215-8328<sup>1</sup>](#),
- [Kathleen A. Christie<sup>5</sup>](#),
- [Claudio Vinegoni](#) [ORCID: orcid.org/0000-0002-9818-3639<sup>1</sup>](#),
- [Wolfram C. Poller<sup>1,2</sup>](#),
- [John E. Mindur](#) [ORCID: orcid.org/0000-0002-2212-3904<sup>1</sup>](#),
- [Christopher T. Chan<sup>1</sup>](#),
- [Shun He<sup>1</sup>](#),
- [Henrike Janssen<sup>1</sup>](#),
- [Lai Ping Wong<sup>6,7</sup>](#),
- [Jeffrey Downey<sup>1</sup>](#),
- [Sumnima Singh<sup>1</sup>](#),
- [Atsushi Anzai](#) [ORCID: orcid.org/0000-0002-7474-1309<sup>1</sup>](#),
- [Florian Kahles<sup>1</sup>](#),
- [Mehdi Jorfi](#) [ORCID: orcid.org/0000-0003-4724-6190<sup>4</sup>](#),
- [Paolo Fumene Feruglio<sup>8</sup>](#),
- [Ruslan I. Sadreyev<sup>6,9</sup>](#),
- [Ralph Weissleder](#) [ORCID: orcid.org/0000-0003-0828-4143<sup>1</sup>](#),

- [Benjamin P. Kleinstiver](#) ORCID: [orcid.org/0000-0002-5469-0655<sup>5</sup>](https://orcid.org/0000-0002-5469-0655),
- [Matthias Nahrendorf](#) ORCID: [orcid.org/0000-0002-4021-1887<sup>1</sup>](https://orcid.org/0000-0002-4021-1887),
- [Rudolph E. Tanzi](#) ORCID: [orcid.org/0000-0002-7032-1454<sup>4</sup>](https://orcid.org/0000-0002-7032-1454) &
- [Filip K. Swirski](#) ORCID: [orcid.org/0000-0002-3163-9152<sup>1,2</sup>](https://orcid.org/0000-0002-3163-9152)

[Nature](#) volume **595**, pages 701–706 (2021) [Cite this article](#)

- 13k Accesses
- 202 Altmetric
- [Metrics details](#)

## Subjects

- [Astrocyte](#)
- [Interleukins](#)
- [Microglial cells](#)
- [Neurodegeneration](#)
- [Neuroimmunology](#)

## Abstract

Communication within the glial cell ecosystem is essential for neuronal and brain health<sup>1,2,3</sup>. The influence of glial cells on the accumulation and clearance of β-amyloid (Aβ) and neurofibrillary tau in the brains of individuals with Alzheimer’s disease (AD) is poorly understood, despite growing awareness that these are therapeutically important interactions<sup>4,5</sup>. Here we show, in humans and mice, that astrocyte-sourced interleukin-3 (IL-3) programs microglia to ameliorate the pathology of AD. Upon recognition of Aβ deposits, microglia increase their expression of IL-3Rα—the specific receptor for IL-3 (also known as CD123)—making them responsive to IL-3. Astrocytes constitutively produce IL-3, which elicits transcriptional, morphological, and functional programming of microglia to endow them with an acute immune response program, enhanced motility, and the capacity to cluster and clear aggregates of Aβ and tau. These

changes restrict AD pathology and cognitive decline. Our findings identify IL-3 as a key mediator of astrocyte–microglia cross-talk and a node for therapeutic intervention in AD.

## 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

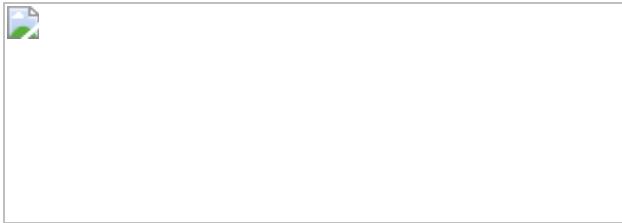
[Rent or Buy](#)

All prices are NET prices.

### Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

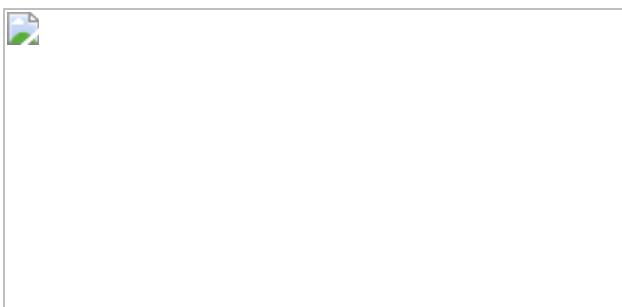
**Fig. 1: IL-3 protects against A $\beta$  accumulation and cognitive impairment in 5xFAD mice.**



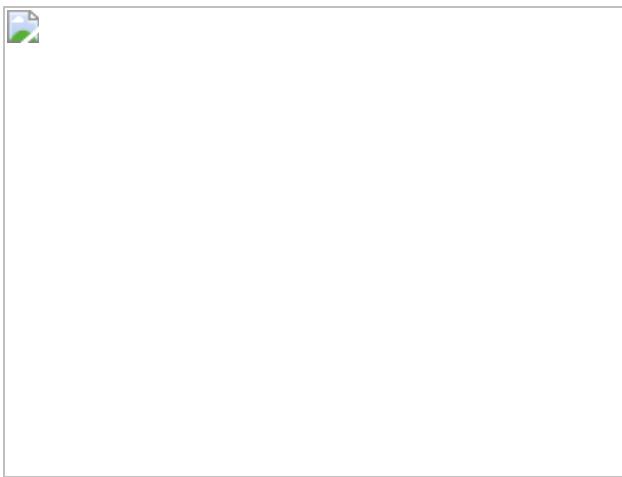
**Fig. 2: Microglia become responsive to astrocyte-derived IL-3 in AD.**



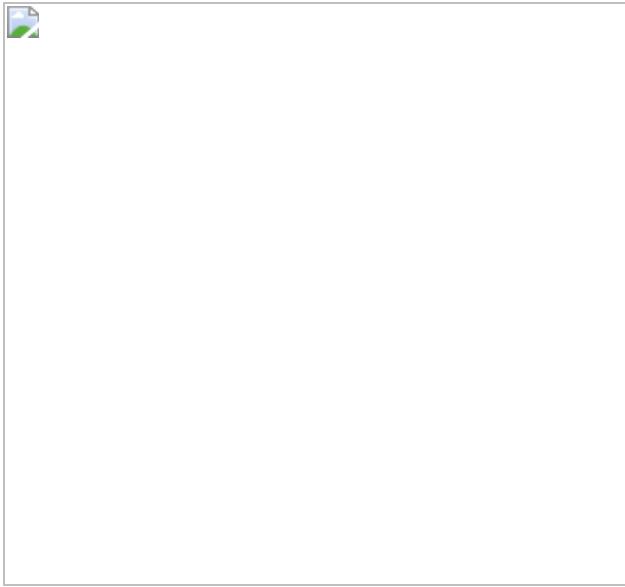
**Fig. 3: IL-3 signalling correlates with disease pathology in the brains of humans with AD.**



**Fig. 4: IL-3 programs microglia to promote motility and clustering of A $\beta$  in the mouse brain and in a 3D human microfluidic triculture system.**



**Fig. 5: Astrocyte IL-3 or microglia IL-3Ra deletion instigate, while IL-3 infusion resolves, A $\beta$  burden and cognitive decline.**



## Data availability

RNA-seq data for microglia from *Il3<sup>-/-</sup>5xFAD* and *5xFAD* mice have been deposited to NCBI-GEO under accession [GSE163289](#). RNA-seq data for WT, *5xFAD*, *Trem2<sup>-/-</sup>* and *Trem2<sup>-/-</sup>5xFAD* mice have been previously described<sup>24</sup> and were deposited to GEO: [GSE132508](#). scRNA sequencing data for homeostatic, DAM stage 1, and DAM stage 2 microglia have been previously described<sup>23</sup> and were deposited to GEO: [GSE98969](#). [Source data](#) are provided with this paper.

## References

1. 1.

Linnerbauer, M., Wheeler, M. A. & Quintana, F. J. Astrocyte crosstalk in CNS inflammation. *Neuron* **108**, 608–622 (2020).

[CAS Article](#) [Google Scholar](#)

2. 2.

Vainchtein, I. D. & Molofsky, A. V. Astrocytes and microglia: in sickness and in health. *Trends Neurosci.* **43**, 144–154 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Castellani, G. & Schwartz, M. Immunological features of non-neuronal brain cells: implications for Alzheimer's disease immunotherapy. *Trends Immunol.* **41**, 794–804 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Fakhoury, M. Microglia and astrocytes in Alzheimer's disease: implications for therapy. *Curr. Neuropharmacol.* **16**, 508–518 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Long, J. M. & Holtzman, D. M. Alzheimer disease: an update on pathobiology and treatment strategies. *Cell* **179**, 312–339 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Mindur, J. E. & Swirski, F. K. Growth factors as immunotherapeutic targets in cardiovascular disease. *Arterioscler. Thromb. Vasc. Biol.* **39**, 1275–1287 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Gómez Ravetti, M. & Moscato, P. Identification of a 5-protein biomarker molecular signature for predicting Alzheimer's disease. *PLoS ONE* **3**, e3111 (2008).

[ADS](#) [Article](#) [Google Scholar](#)

8. 8.

Ray, S. et al. Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. *Nat. Med.* **13**, 1359–1362 (2007).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Britschgi, M. et al. Modeling of pathological traits in Alzheimer's disease based on systemic extracellular signaling proteome. *Mol. Cell. Proteomics* **10**, 008862 (2011).

[Article](#) [Google Scholar](#)

10. 10.

Soares, H. D. et al. Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer disease. *Arch. Neurol.* **69**, 1310–1317 (2012).

[Article](#) [Google Scholar](#)

11. 11.

Huberman, M. et al. Correlation of cytokine secretion by mononuclear cells of Alzheimer patients and their disease stage. *J. Neuroimmunol.* **52**, 147–152 (1994).

[CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Kiddle, S. J. et al. Plasma based markers of [11C] PiB-PET brain amyloid burden. *PLoS ONE* **7**, e44260 (2012).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Frei, K., Bodmer, S., Schwerdel, C. & Fontana, A. Astrocytes of the brain synthesize interleukin 3-like factors. *J. Immunol.* **135**, 4044–4047 (1985).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Frei, K., Bodmer, S., Schwerdel, C. & Fontana, A. Astrocyte-derived interleukin 3 as a growth factor for microglia cells and peritoneal macrophages. *J. Immunol.* **137**, 3521–3527 (1986).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Zambrano, A., Otth, C., Maccioni, R. B. & Concha, I. I. IL-3 controls tau modifications and protects cortical neurons from neurodegeneration. *Curr. Alzheimer Res.* **7**, 615–624 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Zambrano, A., Otth, C., Mujica, L., Concha, I. I. & Maccioni, R. B. Interleukin-3 prevents neuronal death induced by amyloid peptide. *BMC Neurosci.* **8**, 82 (2007).

[Article](#) [Google Scholar](#)

17. 17.

Herisson, F. et al. Direct vascular channels connect skull bone marrow and the brain surface enabling myeloid cell migration. *Nat. Neurosci.* **21**, 1209–1217 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Gate, D. et al. Clonally expanded CD8 T cells patrol the cerebrospinal fluid in Alzheimer's disease. *Nature* **577**, 399–404 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Zenaro, E. et al. Neutrophils promote Alzheimer's disease-like pathology and cognitive decline via LFA-1 integrin. *Nat. Med.* **21**, 880–886 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Pasciuto, E. et al. Microglia require CD4 T cells to complete the fetal-to-adult transition. *Cell* **182**, 625–640 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Anzai, A. et al. Self-reactive CD4<sup>+</sup> IL-3<sup>+</sup> T cells amplify autoimmune inflammation in myocarditis by inciting monocyte chemotaxis. *J. Exp. Med.* **216**, 369–383 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Zhou, Y. et al. Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat. Med.* **26**, 131–142 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Keren-Shaul, H. et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Griciuc, A. et al. TREM2 acts downstream of CD33 in modulating microglial pathology in Alzheimer's disease. *Neuron* **103**, 820–835 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Choi, S. H. et al. A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature* **515**, 274–278 (2014).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Park, J. et al. A 3D human triculture system modeling neurodegeneration and neuroinflammation in Alzheimer's disease. *Nat. Neurosci.* **21**, 941–951 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Oakley, H. et al. Intraneuronal β-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: potential factors in amyloid plaque formation. *J. Neurosci.* **26**, 10129–10140 (2006).

[CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Weber, G. F. et al. Interleukin-3 amplifies acute inflammation and is a potential therapeutic target in sepsis. *Science* **347**, 1260–1265 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Turnbull, I. R. et al. Cutting edge: TREM-2 attenuates macrophage activation. *J. Immunol.* **177**, 3520–3524 (2006).

[CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Bae, S., Park, J. & Kim, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Hsiao, T. et al. Inference of CRISPR edits from Sanger trace data. Preprint at <https://doi.org/10.1101/251082> (2018).

33. 33.

Kleinstiver, B. P. et al. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* **37**, 276–282 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

34. 34.

Rohland, N. & Reich, D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* **22**, 939–946 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

35. 35.

Robbins, C. S. et al. Local proliferation dominates lesional macrophage accumulation in atherosclerosis. *Nat. Med.* **19**, 1166–1172 (2013).

[CAS](#) [Article](#) [Google Scholar](#)

36. 36.

DeVos, S. L. & Miller, T. M. Direct intraventricular delivery of drugs to the rodent central nervous system. *J. Vis. Exp.* **75**, e50326 (2013).

[Google Scholar](#)

37. 37.

McQuade, A. et al. Development and validation of a simplified method to generate human microglia from pluripotent stem cells. *Mol. Neurodegener.* **13**, 67 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

38. 38.

Griciuc, A. et al. Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. *Neuron* **78**, 631–643 (2013).

[CAS](#) [Article](#) [Google Scholar](#)

39. 39.

Kraeuter, A. K., Guest, P. C. & Sarnyai, Z. The Y-maze for assessment of spatial working and reference memory in mice. *Methods Mol. Biol.* **1916**, 105–111 (2019).

40. 40.

Vorhees, C. V. & Williams, M. T. Morris water maze: procedures for assessing spatial and related forms of learning and memory. *Nat. Protocols* **1**, 848–858 (2006).

[Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank the HCI-CRM Flow Cytometry Core Facility at the Massachusetts General Hospital for assistance in cell sorting; the MGH DF/HCC Specialized Histopathology Services core and the Hope Babette Tang Histology Facility at the Massachusetts Institute of Technology for tissue sectioning and histology services; the MGH NexGen sequencing and Bioinformatics facility for RNA-seq experiments and analysis; L. Wu and the Harvard Genome Modification Facility for help generating mice; G. Wojtkiewicz for help with imaging software; and K. Joyes for copy editing. This work was funded by the Cure Alzheimer’s Fund, the National Institutes of Health (NIH) R35 HL135752, P01 HL131478, P01 HL142494, and the Patricia and Scott Eston MGH Research Scholar (to F.K.S.); NIH K99/R00 HL151750, R01 HL158534, and a Canadian Institutes of Health Research Banting Fellowship (to C.S.M.); HL142494 (to M.N. and B.P.K.); NIH R35HL139598 (to M.N.); NIH Ruth L. Kirschstein National Research Service Award Individual Predoctoral Fellowship F31HL147364 (to J.E.M.); and NIH R00 CA218870 (to B.P.K.).

## Author information

## Affiliations

1. Center for Systems Biology and Department of Radiology,  
Massachusetts General Hospital and Harvard Medical School, Boston,  
MA, USA

Cameron S. McAlpine, Yoshiko Iwamoto, Máté G. Kiss, Claudio Vinegoni, Wolfram C. Poller, John E. Mindur, Christopher T. Chan, Shun He, Henrike Janssen, Jeffrey Downey, Sumnima Singh, Atsushi Anzai, Florian Kahles, Ralph Weissleder, Matthias Nahrendorf & Filip K. Swirski

2. Cardiovascular Research Institute and Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Cameron S. McAlpine, Wolfram C. Poller & Filip K. Swirski

3. Friedman Brain Institute and Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Cameron S. McAlpine

4. Genetics and Aging Research Unit, McCance Center for Brain Health, Mass General Institute for Neurodegenerative Disease, Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA

Joseph Park, Ana Griciuc, Eunhee Kim, Se Hoon Choi, Mehdi Jorfi & Rudolph E. Tanzi

5. Center for Genomic Medicine, Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Kathleen A. Christie & Benjamin P. Kleinstiver

6. Department of Molecular Biology, Massachusetts General Hospital,  
Boston, MA, USA

Lai Ping Wong & Ruslan I. Sadreyev

7. Department of Genetics, Harvard Medical School, Boston, MA, USA

Lai Ping Wong

8. Department of Neuroscience, Biomedicine and Movement Sciences,  
University of Verona, Verona, Italy

Paolo Fumene Feruglio

9. Department of Pathology, Massachusetts General Hospital and  
Harvard Medical School, Boston, MA, USA

Ruslan I. Sadreyev

## Authors

1. Cameron S. McAlpine  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Joseph Park  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Ana Griciuc  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Eunhee Kim  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Se Hoon Choi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Yoshiko Iwamoto

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Máté G. Kiss

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Kathleen A. Christie

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Claudio Vinegoni

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Wolfram C. Poller

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. John E. Mindur

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Christopher T. Chan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Shun He

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Henrike Janssen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Lai Ping Wong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Jeffrey Downey

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Sumnima Singh

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Atsushi Anzai

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Florian Kahles

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Mehdi Jorfi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Paolo Fumene Feruglio  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Ruslan I. Sadreyev  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Ralph Weissleder  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Benjamin P. Kleinstiver  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Matthias Nahrendorf  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Rudolph E. Tanzi  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Filip K. Swirski  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

C.S.M., J.P., A.G., E.K., S.H.C., Y.I., M.G.K., K.A.C., C.V., W.C.P., J.E.M., C.T.C., S.H., H.J., L.P.W., J.D., S.S., A.A., F.K., M.J., and P.F.F. conducted experiments, and collected and analysed data. C.S.M., J.P., A.G., R.I.S., R.W., B.P.K., M.N., R.E.T., and F.K.S. conceptualized and designed experiments, discussed results, and interpreted data. C.S.M. and F.K.S. designed figures and wrote the manuscript. R.E.T. and F.K.S. supervised, directed, and managed the study.

## Corresponding authors

Correspondence to [Rudolph E. Tanzi](#) or [Filip K. Swirski](#).

## Ethics declarations

### Competing interests

C.S.M., F.K.S., and R.E.T. are inventors on a patent application filed by Mass General Brigham that describes targeting IL-3 signalling in AD (invention record no. 2020-568). B.P.K. is an inventor on patent applications filed by Mass General Brigham that describe genome engineering technologies and methods, is an advisor to Acrigen Biosciences, and consults for Avectas Inc. and ElevateBio.

## Additional information

**Peer review information** *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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

## Extended data figures and tables

[Extended Data Fig. 1 Analysis of \*Il3<sup>-/-</sup>\* mice.](#)

**a**, FITC-dextran (mol. wt. 4000) was injected i.v. into WT and *Il3<sup>-/-</sup>* mice before death. Blood–brain barrier integrity was determined by measuring the FITC signal in brain homogenate ( $n = 6$  WT mice;  $n = 4$  *Il3<sup>-/-</sup>* mice). **b**, Before death, WT and *Il3<sup>-/-</sup>* mice were injected i.v. with an anti-GR1 antibody conjugated to PE to label all circulating monocytes and neutrophils. PE signal among CD45<sup>+</sup> cells was assessed in the brain by flow cytometry ( $n = 7$  WT mice;  $n = 4$  *Il3<sup>-/-</sup>* mice). **c**, Doublecortin staining and quantification in the hippocampus of WT and *Il3<sup>-/-</sup>* mice at 4 months of age ( $n = 3$ ). **d**, Absence of caspase 3 staining in the hippocampus of WT and *Il3<sup>-/-</sup>* mice along with a representative image of rare positively stained cells from the thalamus ( $n = 3$ ). **e**, Assessment of MHCII<sup>+</sup> microglia in the brains of WT and *Il3<sup>-/-</sup>* mice ( $n = 7$  WT mice;  $n = 4$  *Il3<sup>-/-</sup>* mice). **f**, Analysis of Ki67<sup>+</sup> proliferating microglia ( $n = 5$ ). **g**, Time in new arm during Y-maze testing ( $n = 7$ ). Groups of mice are of evenly mixed sex. Mean ± s.e.m.

## Extended Data Fig. 2 Haematopoiesis and peripheral immune cell dynamics in WT, 5xFAD, and *Il3<sup>-/-</sup>*5xFAD mice.

**a**, Average swim speed during acquisition days of Morris water maze ( $n = 10$  5xFAD mice;  $n = 9$  *Il3<sup>-/-</sup>*5xFAD mice). **b**, Flow cytometry assessment of blood leukocytes ( $n = 9$  WT mice;  $n = 12$  5xFAD mice;  $n = 8$  *Il3<sup>-/-</sup>*5xFAD mice). One-way ANOVA. **c**, Flow cytometry analysis of Lin<sup>-</sup>SCA1<sup>+</sup>cKIT<sup>+</sup> cells (LSKs), multi-potent progenitors (MPP)-4 and -3, short-term haematopoietic stem cells (StHSCs), long-term HSCs (LtHSCs), common myeloid progenitors (CMPs), granulocyte macrophage progenitors (GMPs), and monocyte dendritic progenitors (MDPs) in the bone marrow of 5-month-old WT, 5xFAD, and *Il3<sup>-/-</sup>*5xFAD mice ( $n = 9$  WT mice;  $n = 11$  5xFAD mice;  $n = 8$  *Il3<sup>-/-</sup>*5xFAD mice). One-way ANOVA. **d**, FITC-dextran (mol. wt. 4,000) was injected i.v. into 5xFAD and *Il3<sup>-/-</sup>*5xFAD mice before death. Blood–brain barrier integrity was determined by measuring the FITC signal in brain homogenates ( $n = 4$ ). **e**, WT, 5xFAD, *Il3<sup>-/-</sup>*5xFAD mice were joined by parabiosis with UbiGFP mice from the age of 2 to 6 months (4 months total). **f**, GFP chimerism in blood Ly6C<sup>hi</sup> monocytes and brain CD45<sup>+</sup> cells was assessed by flow cytometry in the WT, 5xFAD, and *Il3<sup>-/-</sup>*

$-5xFAD$  parabionts ( $n = 4$  WT mice;  $n = 5$   $5xFAD$  mice;  $n = 3$   $Il3^{-/-}5xFAD$  mice). Groups of mice are of evenly mixed sex. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . Mean  $\pm$  s.e.m.

### Extended Data Fig. 3 Generation and validation of $Il3^{GFPfl/fl} Aldh1l1Cre^{ERT2} 5xFAD$ and $Il3ra^{fl/fl} Cx3cr1Cre^{ERT2} 5xFAD$ mice.

**a, b**, Schematics of the endogenous loci and editing strategies to generate conditional/reporter models for  $Il3$  (a) and  $Il3ra$  (b). Mice were generated by excising large fragments of the endogenous loci by co-delivery of SpCas9 and two gRNAs, and in the presence of a long single-stranded DNA donor encoding a loxP-cDNA-P2A-EGFP-loxP cassette. **c, d**, Representative Sanger sequencing traces validating insertion of the loxP-cDNA-P2A-EGFP-loxP cassettes at the endogenous loci of  $Il3$  (c) and  $Il3ra$  (d). Missense mutation quenches the GFP signal in  $Il3ra$ -targeted mice. **e**, GFP signal in ex vivo-stimulated splenic and lymph node T cells, known IL-3 sources, from WT and  $Il3^{GFPfl/fl}$  mice. **f**, Flow cytometry analysis of astrocyte IL-3 production in 5-month-old  $Il3^{GFPfl/fl} 5xFAD$  and  $Il3^{GFPfl/fl} Aldh1l1Cre^{ERT2} 5xFAD$  mice injected with tamoxifen. **g**, qPCR analysis of  $Il3$  mRNA expression in sorted astrocytes ( $n = 4$   $Il3^{GFPfl/fl} 5xFAD$  mice;  $n = 5$   $Il3^{GFPfl/fl} Aldh1l1Cre^{ERT2} 5xFAD$  mice). **h**, CSF IL-3 levels. Filled circles, males; open circles, females ( $n = 6$   $Il3^{GFPfl/fl} 5xFAD$  mice;  $n = 9$   $Il3^{GFPfl/fl} Aldh1l1Cre^{ERT2} 5xFAD$  mice). **i**, STAT5 phosphorylation in ex vivo heart macrophages from  $Il3ra^{fl/fl}$  mice simulated with rIL-3. **j**, Flow cytometry analysis of microglia IL-3Ra production in 5-month-old  $Il3ra^{fl/fl} 5xFAD$  and  $Il3ra^{fl/fl} Cx3cr1Cre^{ERT2} 5xFAD$  mice injected with tamoxifen. **k**, qPCR analysis of  $Il3ra$  mRNA expression in sorted microglia ( $n = 4$   $Il3ra^{fl/fl} 5xFAD$  mice;  $n = 5$   $Il3ra^{fl/fl} Cx3cr1Cre^{ERT2} 5xFAD$  mice). All  $Il3^{GFPfl/fl} 5xFAD$ ,  $Il3^{GFPfl/fl} Aldh1l1Cre^{ERT2} 5xFAD$ ,  $Il3ra^{fl/fl} 5xFAD$ , and  $Il3ra^{fl/fl} Cx3cr1Cre^{ERT2} 5xFAD$  mice were injected with tamoxifen beginning at 2 months of age. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Mann–Whitney U-tests. Mean  $\pm$  s.e.m.

## Extended Data Fig. 4 Flow cytometry gating strategy and histology controls.

**a**, Gating strategy used to identify cell populations in the brains of all mice except *Aldh1l1*<sup>GFP</sup> mice. **b**, Backgating of GFP<sup>+</sup> astrocytes in *Il3*<sup>GFPfl/fl</sup> mice. **c**, Gating strategy used to identify cell populations in the brain of *Aldh1l1*<sup>GFP</sup> mice. **d**, Gating and isotype control plots for microglia IL-3R $\alpha$  staining. **e**, Representative images of IgG control antibody staining and IL-3 staining in the brain.

## Extended Data Fig. 5 Astrocyte IL-3 and microglia IL-3R $\alpha$ production.

**a**, *Il3* expression in astrocytes sorted from WT and 5xFAD mice at 5 and 12 months of age ( $n = 5$  5-month-old WT mice;  $n = 4$  5-month-old 5xFAD mice;  $n = 3$  12-month-old WT mice;  $n = 4$  12-month-old 5xFAD mice). **b**, Tissue IL-3 levels in various brain regions in WT mice at 4 months of age ( $n = 4$ ). **c**, Scheme of daily i.p. LPS injection into WT mice over 4 days and qPCR analysis of *C3* and *Gfap* in sorted astrocytes after LPS injection ( $n = 4$  PBS mice;  $n = 3$  LPS mice for *C3*;  $n = 4$  LPS mice for *Gfap* and *Il3*). Two-tailed Mann–Whitney *U*-tests. **d**, CSF IL-3 levels ( $n = 8$  WT PBS mice;  $n = 6$  LPS mice). **e**, Representative images of GFAP<sup>+</sup> astrocytes in the cortex of WT, *Il3*<sup>-/-</sup>, 5xFAD, and *Il3*<sup>-/-</sup> 5xFAD mice. **f**, Representative images of A $\beta$  deposits (6E10) and astrocytes (GFAP) in 5xFAD and *Il3*<sup>-/-</sup> 5xFAD mice. **g**, Proportion of IL-3R $\alpha$ <sup>+</sup> microglia in the brain of WT mice at various ages ( $n = 4$ ). **h**, *Il3ra* transcript expression in brain homogenate from WT mice at various ages ( $n = 4$ ). **i**, Proportion of IL-3R $\alpha$ <sup>+</sup> macrophages in heart, liver, lung (interstitial and alveolar), and brain of WT and 5xFAD mice at 8 months of age ( $n = 3$  WT and 5xFAD mice for all tissues except brain where  $n = 4$ ). One-way ANOVA. \* $P < 0.05$ , \*\* $P < 0.001$ . Mean  $\pm$  s.e.m.

## Extended Data Fig. 6 *Il3ra* expression in microglia and characterization of IL-3R $\alpha$ <sup>hi</sup> and IL-3R $\alpha$ <sup>lo</sup> microglia.

**a**, Analysis of *Il3ra* and other cytokine receptors expressed in homeostatic, DAM stage 1, and DAM stage 2 microglia (data from ref. [23](#)). **b**, Gating strategy for IL-3R $\alpha^{\text{hi}}$  and IL-3R $\alpha^{\text{lo}}$  microglia from 5xFAD mice. **c**, Flow cytometry analysis of IL-3R $\alpha^{\text{hi}}$  and IL-3R $\alpha^{\text{lo}}$  microglia ( $n = 6$  except CD11c and CCL2 where  $n = 3$ ). **d**, mRNA transcript expression in sorted IL-3R $\alpha^{\text{hi}}$  and IL-3R $\alpha^{\text{lo}}$  microglia ( $n = 7$  for IL-3R $\alpha^{\text{lo}}$  microglia except that  $n = 6$  for *Il6*, *Itgam*, *Ptprc*, *Cd36*, and *Cd68*, and  $n = 5$  for *Il10*;  $n = 7$  for IL-3R $\alpha^{\text{hi}}$  microglia except that  $n = 6$  for *Il6*, *Ptprc* and *Cd68*, and  $n = 5$  for *Il10*). \* $P < 0.05$ , \*\* $P < 0.01$ , two-tailed Mann–Whitney *U*-tests. Mean ± s.e.m.

### **Extended Data Fig. 7 IL-3 does not influence proliferation, recognition and phagocytosis of A $\beta$ , or production of inflammatory cytokines in microglia.**

**a**, BrdU incorporation into microglia ( $n = 6$  5xFAD;  $n = 7$  *Il3* $^{-/-}$  5xFAD). Two-tailed Mann–Whitney *U*-tests. **b**, Heat map of microglia RNA-seq expression data for genes important for cell cycle and proliferation. **c**, Assessment of ex vivo phagocytosis of A $\beta$ 42 conjugated to a pH-sensitive dye (PHrodo Red) in sorted microglia lacking *Il3* or stimulated with rIL-3 ( $n = 4$ ). **d**, Heat map of microglia RNA-seq expression data for genes important for A $\beta$  recognition and phagocytosis. **e**, Heat map of microglia RNA-seq expression data for inflammatory cytokines. Groups of mice are of evenly mixed sex. Mean ± s.e.m.; ns, not significant.

### **Extended Data Fig. 8 Human AD iPS triculture system.**

**a**, Representative images demonstrating p-tau (PHF1) localization with neurons (TUJ1) and the presence of astrocytes (GFAP) in the central chamber, and microglia (P2RY12) in the side chamber. **b**, ELISA quantification of A $\beta$ 40, A $\beta$ 38, and A $\beta$ 42 in medium ( $n = 3$  per group). **c**, *IL3RA* expression in human iPS microglia exposed to A $\beta$  ( $n = 2$ ). **d**, Chemokine and cytokine levels in the medium of the human AD iPS triculture system ( $n = 3$  except  $n = 2$  for CCL13 AD astrocytes + neurons). Mean ± s.e.m.

## Extended Data Fig. 9 rIL-3 delivery to the cortex or periphery of 5xFAD mice and summary figure.

**a**, Recombinant IL-3 or PBS was delivered into the cortex of 5xFAD mice. Three days later, localization of microglia to A $\beta$  aggregates was assessed ( $n = 7$  PBS mice;  $n = 6$  rIL-3 mice). Two-tailed Mann–Whitney  $U$ -tests. **b**, Scheme of rIL-3 delivery intraperitoneally twice a week for 10 weeks to 5xFAD mice. **c**, Prior to death, Y-maze behavioural testing was performed and time in the new arm was quantified ( $n = 7$  PBS mice;  $n = 8$  rIL-3 mice). **d**, The amount of A $\beta$  in the cortex of mice was quantified by analysing histological sections ( $n = 6$ ). Groups of mice are of evenly mixed sex. \*\* $P < 0.01$ . Mean  $\pm$  s.e.m. **e**, Model of the role of IL-3 in AD. Astrocytes produce IL-3. In response to A $\beta$ , TREM2 signalling increases microglia IL-3R $\alpha$ , rendering microglia responsive to astrocyte-derived IL-3. IL-3 signalling instigates transcriptional and functional programming of microglia, which leads to a signature of immune regulation, motility, and migration. IL-3-dependent programming promotes clustering of microglia around A $\beta$ , which enables clearance of A $\beta$  and mitigation of AD pathology.

### **Extended Data Table 1 Characteristics of control donors and patients with AD**

[Full size table](#)

## **Supplementary information**

### Supplementary Information

This file contains Supplementary Tables 1-5 and Supplementary Sequences 1-2.

### Reporting Summary

### Video 1

Three-dimensional confocal imaging of optically-cleared mouse cortex stained for A $\beta$  and microglia.

# Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Fig. 5](#)

# Rights and permissions

[Reprints and Permissions](#)

# About this article



Check for  
updates

# Cite this article

McAlpine, C.S., Park, J., Griciuc, A. *et al.* Astrocytic interleukin-3 programs microglia and limits Alzheimer's disease. *Nature* **595**, 701–706 (2021). <https://doi.org/10.1038/s41586-021-03734-6>

[Download citation](#)

- Received: 29 December 2020
- Accepted: 17 June 2021
- Published: 14 July 2021

- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03734-6>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

## Associated Content

### [A protective signal between the brain's supporting cells in Alzheimer's disease](#)

- Jerika J. Barron
- Anna V. Molofsky

### [Astrocytic IL-3 could help microglia protect against Alzheimer disease](#)

- Heather Wood

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03734-6>

---

Spread of a SARS-CoV-2 variant through Europe in the summer of 2020

[Download PDF](#)

- Article
- [Published: 07 June 2021](#)

# Spread of a SARS-CoV-2 variant through Europe in the summer of 2020

- [Emma B. Hodcroft](#) [ORCID: orcid.org/0000-0002-0078-2212](#)<sup>1,2,3</sup>,
- [Moira Zuber](#) [ORCID: orcid.org/0000-0002-4275-8739](#)<sup>1</sup>,
- [Sarah Nadeau](#) [ORCID: orcid.org/0000-0003-1008-8918](#)<sup>2,4</sup>,
- [Timothy G. Vaughan](#) [ORCID: orcid.org/0000-0001-6220-2239](#)<sup>2,4</sup>,
- [Katharine H. D. Crawford](#) [ORCID: orcid.org/0000-0002-6223-4019](#)<sup>5,6,7</sup>,
- [Christian L. Althaus](#) [ORCID: orcid.org/0000-0002-5230-6760](#)<sup>3</sup>,
- [Martina L. Reichmuth](#) [ORCID: orcid.org/0000-0001-9345-851X](#)<sup>3</sup>,
- [John E. Bowen](#) [ORCID: orcid.org/0000-0003-3590-9727](#)<sup>8</sup>,
- [Alexandra C. Walls](#) [ORCID: orcid.org/0000-0002-9636-8330](#)<sup>8</sup>,
- [Davide Corti](#) [ORCID: orcid.org/0000-0002-5797-1364](#)<sup>9</sup>,
- [Jesse D. Bloom](#) [ORCID: orcid.org/0000-0003-1267-3408](#)<sup>5,6,10</sup>,
- [David Veesler](#) [ORCID: orcid.org/0000-0002-6019-8675](#)<sup>8</sup>,
- [David Mateo](#) [ORCID: orcid.org/0000-0002-1590-4163](#)<sup>11</sup>,
- [Alberto Hernando](#) [ORCID: orcid.org/0000-0003-1180-1068](#)<sup>11</sup>,
- [Iñaki Comas](#) [ORCID: orcid.org/0000-0001-5504-9408](#)<sup>12,13</sup>,
- [Fernando González-Candelas](#) [ORCID: orcid.org/0000-0002-0879-5798](#)<sup>13,14</sup>,
- [SeqCOVID-SPAIN consortium](#),
- [Tanja Stadler](#) [ORCID: orcid.org/0000-0001-6431-535X](#)<sup>2,4 nal</sup> &
- [Richard A. Neher](#) [ORCID: orcid.org/0000-0003-2525-1407](#)<sup>1,2 nal</sup>

*Nature* volume **595**, pages 707–712 (2021) [Cite this article](#)

- 24k Accesses
- 829 Altmetric
- [Metrics details](#)

## Subjects

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

## Abstract

Following its emergence in late 2019, the spread of SARS-CoV-2<sup>1,2</sup> has been tracked by phylogenetic analysis of viral genome sequences in unprecedented detail<sup>3,4,5</sup>. Although the virus spread globally in early 2020 before borders closed, intercontinental travel has since been greatly reduced. However, travel within Europe resumed in the summer of 2020. Here we report on a SARS-CoV-2 variant, 20E (EU1), that was identified in Spain in early summer 2020 and subsequently spread across Europe. We find no evidence that this variant has increased transmissibility, but instead demonstrate how rising incidence in Spain, resumption of travel, and lack of effective screening and containment may explain the variant's success. Despite travel restrictions, we estimate that 20E (EU1) was introduced hundreds of times to European countries by summertime travellers, which is likely to have undermined local efforts to minimize infection with SARS-CoV-2. Our results illustrate how a variant can rapidly become dominant even in the absence of a substantial transmission advantage in favourable epidemiological settings. Genomic surveillance is critical for understanding how travel can affect transmission of SARS-CoV-2, and thus for informing future containment strategies as travel resumes.

[Download PDF](#)

## Main

The pandemic of COVID-19, which is caused by SARS-CoV-2, is the first pandemic for which the spread of a viral pathogen has been globally tracked in near real-time using phylogenetic analysis of viral genome sequences<sup>3,4,5</sup>. SARS-CoV-2 genomes continue to be generated at a rate far greater than for any other pathogen, and more than 950,000 full genomes were available in the Global Initiative on Sharing All Influenza Data (GISAID) database as of April 2021<sup>6</sup>.

In addition to tracking viral spread, these sequences have been used to monitor mutations that might change the transmission, pathogenesis, or antigenic properties of the virus. One mutation in particular, D614G in the spike protein (Nextstrain clade 20A and its descendants), seeded large outbreaks in Europe in early 2020 and subsequently dominated outbreaks in the Americas, thereby largely replacing

previously circulating lineages. This rapid rise led to the suggestion that this variant is more transmissible, which has since been corroborated by phylogenetic<sup>7,8</sup> and experimental evidence<sup>9,10</sup>. Subsequently, three variants of concern—Alpha/501Y.V1/B.1.1.7 (refs. <sup>11,12</sup>), Beta/501Y.V2/B.1.351 (refs. <sup>13,14</sup>) and Gamma/501Y.V3/P.1 (ref. <sup>15</sup>), which have increased transmissibility and/or can partially escape neutralization—were identified at the end of 2020.

Following the global dissemination of SARS-CoV-2 in early 2020<sup>3</sup>, intercontinental travel dropped markedly. Within Europe, however, travel (particularly holiday travel) resumed in the summer. Here we report on SARS-CoV-2 variant 20E (EU1), with an A222V mutation in the spike protein, which first rose in frequency in Spain in early summer 2020 and subsequently spread to multiple locations in Europe, rising in frequency in parallel. This variant and a second variant (20A.EU2, with an S477N mutation in the spike protein) accounted for the majority of sequences in Europe in the autumn of 2020.

## European variants in summer 2020

Figure 1 shows a time-scaled phylogeny of sequences sampled in Europe up to the end of November 2020 and their global context, highlighting the variants discussed here. A cluster of sequences in clade 20A has an additional mutation (spike A222V), shown in orange. We designate this cluster as 20E (EU1) (this cluster consists of lineage B.1.177 and its sublineages<sup>16</sup>).

**Fig. 1: Phylogenetic overview of SARS-CoV-2 in Europe up to the end of November 2020.**

 figure1

Left, the tree shows a representative sample of isolates from Europe coloured by clade and by the variants highlighted in this paper. Clade 20A and its daughter clades 20B and 20C (yellow) carry mutations S:D614G. Variant 20E (EU1) (orange), with mutation S:A222V on a S:D614G background, emerged in early summer 2020 and

became common in many European countries in autumn 2020. A separate variant (20A.EU2; blue) with mutation S:S477N became prevalent in France. Right, the proportion of sequences belonging to each variant (up to the end of November 2020) per country. Tree and visualization were generated using the Nextstrain platform<sup>4</sup> (see [Methods](#)). Map data copyright Google, INEGI (2021).

[Full size image](#)

In addition to 20E (EU1), a variant (20A.EU2; blue in Fig. 1) with several amino acid substitutions, including S:S477N, became common in some European countries, particularly France (Extended Data Fig. 1). The S:S477N substitution has arisen multiple times independently, for example in clade 20F, which dominated the outbreak in Oceania during the southern-hemisphere winter. Residue S477 is close to the receptor binding site (Extended Data Fig. 2) and forms part of the epitope recognized by the S2E12 and C102 neutralizing antibodies<sup>17,18</sup>.

Several other smaller clusters, defined by the spike mutations D80Y, S98F and N439K, have also been found in multiple countries (Extended Data Table 1, Extended Data Fig. 1). Whereas none of these has reached the prevalence of 20E (EU1) or 20A.EU2, some have attracted attention in their own right: S:N439K is present in two larger clusters found across Europe<sup>19</sup> and arose several times independently. Updated phylogenies and further analyses of these and other variants are available at <https://covariants.org/>.

## Characterization of S:A222V

Our analysis here focuses on the variant 20E (EU1), with substitution A222V in the spike protein's domain A (also referred to as the N-terminal domain (NTD))<sup>18,20,21</sup> (Extended Data Fig. 2). This mutation is not known to influence receptor binding or membrane fusion by SARS-CoV-2. However, mutations can sometimes have long-range effects on protein conformation or stability.

To investigate whether the A222V mutation affects the conformation of the SARS-CoV-2 spike glycoprotein, we used enzyme-linked immunosorbent assay (ELISA) to test binding of the mutant ectodomain by the benchmark COVID-19 convalescent patient plasma from the National Institute for Biological Standards and Control, and by neutralizing monoclonal antibodies that recognize the receptor-binding domain (RBD; antibodies S2E12 and S309)<sup>18,22,23</sup> and the NTD (antibody 4A8)<sup>24</sup>. The dose-response curves were indistinguishable for the extodomain trimers of SARS-CoV-2 2PS (a prefusion-stabilized form of the spike protein with two proline substitutions) and SARS-CoV-2 2P A222V D614G S (Extended Data Fig. 3a–d), aligning with results from a recent study<sup>25</sup>. Collectively, these data indicate that the A222V

substitution does not appreciably affect the antigenicity of the SARS-CoV-2 spike protein.

To test whether the A222V mutation had an obvious functional effect on the ability of the spike protein to mediate viral entry, we produced lentiviral particles pseudotyped with spike either containing or lacking the A222V mutation on the background of the D614G mutation and deletion of the end of its cytoplasmic tail. Lentiviral particles with the A222V mutant spike had slightly higher titres than those without (mean 1.3-fold higher), although the difference was not statistically significant after normalization by p24 concentration (Extended Data Fig. 3e–h). Therefore, A222V does not lead to the same large increases in the titres of spike-pseudotyped lentivirus as was observed for the D614G mutation<sup>7,10</sup>. However, this small effect must be interpreted cautiously, as the effects of mutations on viral transmission in humans are not always paralleled by measurements made in simplified experimental systems.

In addition to S:A222V, 20E (EU1) has the amino acid mutations ORF10:V30L, N:A220V and ORF14:L67F. However, there is little evidence for the functional relevance of ORF10 and ORF14<sup>26,27</sup>. Different mutations between positions 180 and 220 in the nucleocapsid (N) protein are observed in almost every major lineage of SARS-CoV-2 and we are not aware of any evidence that these mutations have important phenotypic consequences. Therefore, we examined epidemiological and phylogenetic evidence to explain the spread of 20E (EU1).

## Early observations of 20E (EU1)

The earliest sequences of 20E (EU1) were found in samples collected on 20 June 2020 (seven in Spain and one in the Netherlands). By the end of August, 20E (EU1) sequences had also been detected in Belgium, Switzerland, France, Denmark, the UK, Germany, Latvia, Sweden, Norway and Italy. Sequences of 20E (EU1) from Hong Kong, Australia, New Zealand, and Singapore—presumably representing exports from Europe—were first detected between mid-August and mid-October (Supplementary Table 1).

The proportion of sequences that matched 20E (EU1) in several countries is plotted by ISO week in Fig. 2. This variant first rose in frequency in Spain, reaching around 50% prevalence within a month of the first sequence being detected before rising to 80%. In many European countries, there was a gradual rise starting in mid-July before prevalence settled at a level between 15 and 80% in September or October.

**Fig. 2: Frequency of submitted samples representing 20E (EU1) in selected countries.**



We include the eight countries for which there are at least 200 sequences from 20E (EU1), as well as Norway and France, to illustrate points in the text. The symbol size indicates the number of available sequences by country and time point in a nonlinear manner. In most countries we observe a gradual rise from mid-July that settles to a plateau. By contrast, Norway experienced a sharp peak in summer but seems to have brought cases down quickly, although they began to increase again in September. When the last data point included only very few sequences, it has been dropped for clarity. Frequencies are smoothed using a Gaussian with  $\sigma = 1$  week.

[Full size image](#)

## Expansion and spread across Europe

To quantify the spread of 20E (EU1) across Europe, we constructed a phylogeny (Extended Data Fig. 4a) based on data from samples collected before 30 September 2020 and available from GISAID in January 2021 (see [Methods](#)). The phylogeny is collapsed to group diversity that might stem from within-country transmission into sectors of the pie charts (Extended Data Fig. 4b–d) for selected countries. The tree indicates that 20E (EU1) harbours substantial diversity, and most major genotypes have been observed in many European countries. As it is unlikely that phylogenetic patterns sampled in multiple countries arose independently, it is reasonable to assume that most mutations observed in the tree arose once and were carried (possibly multiple times) between countries. Throughout July and August 2020, Spain had a higher per capita incidence of SARS-CoV-2 infection than most other European countries (Extended Data Fig. 5) and 20E (EU1) was much more prevalent in Spain

than elsewhere, suggesting that Spain is the likely origin of most introductions of 20E (EU1) into other countries.

Epidemiological data from Spain indicate that the earliest sequences in the cluster are associated with two known outbreaks in the northeast of the country. The variant seems to have initially spread among agricultural workers in Aragon and Catalonia, then moved into the local population, where it was able to travel to the Valencia region and on to the rest of the country.

Most basal genotypes have been observed both in Spain and in many other countries, suggesting that they were repeatedly exported. However, the 795 sequences from Spain included in Extended Data Fig. 4a are not likely to represent the full diversity. Variants found only outside Spain may reflect diversity that arose in other countries, or may represent diversity that was present but not sampled in Spain (particularly as some European countries, such as the UK and Denmark, sequence a high proportion of cases). Despite limitations in sampling, Extended Data Fig. 4a clearly shows that most major genotypes in this cluster were distributed to multiple countries, suggesting that identical genotypes were introduced into many countries. This is consistent with the large number of introductions estimated from travel data, discussed below. Although initial introductions of the variant probably originated from Spain, cases of 20E (EU1) outside Spain surpassed those in Spain in late September, and later cross-border transmissions are likely to have originated in other countries (Extended Data Fig. 5b). The [Supplementary Information](#) includes a discussion of travel restrictions in selected European countries and the associated patterns of 20E (EU1) introductions.

Extended Data Figure 4e shows the distribution of sequence clusters compatible with onward transmission within countries outside Spain, and highlights two different patterns. Norway and Iceland, for example, seem to have had only a small number of introductions over the summer that led to substantial further spread. In Extended Data Fig. 4a, the majority of sequences from these countries fall into one sector and the remainder are singletons or very small clusters that did not spread. However, later sequences in Norway or Iceland often cluster more closely with diversity in non-Spanish European countries, which may suggest that further introductions came from third countries (see 20E (EU1) Nextstrain build online: <https://covariants.org/variants/20A.EU1>).

By contrast, countries such as Switzerland, the Netherlands, and the UK have sampled sequences that correspond to a large number of independent introductions and include most major genotypes observed in Spain.

## No evidence for transmission advantage

During a dynamic outbreak, it is particularly difficult to be certain of whether a particular variant is increasing in frequency because it has an intrinsic advantage, or because of epidemiological factors<sup>28</sup>. In fact, it is a tautology that every novel large cluster must have grown recently, and multiple lines of independent evidence are required to demonstrate intrinsically elevated transmission potential.

Initially, 20E (EU1) was dispersed across Europe mainly by travellers to and from Spain. Many EU and Schengen-area countries opened their borders to other countries in the bloc on 15 June 2020. Travel resumed quickly and peaked during July and August (Fig. 3). The number of confirmed cases of SARS-CoV-2 infection in Spain rose from around 10 cases per 100,000 inhabitants per week in early July to 100 cases per 100,000 inhabitants per week in late August, while case numbers remained low in most of Europe during this time. To investigate whether repeated imports are sufficient to explain the rapid rise in frequency of 20E (EU1) and its displacement of other variants, we first estimated the number of expected introductions of 20E (EU1) on the basis of the number of visitors from a particular country to different provinces of Spain and the incidence of SARS-CoV-2 infection in those provinces. Taking reported incidence in the provinces at face value and assuming that returning tourists have a similar incidence, we expect 380 introductions of 20E (EU1) into the UK over the summer (6 July to 27 September; see Supplementary Table 2 and Fig. 3 for tourism summaries<sup>29</sup> and departure statistics<sup>30</sup>). Similarly, for Germany and Switzerland we would expect around 320 and 90 introductions of 20E (EU1), respectively. We then created a simple model that also incorporates the incidence in the country to which travellers are returning and the onward spread of imported infections with 20E (EU1) to estimate the frequency of 20E (EU1) in countries across Europe over time (Fig. 3). This model assumes that 20E (EU1) spread at the same rate as other variants in the resident countries and predicts that the frequencies of 20E (EU1) would start to rise in July, continue to rise during August, and be stable thereafter, consistent with observations in many countries (Fig. 3b).

**Fig. 3: Travel volume and contribution of imported infections.**

---

 **figure3**

---

**a**, Travel from Spain to other European countries resumed in July 2020 (although it was low compared to previous years). Assuming that travel returnees were infected at the average incidence of the Spanish province they visited and transmitted the virus at the rate of their resident population, imports from Spain are expected to account for between 2 and 12% of SARS-CoV-2 infections after the summer. Traveller incidence was calculated using case and travel data at the level of provinces. Note that this

model accounts only for the contribution of summer travel and that stochastic fluctuations and other variants after the summer will result in further variation in the frequency of 20E (EU1). See [Methods](#) and Extended Data Fig. [8](#) for geographic detail.

[Full size image](#)

Although the shape of the expected frequency trajectories from imports in Fig. [3b](#) is consistent with observations, this naive import model underestimates the final observed frequency of 20E (EU1) by between 1- and 12-fold, depending on the country (Extended Data Fig. [6](#)). This discrepancy might be due to either intrinsically faster transmission of 20E (EU1) or underestimation of introductions. Underestimates might result from country-specific reporting, such as the relative ascertainment rate in source and destination populations, and the fact that risk of exposure and onward transmission are likely to be increased by travel-related activities both abroad, en route, and at home. Furthermore, the incidence of SARS-CoV-2 infection in holiday destinations might not be well-represented by the provincial averages used in the model. For example, during the first wave in spring 2020, some ski resorts had exceptionally high incidence and contributed disproportionately to the dispersal of SARS-CoV-2<sup>[31,32](#)</sup>. The facts that the rapid increase in the frequency of 20E (EU1) slowed or stopped in most countries after the summer travel period and that it did not fully replace other variants are consistent with import-driven dynamics with little or no competitive advantage.

The notion that an underestimated incidence in travel returnees rather than faster spread of 20E (EU1) is the major contributor to the above discrepancy is supported by the fact that German authorities reported about 2.2 times as many cases with suspected infection in Spain as the model predicts (982 reported versus 452 estimated from 6 July to 13 September, regardless of variant; Extended Data Fig. [7a](#)). Switzerland reported 131 infections in travel returnees, while the model predicts 130. After we adjusted imports for the 37% of Swiss case reports without exposure information, we find that the model underestimates introductions 1.6-fold. Tourists from countries with small (1–4-fold) and large (8–12-fold) discrepancies tended to visit different destinations in Spain (Extended Data Figs. [6](#), [7c–e](#)), which further suggests that the underestimation of incidence in travel returnees is determined by destination and behaviour.

To investigate the possibility that 20E (EU1) grew faster than other variants following its introduction, we identified introductions of both 20E (EU1) and non-20E (EU1) variants into Switzerland and their downstream Swiss transmission chains. These data suggest that there were 34 or 291 introductions of 20E (EU1), depending on the criterion used to assign sequences to putative transmission chains (see [Methods](#)). Phylodynamic estimates of the effective reproductive number ( $R_e$ ) over time for introductions of 20E (EU1) and for other variants (Extended Data Fig. [8](#)) suggest a

tendency for 20E (EU1) introductions to transiently grow faster. This transient signal of faster growth, however, is more readily explained by behavioural differences and increased travel-associated transmission than intrinsic differences in the virus. We repeated the phylogenetic analysis with a pan-European set of putative introductions and found similar patterns to those observed for Switzerland.

These patterns are also consistent with the fact that Swiss cases with likely exposure in Spain tended to be in younger individuals (median 30 years, interquartile range (IQR) 23–42.25 years) than cases acquired in Switzerland (median 35 years, IQR 24–51 years). Younger individuals tend to have more contacts than older people<sup>33,34</sup>. The association with particular demographics will decay rapidly, and with it any associated increased transmission inferred by phylogenetics.

Most introductions of 20E (EU1) are expected to have occurred towards the end of summer, when incidence in Spain was rising and return travel volume peaked. For countries such as Belgium, the comparatively high incidence of non-20E (EU1) variants at this time and, therefore, a relatively low impact of imported variants (Extended Data Fig. 5) might explain why 20E (EU1) remained at low frequencies in these countries despite high-volume travel to Spain.

Case numbers across Europe started to rise rapidly around the same time as the 20E (EU1) variant started to become prevalent in multiple countries (Extended Data Fig. 5). However, countries where 20E (EU1) was rare (Belgium, France, Czech Republic; Extended Data Fig. 1) saw similarly rapid increases, which suggests that this rise was not driven by any particular lineage and that 20E (EU1) has no substantial difference in transmissibility. Furthermore, in Switzerland  $R_e$  increased in the autumn by a comparable amount for the 20E (EU1) and non-20E (EU1) variants (Extended Data Fig. 8). Although we cannot rule out the possibility that 20E (EU1) had a slight transmission advantage compared to other variants circulating at the time, most of its spread is explained by epidemiological factors. The arrival of autumn and seasonal factors are a more plausible explanation for the resurgence of cases<sup>35</sup>.

## Discussion

The rapid spread of 20E (EU1) and other variants have underscored the importance of a coordinated and systematic sequencing effort to detect, track, and analyse emerging SARS-CoV-2 variants. This becomes even more urgent with the recent detection of several variants of concern<sup>11,12,13,14,15</sup>. It is only through multi-country genomic surveillance that it has been possible to detect and track 20E (EU1) and other variants.

When a new variant is observed, policy-makers need a rapid assessment of whether the new variant increases the transmissibility of the virus, evades pre-existing

immunity or has different clinical properties<sup>36</sup>. In the case of 20E (EU1), none of these seem to have changed substantially, making it an important example of how travel combined with large regional differences in prevalence can lead to substantial and rapid shifts in the variant distribution without a marked transmission advantage. Such shifts—driven predominantly by epidemiological factors—are more likely in a low-incidence setting, where a large fraction of cases can result from introductions. By contrast, the variant of concern Alpha/501Y.V1/B.1.1.7 spread across Europe in late 2020, at which time most countries, including the UK (where it first rose to prominence), reported a high incidence. In such high-incidence settings, travel alone cannot explain a rapid rise in frequency, and the dynamics point to a bona fide transmission advantage. In-depth characterization of a spectrum of such dynamics (no substantial advantage for 20E (EU1), moderate advantage for the D614G mutation<sup>8</sup>, and a strong transmission advantage for Alpha/501Y.V1/B.1.1.7<sup>11,12</sup> and Beta/501Y.V2<sup>13</sup>) will facilitate the assessment of emerging variants in the future.

Finally, our analysis highlights that countries should carefully consider their approach to travel when large-scale international movement resumes across Europe. We show that holiday travel in summer 2020 resulted in unexpectedly high levels of introductions and onward spread across Europe. Whether the 20E (EU1) variant described here spread rapidly owing to a transmission advantage or to epidemiological factors alone, its repeated introduction and rise in prevalence in multiple countries implies that the summer travel guidelines and restrictions were generally not sufficient to prevent onward transmission of introductions. Travel precautions such as quarantine should, in principle, have prevented the spread of SARS-CoV-2 infections acquired abroad, but in practice failed to have the desired effect. Although long-term travel restrictions and border closures are neither tenable nor desirable, the identification of better ways to reduce the risk of introducing variants, and to ensure that those that are introduced do not spread widely, will help countries to maintain their (often hard-won) low levels of SARS-CoV-2 transmission.

## 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.

### Phylogenetic analysis

We used the Nextstrain pipeline for our phylogenetic analyses (<https://github.com/nextstrain/ncov/>)<sup>4</sup>. In brief, we aligned sequences using mafft<sup>37</sup>, subsampled sequences (see below), added sequences from the rest of the world for phylogenetic context based on genomic proximity, reconstructed a phylogeny using IQTree<sup>38</sup> and inferred a time-scaled phylogeny using TreeTime<sup>39</sup>. For computational feasibility, ease of interpretation, and to balance disparate sampling efforts between countries, the Nextstrain-maintained runs sub-sample the available genomes across time and geography, resulting in final builds of ~5,000 genomes each. After sub-sampling, the 20E (EU1) cluster within the Nextstrain build contained 5,145 sequences, 3,369 of which were unique (accounting for missing data in the sequence).

Sequences were downloaded from GISAID at the end of January and analysed using the nextstrain/ncov workflow, using a cutoff date of 30 September 2020 (for Extended Data Fig. 4a) or 30 November 2020 (for all other analyses). These dates were chosen to focus first on introductions over the summer (for 30 September) and then to highlight ongoing circulation through the autumn (30 November) before the spread of the variants of concern identified in December 2020 and January 2021. A table acknowledging the invaluable contributions by many labs is available in the [Supplementary Information](#). The Swiss SARS-CoV-2 sequencing efforts have been described previously<sup>40,41</sup>. The majority of Swiss sequences used here are from the Nadeau et al.<sup>40</sup> dataset and the remainder are available on GISAID.

## Defining the 20E (EU1) cluster

The cluster was initially identified as a monophyletic group of sequences stemming from the larger 20A clade with amino acid substitutions at positions S:A222V, ORF10:V30L, and N:A220V or ORF14:L67F (overlapping reading frame with N), corresponding to nucleotide mutations C22227T, C28932T, and G29645T. In addition, sequences in 20E (EU1) differ from their ancestors by the synonymous mutations T445C, C6286T, and C26801G.

The sub-sampling of the standard Nextstrain analysis means that we were not able to visualize the true size or phylogenetic structure of the cluster in question. To specifically analyse this cluster using almost all available sequences, we designed a specialized build that focused on cluster-associated sequences and their most genetically similar neighbours. For computational reasons, we limited the number of samples to 900 per country per month. As only the UK has more sequences than this for the relevant time period, this results in a random downsampling of sequences from the UK for the months of August, September, and October. Furthermore, we excluded several problematic sequences because of high intra-sample variation, wrong dates, and over-divergence (divergence values were implausible given the provided dates). A

full list of the sequences excluded (and the reasons why) is given on GitHub at [https://github.com/neherlab/2020\\_EU1\\_paper/blob/master/scripts/bad\\_sequences.py](https://github.com/neherlab/2020_EU1_paper/blob/master/scripts/bad_sequences.py).

We identified sequences in the cluster on the basis of the presence of nucleotide substitutions at positions 22227, 28932, and 29645 and used this set as a ‘focal’ sample in the nextstrain/ncov pipeline. This selection excludes any sequences with no coverage or reversions at these positions, but the similarity-based sampling during the Nextstrain run will identify these, as well as any other nearby sequences, and incorporate them into the dataset. We used these three mutations as they included the largest number of sequences that are distinct to the cluster. By this criterion, there are currently 60,316 sequences in the cluster that were sampled before 30 November 2020.

To visualize the changing prevalence of the cluster over time, we plotted the proportion of sequences identified by the four substitutions described above as a fraction of the total number of sequences submitted, per ISO week. The frequencies of other clusters were identified in an analogous way.

## Phylogeny and geographic distribution

The size of the cluster and the number of unique mutations among individual sequences mean that interpreting overall patterns and connections between countries is not straightforward. We aimed to create a simplified version of the tree that focuses on connections between countries and de-emphasizes onward transmission within a country. As our focal build contained ‘background’ sequences that do not fall within the cluster, we used only the monophyletic clade containing the four amino-acid changes and three synonymous nucleotide changes that identify the cluster. Then, subtrees that contained only sequences from one country were collapsed into the parent node. The resulting phylogeny contains only mixed-country nodes and single-country nodes that have mixed-country nodes as children. (An illustrative example of this collapsing can be seen in Extended Data Fig. 4b–d.) Nodes in this tree therefore represent ancestral genotypes of subtrees: sequences represented within a node may have further diversified within their country, but share a set of common mutations. We count all sequences in the subtrees towards the geographic distribution represented in the pie charts in Extended Data Fig. 4a.

This tree allows us to infer lower bounds for the number of introductions to each country, and to identify plausible origins of those introductions. It is important to remember that, particularly for countries other than the UK, the full circulating diversity of the variant is probably not being captured, and therefore intermediate transmissions cannot be ruled out. In particular, the closest relative of a particular sequence will often have been sampled in the UK simply because sequencing efforts in the UK exceed those of most other countries by orders of magnitude. It is, however,

not our goal to identify all introductions but to investigate large-scale patterns of spread in Europe.

## Travel volume and destination

Mobile phone roaming data were used to estimate the number of visitors from a given country that departed from a given province for each calendar week. The mobile phone record dataset contains approximately 13 million devices, with more than 2.6 million roammers. A visitor was considered to be departing the country during a given week if they were not seen in the dataset for the next eight weeks. The nationality of a visitor was inferred from the Mobile Country Code (MCC). The total number of unique visitors was aggregated for each province and each week in the period of study; these totals were then scaled using official statistics as a reference to account for the partial coverage of the dataset.

## Estimation of contributions from imports

To estimate how the frequency of 20E (EU1) is expected to change in country  $X$  as a result of travel, we considered the following simple model: A fraction  $\alpha_i$  of the population of  $X$  returns from Spain every week  $i$  (estimated from roaming data, see above) and is infected with 20E (EU1) with a probability  $p_i$  given by its per capita weekly incidence in Spain. Incidence is the weighted average over incidence in Spanish provinces by the distribution of visitors across the provinces. The week-over-week fold-change of the epidemic in  $X$  is calculated as  $g_i = (c_i - \alpha_i p_i)/c_{i-1}$ , where  $c_i$  is the per capita incidence in week  $i$  in  $X$ . This fold-change captures the local growth of the epidemic in country  $X$ . The total number of 20E (EU1) cases  $v_i$  in week  $i$  is hence  $v_i = g_i v_{i-1} + p_i \alpha_i$ , while the total number of non-20E (EU1) cases is  $r_i = g_i r_{i-1}$ . Running this recursion from mid-June to November results in the frequency trajectories in Fig. 3.

From 1 June 2020 to 30 September 2020, the Swiss Federal Office of Public Health (FOPH) reported 23,199 confirmed SARS-CoV-2 infections. Of these cases, 14,583 (62.9%) provided information about their likely place of exposure and country of infection in a clinical registration form. Of these, 3,304 (22.7%) reported exposure abroad and 136 (0.9%) named Spain as the country of infection. The Robert-Koch Institute reported statistics on likely country of infection by calendar week in their daily situation reports<sup>42</sup>.

## Phylodynamic analysis of Swiss transmission chains

We identified introductions into Switzerland and downstream Swiss transmission chains by considering a tree of all available Swiss sequences combined with foreign sequences with high similarity to Swiss sequences (full procedure described previously<sup>40</sup>). Putative transmission chains were defined as majority Swiss clades allowing for at most three ‘exports’ to third countries. Identification of transmission chains is complicated by polytomies in SARS-CoV-2 phylogenies and we bounded the resulting uncertainty by either (i) considering all subtrees descending from the polytomy as separate introductions (called ‘max’ in Extended Data Fig. 8) or (ii) aggregating all into a single introduction (called ‘min’)<sup>40</sup>. We further extended this analysis to include a pan-European dataset consisting of putative transmission chains defined via the collapsed phylogenies discussed above. Specifically, each section of a pie chart, which corresponds to a country-specific collection of sequences, was taken as a single introduction. Non-20E (EU1)  $R_e$  estimates were obtained from case data and the estimated frequency of 20E (EU1) in different countries.

The phylodynamic analysis of the transmission chains was performed using BEAST2 with a birth–death-model tree prior<sup>43,44</sup>. 20E (EU1) and non-20E (EU1) variants share a sampling probability and  $\log R_e$  has an Ornstein–Uhlenbeck prior as described in ref. 40 (but note a different smoothing prior was used there).

## ELISA

We coated 384-well Maxisorp plates (Thermo Fisher) overnight at room temperature with 3 µg/ml of SARS-CoV-2 S2P<sup>45</sup> or SARS-CoV-2 A222V D614G S2P in 20 mM Tris pH 8 and 150 mM NaCl, produced as previously described<sup>21</sup>. In brief, Expi293F cells were transiently transfected with a plasmid containing the spike protein and supernatant was clarified six days later before Ni Sepharose resin purification and flash freezing. Gibco (Fisher) Expi293F cells were used for protein production and have not been authenticated or tested for mycoplasma contamination. They are not in the database of commonly misidentified cell lines. Plates were slapped dry and blocked with Blocker Casein in TBS (Thermo Fisher) for 1 h at 37 °C. Plates were slapped dry and 1 µM S2E12<sup>18</sup>, S309<sup>22</sup>, or 4A8<sup>24</sup> IgG or 1:4 NIBSC human plasma (20/130; <https://www.nibsc.org/documents/ifu/20-130.pdf>) was serially diluted 1:3 in TBST and incubated for one hour at 37 °C. Plates were washed 4× with TBST using a 405 TS Microplate Washer (BioTek) followed by addition of 1:5,000 goat anti-human Fc IgG-HRP (Thermo Fisher) for one hour at 37 °C. Plates were washed 4× and TMB Microwell Peroxidase (Seracare) was added. The reaction was quenched after 1–2 min with 1 N HCl and the A450 of each well was read using a Varioskan Lux plate reader (Thermo Fisher).

## Pseudotyped lentivirus production and titres

The S:A222V mutation was introduced into the protein-expression plasmid HDM-Spiked21-D614G, which encodes a codon-optimized spike from Wuhan-Hu-1 (GenBank NC 045512) with a 21-amino acid cytoplasmic tail deletion and the D614G mutation<sup>46</sup>. This plasmid is also available on AddGene (plasmid 158762). We made two different versions of the A222V mutant that differed only in which codon was used to introduce the valine mutation (either GTT or GTC). The sequences of these plasmids (HDM Spike-d21D614G-A222V-GTT and HDM Spike-d21-D614G-A222V-GTC) are available as supplement files at [https://github.com/neherlab/2020\\_EU1\\_paper/tree/master/plasmid\\_data](https://github.com/neherlab/2020_EU1_paper/tree/master/plasmid_data).

Spike-pseudotyped lentiviruses were produced as described<sup>46</sup>. Two separate plasmid preps of the A222V (GTT) spike and one plasmid prep of the A222V (GTC) spike were each used in duplicate to produce six replicates of A222V spike-pseudotyped lentiviruses. Three plasmid preps of the initial D614G spike plasmid (with the 21-amino acid cytoplasmic tail truncation) were each used once used to make three replicates of D614G spike-pseudotyped lentiviruses. All viruses were titred in duplicate.

Lentiviruses were produced with both Luciferase IRES ZsGreen and ZsGreen-only lentiviral backbones<sup>46</sup>, and then titred using luciferase signal or percentage of fluorescent cells, respectively. All viruses were titred in 293T-ACE2 cells (BEI NR-52511) as described<sup>47</sup>, with the following modifications. Viruses containing luciferase were titred starting at a 1:10 dilution followed by five serial twofold dilutions. The Promega BrightGlo luciferase system was used to measure relative luciferase units (RLUs) ~65 h post-infection and RLUs per ml were calculated at each dilution then averaged across all dilutions for each virus. Viruses containing only ZsGreen were titred starting at a 1:3 dilution followed by four serial fivefold dilutions. The 1:375 dilution was visually determined to be ~1% positive about 65 h post-infection and was used to calculate the percent of infected cells using flow cytometry (BD FACSCelesta cell analyser). Viral titres were then calculated using the percentage of green cells via the Poisson formula. To normalize viral titres by lentiviral particle production, p24 concentration (in pg/ml) was quantified by ELISA according to the manufacturer's instructions (Advanced Bioscience Laboratories Cat. 5421). All viral supernatants were measured in technical duplicates at a 1:100,000 dilution.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

Sequence data were obtained from GISAID and tables listing all accession numbers of sequences are available in the [Supplementary Information](#).

## Code availability

Code used for the above analyses is available on GitHub at [https://github.com/neherlab/2020\\_EU1\\_paper](https://github.com/neherlab/2020_EU1_paper). The code used to run the cluster builds is available on GitHub at [https://github.com/emmahodcroft/ncov\\_cluster](https://github.com/emmahodcroft/ncov_cluster).

## References

1. 1.

WHO Emergency Committee. *Statement on the Second Meeting of the International Health Regulations (2005) Emergency Committee Regarding the Outbreak of Novel Coronavirus (2019-nCoV)* [https://www.who.int/news-room/detail/01-02-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/01-02-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)) (2020).

2. 2.

Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Worobey, M. et al. The emergence of SARS-CoV-2 in Europe and North America. *Science* **370**, 564–570 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

5. 5.

du Plessis, L. et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **371**, 708–712 (2021).

[ADS](#) [Article](#) [Google Scholar](#)

6. 6.

Shu, Y. & McCauley, J. GISAID: global initiative on sharing all influenza data — from vision to reality. *Euro Surveill.* **22**, 30494 (2017).

[Article](#) [Google Scholar](#)

7. 7.

Korber, B. et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812–827.e19 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Volz, E. et al. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell* **184**, 64–75.e11 (2020).

[Article](#) [Google Scholar](#)

9. 9.

Plante, J. A. et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592**, 116–121 (2021).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Yurkovetskiy, L. et al. Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell* **183**, 739–751.e8 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Davies, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, eabg3055 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Volz, E. et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* **593**, 266–269 (2021).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Pearson, C. A. B. et al. Estimates of severity and transmissibility of novel SARS-CoV-2 variant 501Y.V2 in South Africa. Preprint at <https://cmmid.github.io/topics/covid19/sa-novel-variant.html> (2021).

14. 14.

Tegally, H. et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438–443 (2021).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Sabino, E. C. et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* **397**, 452–455 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Tortorici, M. A. et al. Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **370**, 950–957 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Thomson, E. C. et al. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**, 1171–1187 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

McCallum, M., Walls, A. C., Bowen, J. E., Corti, D. & Veesler, D. Structure-guided covalent stabilization of coronavirus spike glycoprotein trimers in the closed conformation. *Nat. Struct. Mol. Biol.* **27**, 942–949 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292.e6 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Pinto, D. et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Walls, A. C. et al. Elicitation of potent neutralizing antibody responses by designed protein nanoparticle vaccines for SARS-CoV-2. *Cell* **183**, 1367–1382.e17 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Chi, X. et al. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**, 650–655 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

25. 25.

McCallum, M. et al. N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332–2347.e16 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Finkel, Y. et al. The coding capacity of SARS-CoV-2. *Nature* **589**, 125–130 (2021).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Pancer, K. et al. The SARS-CoV-2 ORF10 is not essential in vitro or in vivo in humans. *PLoS Pathog.* **16**, e1008959 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Grubaugh, N. D., Hanage, W. P. & Rasmussen, A. L. Making sense of mutation: what D614G means for the COVID-19 pandemic remains unclear. *Cell* **182**, 794–795 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Instituto Nacional de Estadística. *Hotel Industry and Tourism – Tourist Movement on Borders Survey Frontur*. (accessed December 2020);  
[https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica\\_C&cid=1254736176996&menu=resultados&secc=1254736195568&idp=1254735576863](https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176996&menu=resultados&secc=1254736195568&idp=1254735576863)

30. 30.

Aena.es. *Air Traffic Statistics* (accessed January 2021);  
<https://portal.aena.es/en/corporate/air-traffic-statistics.html>

31. 31.

Correa-Martínez, C. L. et al. A pandemic in times of global tourism: superspreading and exportation of COVID-19 cases from a ski area in austria. *J. Clin. Microbiol.* **58**, e00588-20 (2020).

[Article](#) [Google Scholar](#)

32. 32.

Knabl, L. et al. High SARS-CoV-2 seroprevalence in children and adults in the Austrian ski resort Ischgl. Preprint at  
<https://doi.org/10.1101/2020.08.20.20178533> (2020).

33. 33.

Mossong, J. et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).

[Article](#) [Google Scholar](#)

34. 34.

Jarvis, C. I. et al. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med.* **18**, 124 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

35. 35.

Neher, R. A., Dyrdak, R., Druelle, V., Hodcroft, E. B. & Albert, J. Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. *Swiss Med. Wkly.* **150**, w20224 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

36. 36.

Lauring, A. S. & Hodcroft, E. B. Genetic variants of SARS-CoV-2—what do they mean? *J. Am. Med. Assoc.* **325**, 529–531 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

37. 37.

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).

[CAS](#) [Article](#) [Google Scholar](#)

38. 38.

Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

39. 39.

Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, vex042 (2018).

[Article](#) [Google Scholar](#)

40. 40.

Nadeau, S. et al. Quantifying SARS-CoV-2 spread in Switzerland based on genomic sequencing data. Preprint at <https://doi.org/10.1101/2020.10.14.20212621> (2020).

41. 41.

Stange, M. et al. SARS-CoV-2 outbreak in a tri-national urban area is dominated by a B.1 lineage variant linked to mass gathering events. *PLOS Pathog.* **17**, e1009374 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

42. 42.

Robert Koch Institute. *Aktueller Lage-/Situationsbericht des RKI zu COVID-19* [https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Situationsberichte/Gesamt.html](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Gesamt.html) (2020).

43. 43.

Stadler, T., Kühnert, D., Bonhoeffer, S. & Drummond, A. J. Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc. Natl Acad. Sci. USA* **110**, 228–233 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

44. 44.

Bouckaert, R. et al. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* **15**, e1006650 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

45. 45.

Pallesen, J. et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl Acad. Sci. USA* **114**, E7348–E7357 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

46. 46.

Greaney, A. J. et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe* **29**, 44–57.e9 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

47. 47.

Crawford, K. H. D. et al. Protocol and reagents for pseudotyping lentiviral particles with SARS-CoV-2 spike protein for neutralization assays. *Viruses* **12**, 513 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank researchers, clinicians, and public health authorities for making SARS-CoV-2 sequence data available in a timely manner; the COVID-19 Genomics UK consortium for their sequencing efforts, which have provided a third of the sequences currently publicly available; and the Swiss Federal Office of Public Health (FOPH) for providing access to their data. This work was supported by the Swiss National Science Foundation (SNSF) through grant numbers 31CA30 196046 (to R.A.N., E.B.H., and C.L.A.) and 31CA30 196267 (to T.S.), the European Union's Horizon 2020 research and innovation programme project EpiPose (no 101003688) (to M.L.R. and C.L.A.), core funding by the University of Basel and ETH Zürich, the National Institute of General Medical Sciences (R01GM120553 to D.V.), the National Institute of Allergy and Infectious Diseases (DP1AI158186 and HHSN272201700059C to D.V.), a Pew Biomedical Scholars Award (to D.V.), an Investigators in the Pathogenesis of Infectious Disease Awards from the Burroughs Wellcome Fund (to D.V. and J.D.B.), a Fast Grants award (to D.V.), and NIAID grants R01AI141707 (to J.D.B.) and F30AI149928 (to K.H.D.C.). SeqCOVID-SPAIN is funded by the Instituto de Salud Carlos III project COV20/00140, Spanish National Research Council and ERC StG 638553 to I.C. and BFU2017-89594R from MICIN to F.G.C. J.D.B. is an Investigator of the Howard Hughes Medical Institute.

## Author information

### Author notes

1. These authors jointly supervised this work: Tanja Stadler, Richard A. Neher

### Affiliations

1. Biozentrum, University of Basel, Basel, Switzerland

Emma B. Hodcroft, Moira Zuber & Richard A. Neher

2. Swiss Institute of Bioinformatics, Basel, Switzerland

Emma B. Hodcroft, Sarah Nadeau, Timothy G. Vaughan, Tanja Stadler & Richard A. Neher

3. Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

Emma B. Hodcroft, Christian L. Althaus & Martina L. Reichmuth

4. Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland

Sarah Nadeau, Timothy G. Vaughan & Tanja Stadler

5. Division of Basic Sciences and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Katharine H. D. Crawford & Jesse D. Bloom

6. Department of Genome Sciences, University of Washington, Seattle, WA, USA

Katharine H. D. Crawford & Jesse D. Bloom

7. Medical Scientist Training Program, University of Washington, Seattle, WA, USA

Katharine H. D. Crawford

8. Department of Biochemistry, University of Washington, Seattle, WA, USA

John E. Bowen, Alexandra C. Walls & David Veesler

9. Humabs Biomed SA, a subsidiary of Vir Biotechnology, Bellinzona, Switzerland

Davide Corti

10. Howard Hughes Medical Institute, Seattle, WA, USA

Jesse D. Bloom

11. Kido Dynamics SA, Lausanne, Switzerland

David Mateo & Alberto Hernando

12. Tuberculosis Genomics Unit, Biomedicine Institute of Valencia (IBV-CSIC), Valencia, Spain

Iñaki Comas, Galo Adrian Goig, Álvaro Chiner-Oms, Irving Cancino-Muñoz, Mariana Gabriela López, Manuela Torres-Puente, Inmaculada Gomez-Navarro, Santiago Jiménez-Serrano & Jordi Pérez-Tur

13. CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

Iñaki Comas, Fernando González-Candelas, Fernando González-Candelas, María Alma Bracho, Giuseppe D'Auria, Alfredo Mayor, Elisa Martró & Vicente Martín-Sánchez

14. Joint Research Unit “Infection and Public Health” FISABIO-University of Valencia, Institute for Integrative Systems Biology (I2SysBio), Valencia, Spain

Fernando González-Candelas, Fernando González-Candelas, Lidia Ruiz-Roldán, María Alma Bracho & Neris García-González

15. FISABIO, Servicio de Secuenciación, Valencia, Spain

Llúcia Martínez-Priego, Inmaculada Galán-Vendrell, Paula Ruiz-Hueso, Griselda De Marco, Maria Loreto Ferrús, Sandra Carbó-Ramírez & Giuseppe D’Auria

16. Instituto de Biología Integrativa de Sistemas, I2SysBio (CSIC-Universitat de València), Valencia, Spain

Mireia Coscollá & Paula Ruiz-Rodríguez

17. Servicio de Vigilancia y Control Epidemiológico, Dirección General de Salud Pública y Adicciones, Conselleria de Sanitat Universal i Salut Pública, Generalitat Valenciana, Valencia, Spain

Francisco Javier Roig-Sena

18. Real Jardín Botánico, Consejo Superior de Investigaciones Científicas, Madrid, Spain

Isabel Sanmartín

19. Genomes and Disease, Centre for Research in Molecular Medicine and Chronic Diseases (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

Daniel Garcia-Souto, Ana Pequeno-Valtierra, Jose M. C. Tubio & Jorge Rodríguez-Castro

20. Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

Daniel Garcia-Souto & Jose M. C. Tubio

21. Cancer Ageing and Somatic Mutation Programme, Wellcome Sanger Institute, Cambridge, UK

Daniel Garcia-Souto

22. Servei de Microbiologia, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

Nuria Rabella, Ferrán Navarro & Elisenda Miró

23. CREPIMC, Institut d'Investigació Biomèdica Sant Pau, Barcelona, Spain

Nuria Rabella, Ferrán Navarro & Elisenda Miró

24. Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, Cerdanyola, Spain

Nuria Rabella & Ferrán Navarro

25. Servicio de Microbiología, H.U. Puerta del Mar, Cádiz, Spain

Manuel Rodríguez-Iglesias, Fátima Galán-Sánchez & Salud Rodriguez-Pallares

26. INIBICA, Instituto de Investigación Biomédica de Cádiz, Cádiz, Spain

Manuel Rodríguez-Iglesias, Fátima Galán-Sánchez & Salud Rodriguez-Pallares

27. Departamento de Biomedicina, Biotecnología y Salud Pública, Facultad de Medicina, Universidad de Cádiz, Cádiz, Spain

Manuel Rodríguez-Iglesias & Fátima Galán-Sánchez

28. Plataforma de Genómica y Bioinformática, Centro de Investigación Biomédica de La Rioja (CIBIR), Logroño, Spain

Maria de Toro & María Bea Escudero

29. Laboratorio de Microbiología, Hospital San Pedro, Logroño, Spain

José Manuel Azcona-Gutiérrez & Miriam Blasco Alberdi

30. ISGlobal, Institute for Global Health, Barcelona, Spain

Alfredo Mayor, Alberto L. García-Basteiro, Gemma Moncunill, Carlota Dobaño, Pau Cisteró, Miguel J. Martínez, Jordi Vila, Elisa Rubio, Aida Peiró-Mestres & Jessica Navero-Castillejos

31. Centro de Investigação em Saúde de Manhiça (CISM), Maputo, Mozambique

Alfredo Mayor & Alberto L. García-Basteiro

32. Microbiology Department, Hospital Clinic I Provincial de Barcelona, Barcelona, Spain

Alfredo Mayor, Alberto L. García-Basteiro, Gemma Moncunill, Carlota Dobaño, Pau Cisteró & María Ángeles Marcos

33. Servicio de Microbiología Clínica y Enfermedades Infecciosas, Hospital General Universitario Gregorio Marañón, Madrid, Spain

Darío García-de-Viedma, Laura Pérez-Lago, Marta Herranz, Jon Sicilia, Pilar Catalán-Alonso & Patricia Muñoz

34. Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

Darío García-de-Viedma, Laura Pérez-Lago, Marta Herranz, Jon Sicilia, Pilar Catalán-Alonso & Patricia Muñoz

35. CIBER Enfermedades Respiratorias (CIBERES), Madrid, Spain

Darío García-de-Viedma, Marta Herranz, Pilar Catalán-Alonso & Patricia Muñoz

36. Servicio de Microbiología Clínica, Hospital San Pedro de Alcántara, Cáceres, Spain

Cristina Muñoz-Cuevas & Guadalupe Rodríguez-Rodríguez

37. Servicio Extremeño de Salud, Cáceres, Spain

Cristina Muñoz-Cuevas & Guadalupe Rodríguez-Rodríguez

38. Servicio de Microbiología, Hospital Dr Peset, Valencia, Spain

Juan Alberola-Enguidanos, Jose Miguel Nogueira & Juan José Camarena

39. Conselleria de Sanitat i Consum, Generalitat Valenciana, Valencia, Spain

Juan Alberola-Enguidanos, Jose Miguel Nogueira & Juan José Camarena

40. Departamento Microbiología, Facultad de Medicina, Universidad de Valencia, Valencia, Spain

Juan Alberola-Enguidanos, Jose Miguel Nogueira, Juan José Camarena, David Navarro & Concepción Gimeno-Cardona

41. Servicio de Microbiología Clínica Hospital Universitario Miguel Servet, Zaragoza, Spain

Antonio Rezusta, Alexander Tristáncho-Baró, Ana Milagro, Nieves Felisa Martínez-Cameo & Yolanda Gracia-Grataloup

42. Instituto de Investigación Sanitaria de Aragón, Centro de Investigación Biomédica de Aragón (CIBA), Zaragoza, Spain

Antonio Rezusta & Alexander Tristáncho-Baró

43. Facultad de Medicina, Universidad de Zaragoza, Zaragoza, Spain

Antonio Rezusta, Rafael Benito & Sonia Algarate

44. Servicio de Microbiología, Laboratori Clínic Metropolitana Nord, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain

Elisa Martró, Antoni E. Bordoy, Anna Not & Adrián Antuori-Torres

45. Hospital Clínico Universitario Lozano Blesa, Zaragoza, Spain

Rafael Benito, Sonia Algarate & Jessica Bueno

46. Servicio de Enfermedades Infecciosas y Microbiología Clínica, Clínica Universidad de Navarra, Pamplona, Spain

Jose Luis del Pozo

47. Servicio de Microbiología, Hospital Universitario Central de Asturias, Oviedo, Spain

Jose Antonio Boga, Cristián Castelló-Abietar, Susana Rojo-Alba, Marta Elena Alvarez-Argüelles & Santiago Melon

48. Grupo de Microbiología Traslacional Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Spain

Jose Antonio Boga, Cristián Castelló-Abietar, Susana Rojo-Alba, Marta Elena Alvarez-Argüelles & Santiago Melon

49. Servicio de Microbiología, Hospital Universitario Cruces, Bilbao, Spain

Maitane Aranzamendi-Zaldumbide

50. Grupo de Microbiología y Control de Infección, Instituto de Investigación Sanitaria Biocruces, Bizkaia, Spain  
Maitane Aranzamendi-Zaldumbide
51. Servicio de Microbiología & CORE de Biología Molecular, CDB, Hospital Clínic, Barcelona, Spain  
Andrea Vergara-Gómez
52. Centro de Investigación en Sanidad Animal, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, O.A., M.P. - INIA, Valdeolmos, Spain  
Jovita Fernández-Pinero
53. Departamento de Microbiología, Hospital Clínic de Barcelona, Barcelona, Spain  
Miguel J. Martínez, Jordi Vila, Elisa Rubio, Aida Peiró-Mestres & Jessica Navero-Castillejos
54. CINBIO, Universidade de Vigo, Vigo, Spain  
David Posada, Diana Valverde, Nuria Estévez-Gómez, Iria Fernandez-Silva, Loretta de Chiara, Pilar Gallego-García & Nair Varela
55. Department of Biochemistry, Genetics, and Immunology, Universidade de Vigo, Vigo, Spain  
David Posada, Diana Valverde, Iria Fernandez-Silva & Loretta de Chiara
56. Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-UVIGO, Vigo, Spain  
David Posada & Diana Valverde
57. Hospital General Universitario de Castellón, Castellón, Spain  
Rosario Moreno & María Dolores Tirado
58. IdISSC/Hospital Clínico San Carlos, Madrid, Spain  
Ulises Gomez-Pinedo
59. Hospital Marqués de Valdecilla - IDIVAL, Santander, Spain

Mónica Gozalo-Margüello, Maria Eliecer-Cano, José Manuel Méndez-Legaza, Jesus Rodríguez-Lozano, María Siller & Daniel Pablo-Marcos

60. Servicio de Microbiología, Hospital Universitario Son Espases, Palma de Mallorca, Spain

Antonio Oliver, Jordi Reina & Carla López-Causapé

61. Instituto de Investigación Sanitaria de las Islas Baleares, Palma de Mallorca, Spain

Antonio Oliver & Carla López-Causapé

62. Servicio de Microbiología, Hospital Universitario de Álava, Osakidetza-Servicio Vasco de Salud, Vitoria-Gasteiz (Álava), Spain

Andrés Canut-Blasco, Silvia Hernández-Crespo, María Luz A. Cordón, María-Concepción Lecároz-Agara, Carmen Gómez-González, Amaia Aguirre-Quiñonero, José Israel López-Mirones, Marina Fernández-Torres & María Rosario Almela-Ferrer

63. Servicio Microbiología, Departamento de Salud de Elche-Hospital General, Elche, Spain

Nieves Gonzalo-Jiménez, María Montserrat Ruiz-García, Antonio Galiana & Judith Sanchez-Almendro

64. Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández, Elche, Spain

Maria Montserrat Ruiz-García

65. Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana, Elche, Spain

Antonio Galiana & Judith Sanchez-Almendro

66. Biodonostia, Osakidetza, Hospital Universitario Donostia, Servicio de Microbiología, San Sebastián, Spain

Gustavo Cilla, Milagrosa Montes, Luis Piñeiro, Ane Sorarain & José María Marimón

67. Hospital Universitario y Politécnico La Fe, Servicio de Microbiología, Valencia, Spain

Maria Dolores Gomez-Ruiz, José Luis López-Hontangas & Eva M. González Barberá

68. Servicio de Microbiología, Hospital Universitario Virgen de las Nieves, Granada, Spain

José María Navarro-Marí, Irene Pedrosa-Corral & Sara Sanbonmatsu-Gámez

69. Hospital Universitario Virgen de las Nieves, Instituto de Investigación Biosanitaria ibs, Granada, Spain

José María Navarro-Marí, Irene Pedrosa-Corral & Sara Sanbonmatsu-Gámez

70. Hospital Universitario de Gran Canaria Dr. Negrín, Las Palmas de Gran Canaria, Spain

Carmen Pérez-González, Francisco Chamizo-López & Ana Bordes-Benítez

71. Microbiology Service, Hospital Clínico Universitario, INCLIVA Research Institute, Valencia, Spain

David Navarro, Eliseo Albert & Ignacio Torres

72. Laboratorio de Microbiología, Hospital General Universitario de Elda, Elda, Spain

Isabel Gascón, Cristina Juana Torregrosa-Hetland, Eva Pastor-Boix & Paloma Cascales-Ramos

73. Servicio de Microbiología. Consorcio Hospital General Universitario de Valencia, Valencia, Spain

Begoña Fuster-Escrivá, Concepción Gimeno-Cardona, María Dolores Ocete & Rafael Medina-Gonzalez

74. Laboratorio Biología Molecular, Área de Diagnóstico Biológico, Hospital Universitario La Ribera, Alzira, Valencia, Spain

Julia González-Cantó & Olalla Martínez-Macias

75. Servicio de Microbiología, Hospital Regional Universitario de Málaga, Málaga, Spain

Begoña Palop-Borrás, Inmaculada de Toro, María Concepción Mediavilla-Gradolph & Mercedes Pérez-Ruiz

76. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, O.A., M.P. - INIA, Madrid, Spain

Óscar González-Recio & Mónica Gutiérrez-Rivas

77. Hospital General Universitario de Albacete, Albacete, Spain

Encarnación Simarro-Córdoba, Julia Lozano-Serra & Lorena Robles-Fonseca

78. Hospital Universitario San Cecilio, Granada, Spain

Adolfo de Salazar, Laura Viñuela-González, Natalia Chueca, Federico García & Cristina Gómez-Camarasa

79. Animal Health Department, Universidad de León, León, Spain

Ana Carvajal, Raul de la Puente, Juan-Miguel Fregeneda-Grandes & Héctor Argüello

80. Research Group on Gene-Environment Interactions and Health, Institute of Biomedicine (IBIOMED), Universidad de León, León, Spain

Vicente Martín-Sánchez, Antonio José Molina & Tania Fernández-Villa

81. Servicio de Microbiología, Hospital Arnau de Vilanova, Valencia, Spain

Maria Amparo Farga-Martí & Victoria Domínguez-Márquez

82. Hospital Clínico Universitario de Santiago de Compostela, Santiago de Compostela, Spain

José Javier Costa-Alcalde, Rocío Trastoy, Gema Barbeito-Castiñeiras, Amparo Coira, María Luisa Pérez-del-Molino & Antonio Aguilera

83. Biomedical Research Institute of Barcelona (IIBB), Spanish National Research Council (CSIC), Barcelona, Spain

Anna M. Planas

84. Servicio de Enfermedades Infecciosas, Hospital Clínic de Barcelona, Barcelona, Spain

Alex Soriano

85. Biomedical Research Institute Sant Pau (IIB Sant Pau), Barcelona, Spain

Israel Fernandez-Cádenas

86. Institut of Global Health of Barcelona (ISGlobal), Barcelona, Spain

Maria Ángeles Marcos

87. Servicio de Microbiología, Hospital Clínico Universitario Virgen de la Arrixaca, Departamento de Genética y Microbiología, Universidad de Murcia, Murcia, Spain

Antonio Moreno-Docón

88. Hospital Universitario 12 de Octubre, Madrid, Spain

Esther Viedma

89. Hospital Universitario La Paz, Madrid, Spain

Jesús Mingorance

90. Hospital Universitario Ramón y Cajal, Madrid, Spain

Juan Carlos Galán-Montemayor

91. Laboratorio de Microbiología, Hospital Marina Baixa, Villajoyosa, Spain

Mónica Parra-Grande

## Authors

1. Emma B. Hodcroft

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Moira Zuber

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Sarah Nadeau

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Timothy G. Vaughan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Katharine H. D. Crawford

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Christian L. Althaus

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Martina L. Reichmuth

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. John E. Bowen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Alexandra C. Walls

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Davide Corti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Jesse D. Bloom

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. David Veesler

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. David Mateo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Alberto Hernando

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Iñaki Comas

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Fernando González-Candelas

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Tanja Stadler

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Richard A. Neher

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## **Consortia**

### **SeqCOVID-SPAIN consortium**

- Fernando González-Candelas
- , Galo Adrian Goig
- , Álvaro Chiner-Oms
- , Irving Cancino-Muñoz
- , Mariana Gabriela López
- , Manuela Torres-Puente
- , Inmaculada Gomez-Navarro
- , Santiago Jiménez-Serrano
- , Lidia Ruiz-Roldán
- , María Alma Bracho
- , Neris García-González
- , Llúcia Martínez-Priego
- , Inmaculada Galán-Vendrell
- , Paula Ruiz-Hueso
- , Griselda De Marco
- , Maria Loreto Ferrús
- , Sandra Carbó-Ramírez
- , Giuseppe D'Auria
- , Mireia Coscollá
- , Paula Ruiz-Rodríguez
- , Francisco Javier Roig-Sena
- , Isabel Sanmartín
- , Daniel Garcia-Souto
- , Ana Pequeno-Valtierra
- , Jose M. C. Tubio
- , Jorge Rodríguez-Castro
- , Nuria Rabella
- , Ferrán Navarro
- , Elisenda Miró
- , Manuel Rodríguez-Iglesias
- , Fátima Galán-Sánchez
- , Salud Rodriguez-Pallares

- , María de Toro
- , María Bea Escudero
- , José Manuel Azcona-Gutiérrez
- , Miriam Blasco Alberdi
- , Alfredo Mayor
- , Alberto L. García-Basteiro
- , Gemma Moncunill
- , Carlota Dobaño
- , Pau Cisteró
- , Darío García-de-Viedma
- , Laura Pérez-Lago
- , Marta Herranz
- , Jon Sicilia
- , Pilar Catalán-Alonso
- , Patricia Muñoz
- , Cristina Muñoz-Cuevas
- , Guadalupe Rodríguez-Rodríguez
- , Juan Alberola-Enguidanos
- , Jose Miguel Nogueira
- , Juan José Camarena
- , Antonio Rezusta
- , Alexander Tristáncho-Baró
- , Ana Milagro
- , Nieves Felisa Martínez-Cameo
- , Yolanda Gracia-Grataloup
- , Elisa Martró
- , Antoni E. Bordoy
- , Anna Not
- , Adrián Antuori-Torres
- , Rafael Benito
- , Sonia Algarate
- , Jessica Bueno
- , Jose Luis del Pozo
- , Jose Antonio Boga
- , Cristián Castelló-Abietar
- , Susana Rojo-Alba
- , Marta Elena Alvarez-Argüelles

- , Santiago Melon
- , Maitane Aranzamendi-Zaldumbide
- , Andrea Vergara-Gómez
- , Jovita Fernández-Pinero
- , Miguel J. Martínez
- , Jordi Vila
- , Elisa Rubio
- , Aida Peiró-Mestres
- , Jessica Navero-Castillejos
- , David Posada
- , Diana Valverde
- , Nuria Estévez-Gómez
- , Iria Fernandez-Silva
- , Loretta de Chiara
- , Pilar Gallego-García
- , Nair Varela
- , Rosario Moreno
- , Maria Dolores Tirado
- , Ulises Gomez-Pinedo
- , Mónica Gozalo-Margüello
- , María Eliecer-Cano
- , José Manuel Méndez-Legaza
- , Jesus Rodríguez-Lozano
- , María Siller
- , Daniel Pablo-Marcos
- , Antonio Oliver
- , Jordi Reina
- , Carla López-Causapé
- , Andrés Canut-Blasco
- , Silvia Hernáez-Crespo
- , María Luz A. Cordón
- , María-Concepción Lecároz-Agara
- , Carmen Gómez-González
- , Amaia Aguirre-Quiñonero
- , José Israel López-Mirones
- , Marina Fernández-Torres
- , María Rosario Almela-Ferrer

- , Nieves Gonzalo-Jiménez
- , Maria Montserrat Ruiz-García
- , Antonio Galiana
- , Judith Sanchez-Almendro
- , Gustavo Cilla
- , Milagrosa Montes
- , Luis Piñeiro
- , Ane Sorarrain
- , José María Marimón
- , Maria Dolores Gomez-Ruiz
- , José Luis López-Hontangas
- , Eva M. González Barberá
- , José María Navarro-Marí
- , Irene Pedrosa-Corral
- , Sara Sanbonmatsu-Gámez
- , Carmen Pérez-González
- , Francisco Chamizo-López
- , Ana Bordes-Benítez
- , David Navarro
- , Eliseo Albert
- , Ignacio Torres
- , Isabel Gascón
- , Cristina Juana Torregrosa-Hetland
- , Eva Pastor-Boix
- , Paloma Cascales-Ramos
- , Begoña Fuster-Escrivá
- , Concepción Gimeno-Cardona
- , María Dolores Ocete
- , Rafael Medina-Gonzalez
- , Julia González-Cantó
- , Olalla Martínez-Macias
- , Begoña Palop-Borrás
- , Inmaculada de Toro
- , Maria Concepción Mediavilla-Gradolph
- , Mercedes Pérez-Ruiz
- , Óscar González-Recio
- , Mónica Gutiérrez-Rivas

- , Encarnación Simarro-Córdoba
- , Julia Lozano-Serra
- , Lorena Robles-Fonseca
- , Adolfo de Salazar
- , Laura Viñuela-González
- , Natalia Chueca
- , Federico García
- , Cristina Gómez-Camarasa
- , Ana Carvajal
- , Raul de la Puente
- , Vicente Martín-Sánchez
- , Juan-Miguel Fregeneda-Grandes
- , Antonio José Molina
- , Héctor Argüello
- , Tania Fernández-Villa
- , María Amparo Farga-Martí
- , Victoria Domínguez-Márquez
- , José Javier Costa-Alcalde
- , Rocío Trastoy
- , Gema Barbeito-Castiñeiras
- , Amparo Coira
- , María Luisa Pérez-del-Molino
- , Antonio Aguilera
- , Anna M. Planas
- , Alex Soriano
- , Israel Fernandez-Cádenas
- , Jordi Pérez-Tur
- , María Ángeles Marcos
- , Antonio Moreno-Docón
- , Esther Viedma
- , Jesús Mingorance
- , Juan Carlos Galán-Montemayor
- & Mónica Parra-Grande

## Contributions

E.B.H. identified the cluster, led the analysis, created figures, and drafted the manuscript. R.A.N. analysed data, created figures, and drafted the manuscript. M.Z., S.N., T.G.V., C.L.A., T.S., and M.L.R. analysed data and created figures. D.V. investigated structural aspects and created figures. J.D.B., J.E.B., A.C.W., D.C., and K.H.D.C. performed experimental assays and created figures. I.C. and F.G.C. interpreted the origins of the cluster and contributed data. D.M. and A.H. contributed and interpreted data. All authors contributed to and approved the final manuscript.

## Corresponding authors

Correspondence to [Emma B. Hodcroft](#) or [Richard A. Neher](#).

## Ethics declarations

## Competing interests

D.V. is a consultant for Vir Biotechnology. D.C. is an employee of Vir Biotechnology and may hold shares in Vir Biotechnology. The Veesler laboratory has received an unrelated sponsored research agreement from Vir Biotechnology. A.H. is a co-founder of Kido Dynamics and D.M. is employed by Kido Dynamics. The other authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

## Extended Data Fig. 1 Variant dynamics in different European countries.

In countries with at least ten sequences that fall into any of the defined clusters, the proportion of sequences per ISO week that fall into each cluster is shown.

## Extended Data Fig. 2 Structure model of the SARS-CoV-2 spike protein.

Two orthogonal orientations of the SARS-CoV-2 spike glycoprotein trimer highlighting the position of the variants described in the manuscript and the RBD and NTD (domain A).

## Extended Data Fig. 3 The substitution A222V in the spike protein has no substantial effect on antigenic properties or replication of pseudotyped lentiviruses.

**a**, Binding of a serial dilution of NIBSC convalescent plasma to immobilized SARSCoV-2 2P S (blue) or SARS-CoV-2 2P A222V D614G S (red). **b**, **c**, Binding of serially diluted concentrations of the human neutralizing antibodies S2E12 (**b**) and S309 (**c**) to immobilized SARSCoV-2 2P S (blue) or SARS-CoV-2 2P A222V D614G S (red). **d**, Binding of serially diluted concentrations of the human neutralizing antibody 4A8 to immobilized SARS-CoV-2 2P S (blue) or SARS-CoV-2 2P A222V D614G S (red).  $n = 2$  experiments performed with independent protein preparations (each in duplicate). Each data point consists of a technical duplicate of each antibody or plasma dilution; error bars, s.d. The experiment shown is representative of two independent experiments. **e**, Titres of lentiviral particles carrying luciferase in the viral genome. Horizontal line, mean. **f**, Titres of lentiviral particles carrying the fluorescent protein ZsGreen in the viral genome. Horizontal line, mean. In **e**, **f**, titres with the A222V mutation are on average higher by a factor of 1.3. **g**, Titres of lentiviral particles carrying luciferase in the viral genome normalized by the p24 concentration ( $\text{pg ml}^{-1}$ ) of each viral supernatant. After p24 normalization, the difference in titre shrinks from 1.28- to 1.14-fold, increasing the  $P$  value to 0.16. **h**,

Titres of lentiviral particles carrying ZsGreen in the viral genome normalized by the p24 concentration ( $\text{pg ml}^{-1}$ ) of each viral supernatant.  $P$  values calculated using two-sided  $t$ -test.

### **Extended Data Fig. 4 Collapsed genotype phylogeny and statistics of putative introductions.**

**a**, The phylogeny shown is the subtree of the 20E (EU1) cluster using data from samples collected before 30 September 2020 and available on GISAID as of January 2021, with sequences carrying all six defining mutations. Pie charts show the representation of sequences from selected countries at each node. The size of each pie chart indicates the total number of sequences at each node. Pie chart fractions scale nonlinearly with the true counts (fourth root) to ensure that all countries are visible and branch lengths are jittered to reduce overlap. Although the jitter means that branch lengths should be interpreted with caution, the smallest branches shown in the tree are equal to one mutation. **b–d**, An example of how the pie chart phylogeny was created. The tree is shown in ‘divergence view’ with the branch lengths in mutations. Internal nodes are shown as horizontal lines with other nodes (internal and external) branching from them. If sequences are identical, they align on the horizontal line. In this example zooming in to the Norwegian cluster, the outermost tips are first collapsed down to their parental node (**b**), forming a pie chart that consists only of sequences from Denmark (**c**). This single-country pie chart is collapsed with the next level of nodes (**d**), including more sequences from Denmark and sequences from Norway, to form a multi-country pie chart. **e**, Rank-order plots of sizes of clusters of sequences in the pie chart slices, in different countries, compatible with a single introduction. Countries such as Norway and Iceland have relatively few clusters, with one or two large clusters dominating, which suggests that a small number of introductions dominated 20E (EU1) circulation. Countries such as the UK and Denmark, on the other hand, show many clusters of varying size, indicating that multiple introductions led to onward spread. The legend indicates total number of sequences  $s$  and number of clusters  $c$ .

**Extended Data Fig. 5 Incidence in various countries over the summer.**

**a**, Spain and Belgium had relatively higher incidence from the start of July compared with other countries in Europe. **b**, The estimated total number of EU1 infections (red) outside Spain (countries as in **a**) surpassed the number in Spain in September.

**Extended Data Fig. 6 Rescaled predictions by the import model match observed frequency trajectories.**

In most countries, observations of 20E (EU1) increased in July 2020 and reached a plateau or a slower increase by October 2020. Predictions by the import model need to be scaled by a factor between 1.2 and 12 (see legends in individual plots) to match the observed frequencies (see main text for discussion). Fluctuations on short time scales in the observed frequency of 20E (EU1) are probably due to sampling and the dynamics of local outbreaks. Observed frequencies are subject to variable reporting delays.

**Extended Data Fig. 7 Reported and estimated introductions of 20E (EU1) to Germany and Switzerland and incidence in Spain by province and similarity in Spanish province travel destination of selected European countries.**

**a**, Germany; **b**, Switzerland. Travel estimate is estimated introductions from Spain based on incidence and roaming data. Reported cases are cases with a suspected origin in Spain as reported by the RKI (Robert Koch Institute, 2020) and the Federal Office of Public Health (FOPH) for Germany and Switzerland, respectively. In Switzerland the adjusted rep. cases accounts for the fact that 37% of case reports lack exposure information. **c**, Incidence in Spain in early and mid-summer. **d**, Distributions of visitors to Spain from different countries. **e**, Similarities of destinations in Spain among visitors from different countries in calendar weeks 28–35.

## Extended Data Fig. 8 Phylodynamic analysis of the spread of the 20E (EU1) variant across Europe and in Switzerland.

**a, d**, Across Europe; **b, c, e, f**, in Switzerland. **a–c**, The sizes of putative transmission chains caused by introductions into Europe and Switzerland. Not shown are the number of singletons, which are introductions with no evidence of onward transmission. In Switzerland, these are shown under two extreme definitions of an introduction (min and max; see [Methods](#)). Depending on the min or max definition of introductions, there were between 14 and 236 singletons of 20E (EU1) infection (41% or 81% of all 20E (EU1) introductions) and between 62 and 1,089 singletons of other variants (30% or 79% of all non-20E (EU1) introductions). In Europe, we see 206 20E (EU1) singletons (46% of all 20E (EU1) introductions). There were also a small number of larger transmission chains that included more than 53 transmissions (20 across all datasets), which are not shown in the histograms. **d–f**, The effective reproductive number estimated for 20E (EU1) (red) and the non-20E (EU1) variants (blue). In Switzerland, this was done for the two extreme definitions of an introduction. For Europe, non-20E (EU1)  $R_e$  estimates were generated from case numbers. Although there are few data to inform estimates of  $R_e$  for 20E (EU1) in July and it differs little from the prior, there is some evidence that 20E (EU1) was growing faster than other variants in August. However, systematic differences in ascertainment in travel associated cases might confound this inference. From mid-September,  $R_e$  of 20E (EU1) is largely statistically indistinguishable from that of other variants. Shaded areas indicate 95% highest posterior density regions. Notably, the peak in August in the Swiss analysis is larger under the ‘min’ definition (**f**) than under the ‘max’ definition (**e**), consistent with a more conservative definition of a cluster which would then require more onward transmission. See Extended Data Fig. [9](#) for details on all variants circulating in Switzerland May–October 2020.

## Extended Data Fig. 9 Lineages found in a Swiss-focused Nextstrain build.

A lineage is defined as a node present in the tree after the cut-off date of 1 May 2020 with at least ten Swiss sequences as children. Clusters discussed in this manuscript are labelled. Lineages are shown as the proportion of the total number of sequences per week in Switzerland. Striped space in the bottom graph represents lineages with most recent common ancestors dating back before 1 May 2020 and lineages that do not contain at least ten Swiss sequences.

**Extended Data Table 1 Representative mutations of 20E (EU1) (the focus of this study) and other notable variants**

[Full size table](#)

## **Supplementary information**

### [Supplementary Information](#)

This file contains additional information about the first sequences identified for selected countries across Europe, as well as information about travel restrictions and quarantine requirements over the summer of 2020. It also includes Supplementary Tables 1-2.

### [Reporting Summary](#)

### [Supplementary Information](#)

This file contains contains the attributions for the data providers and submitters for the sequences used in this analysis.

## **Rights and permissions**

### [Reprints and Permissions](#)

## **About this article**



Check for  
updates

## Cite this article

Hodcroft, E.B., Zuber, M., Nadeau, S. *et al.* Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595**, 707–712 (2021). <https://doi.org/10.1038/s41586-021-03677-y>

[Download citation](#)

- Received: 25 November 2020
- Accepted: 28 May 2021
- Published: 07 June 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03677-y>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

## Associated Content

Collection

[Coronavirus](#)

## Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03677-y>.

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

Untangling introductions and persistence in COVID-19 resurgence in Europe  
[Download PDF](#)

- Article
- [Published: 30 June 2021](#)

# Untangling introductions and persistence in COVID-19 resurgence in Europe

- [Philippe Lemey](#) [ORCID: orcid.org/0000-0003-2826-5353<sup>1,2</sup>](#),
- [Nick Ruktanonchai<sup>3,4</sup>](#),
- [Samuel L. Hong](#) [ORCID: orcid.org/0000-0001-6354-4943<sup>1</sup>](#),
- [Vittoria Colizza](#) [ORCID: orcid.org/0000-0002-2113-2374<sup>5</sup>](#),
- [Chiara Poletto](#) [ORCID: orcid.org/0000-0002-4051-1716<sup>5</sup>](#),
- [Frederik Van den Broeck<sup>1,6</sup>](#),
- [Mandev S. Gill<sup>1</sup>](#),
- [Xiang Ji](#) [ORCID: orcid.org/0000-0002-7243-0865<sup>7</sup>](#),
- [Anthony Levasseur<sup>8</sup>](#),
- [Bas B. Oude Munnink](#) [ORCID: orcid.org/0000-0002-9394-1189<sup>9</sup>](#),
- [Marion Koopmans<sup>9</sup>](#),
- [Adam Sadilek](#) [ORCID: orcid.org/0000-0003-0784-7098<sup>10</sup>](#),
- [Shengjie Lai](#) [ORCID: orcid.org/0000-0001-9781-8148<sup>3</sup>](#),
- [Andrew J. Tatem](#) [ORCID: orcid.org/0000-0002-7270-941X<sup>3</sup>](#),
- [Guy Baele](#) [ORCID: orcid.org/0000-0002-1915-7732<sup>1</sup>](#),
- [Marc A. Suchard](#) [ORCID: orcid.org/0000-0001-9818-479X<sup>11,12,13</sup>](#) &
- [Simon Dellicour](#) [ORCID: orcid.org/0000-0001-9558-1052<sup>1,14</sup>](#)

*Nature* volume 595, pages 713–717 (2021) [Cite this article](#)

- 16k Accesses
- 254 Altmetric
- [Metrics details](#)

## Subjects

- [Epidemiology](#)
- [Phylogenetics](#)
- [SARS-CoV-2](#)

## Abstract

After the first wave of SARS-CoV-2 infections in spring 2020, Europe experienced a resurgence of the virus starting in late summer 2020 that was deadlier and more difficult to contain<sup>1</sup>. Relaxed intervention measures and summer travel have been implicated as drivers of the second wave<sup>2</sup>. Here we build a phylogeographical model to evaluate how newly introduced lineages, as opposed to the rekindling of persistent lineages, contributed to the resurgence of COVID-19 in Europe. We inform this model using genomic, mobility and epidemiological data from 10 European countries and estimate that in many countries more than half of the lineages circulating in late summer resulted from new introductions since 15 June 2020. The success in onward transmission of newly introduced lineages was negatively associated with the local incidence of COVID-19 during this period. The pervasive spread of variants in summer 2020 highlights the threat of viral dissemination when restrictions are lifted, and this needs to be carefully considered in strategies to control the current spread of variants that are more transmissible and/or evade immunity. Our findings indicate that more effective and coordinated measures are required to contain the spread through cross-border travel even as vaccination is reducing disease burden.

[Download PDF](#)

## Main

Upon successfully curbing transmission in spring 2020, many European countries witnessed a resurgence in cases of COVID-19 in the late summer. The number of COVID-19 infections increased rapidly, and by the end of October, it was clear that the continent was deep into a second epidemic wave. This forced governments to reimpose lockdowns and social restrictions in an effort to contain the resurgence. Although these measures reduced infection rates across Europe<sup>3</sup>, several countries witnessed a stabilization at high levels or even a new surge in infections. The spread of more transmissible variants, in particular B.1.1.7 (Alpha variant or 20I (V1)), which was first identified in the UK<sup>4</sup>, has considerably exacerbated the challenge to contain COVID-19.

Already early on in the pandemic, modelling studies warned about new waves due to partial relaxation of restrictions<sup>5</sup> or seasonal variations<sup>6</sup>. By mid-April, the European Commission constructed a roadmap to lifting coronavirus containment measures<sup>7</sup>, recommending a cautious and coordinated manner to revive social and economic activities. However, the early start of the devastating second wave demonstrated that there was insufficient adherence to these measured recommendations. Cross-border travel, and mass tourism in particular, has been implicated as a major instigator of the second wave. Genomic surveillance demonstrated that a new variant (lineage B.1.177<sup>8</sup>, 20E (EU1) (<https://nextstrain.org/>), which emerged in Spain in early summer, has spread to multiple locations in Europe<sup>2</sup>. Although this variant quickly grew into the dominant circulating SARS-CoV-2 strain in several countries, it did not appear to be associated with a higher intrinsic transmissibility<sup>2</sup>.

Although it appears clear that travel considerably contributed to the second wave in Europe, it remains challenging to assess how it may have restructured and reignited the epidemic in the different European countries. Even without resuming travel, relaxing containment measures when low-level transmission is ongoing risks the proliferation of locally circulating strains. Phylodynamic analyses may provide insights into the relative importance of persistence versus the introduction of new lineages, but such analyses are complicated for SARS-CoV-2 for different reasons. Phylogenetic reconstructions may be poorly resolved owing to the relatively limited SARS-CoV-2 sequence diversity<sup>9</sup>. This is further confounded by the degree of genetic mixing that can be expected from unrestricted travel before the lockdowns in spring 2020.

## Mobility data predicts SARS-CoV-2 spread

We analysed SARS-CoV-2 B.1 (20A) genomes from 10 European countries for which a minimal number of genomes from the second wave were already available on 3 November 2020. Using a two-step procedure that relied on subsampling relative to country-specific case counts (see [Methods](#)), we compiled a dataset of close to 4,000 genomes sampled between 29 January and 31 October 2020 (Extended Data Table 1). To achieve maximum resolution in our evolutionary reconstructions, we constructed a Bayesian time-measured phylogeographical model that integrates mobility and epidemiological data. Our approach simultaneously infers phylogenetic history and ancestral movement throughout this history while also identifying the drivers of spatial spread<sup>10</sup>. We used the latter functionality to determine the most appropriate mobility or connectivity measure. Specifically, we considered international air transportation data, the Google COVID-19 Aggregated Mobility Research Dataset (also referred to here as ‘mobility data’), and the social connectedness index of Facebook, as covariates of phylogeographical spread (Extended Data Fig. 1). The Google mobility dataset

contains anonymized mobility flows aggregated over users who have turned on the location history setting, which is off by default (see [Methods](#)). The social connectedness index reflects the structure of social networks and has been suggested to correlate with the geographical spread of COVID-19<sup>11</sup>. To help to inform the phylogenetic coalescent time distribution, we parameterized the viral population size trajectories through time as a function of epidemiological case count data for the countries under investigation.

Analyses using both time-homogeneous and time-inhomogeneous models offered strong support for mobility data as a predictor of spatial diffusion whereas air transportation data and the social connectedness index offered no predictive value (Extended Data Table 2). The fact that mobility data encompassing both air and land-based transport are required to explain COVID-19 spread highlights the need to consider both types of transport in containment strategies. To ensure that containment strategies were accommodated by our reconstructions, we further extended our time-inhomogeneous approach to model biweekly variation in the overall rate of spread between countries as a function of mobility (see [Methods](#) and Extended Data Table 2).

## Dynamic viral transmission through time

We use our probabilistic model of spatial spread informed by genomic, mobility and epidemiological data to characterize the dynamics of spread throughout the epidemic in Europe. We first focus on the ratio of introductions over the total viral flow in and out of each country over time and the genetic structure of country-specific transmission chains (Fig. 1). For the latter, we use a normalized entropy measure that quantifies the degree of phylogenetic interspersion of country-specific transmission chains in the SARS-CoV-2 phylogeny (see [Methods](#)). Although estimates for individual dispersal between pairs of countries can also be obtained (Extended Data Fig. 2), we remain cautious in interpreting these as direct pathways of spread because the genome sampling only covers a restricted set of European countries. The mobility to and from each country within our 10-country sample covers between 64% and 96% of the mobility of these countries to and from all countries within Europe (Extended Data Table 3 and Extended Data Fig. 3), except for Norway (27%), for which other Scandinavian countries account for considerable mobility connections (61%), and the UK (49%), for which Ireland accounts for a large fraction of mobility connections (38%).

**Fig. 1: Mobility, genome sampling, case counts and phylogeographical summaries through time for 10 European countries.**

---

 **figure1**

---

**a**, The country-specific Google mobility influx in the 10 countries during 2-week intervals. **b**, The weekly genome sampling by country used in the phylogeographical analysis. **c**, For each country, the ratio of introductions over the total viral flow from and to that country (in 2-week intervals) and a monthly normalized entropy measure summarizing the phylogenetic structure of country-specific transmission chains are shown. The posterior mean ratios of introductions are depicted with circles that have a size proportional to the total number of transitions from and to that country and the grey surface represents the 95% highest posterior density (HPD) intervals. The posterior mean normalized entropies and 95% HPD intervals are depicted by dotted lines. These normalized entropy measures indicate how phylogenetically structured the epidemic is in each country, and ranges from 0 (perfectly structured, for example, a single country-specific cluster) to 1 (unstructured interspersion of country-specific sequences across the entire SARS-CoV-2 phylogeny). The introduction ratios and normalized entropy measures are superimposed on the incidence of COVID-19 (daily

cases per  $10^6$  people) reported for each country through time (coloured density plot). The two vertical dashed lines represent the summer time interval (15 June and 15 August 2020) for which we subsequently evaluate introductions versus persistence (see Fig. 2).

[Full size image](#)

According to the proportion of introductions, we estimate more viral import than export events for Switzerland, Norway, the Netherlands and Belgium throughout most of the time period under investigation. According to the estimated phylogenetic entropy, these countries also experienced many independent transmission chains since the epidemic started to unfold. This is consistent with country-specific studies; for the first wave in Belgium, for example, about 331 individual introductions were estimated in the ancestry of a limited sample of 740 genomes<sup>12</sup>. For Portugal, we also estimate higher proportions of introductions early in the first wave but with a subsequent decline to predominantly export events. France, Italy and Spain, on the other hand, are characterized by a relatively high viral export during the first wave. The proportion of introductions remained relatively low for Italy and Spain after the first wave, whereas in France these proportions were high from mid-June until the end of July. However, the absolute number of transitions in our sample are low during this time period. These countries also had comparatively lower entropy values early in the epidemic, with an increase for France by the start of summer and a more gradual increase over time for Italy. In Spain, however, the genetic complexity of the SARS-CoV-2 transmission chains remained limited. In the UK and Germany, the viral flow in and out of the country was initially relatively balanced. A recent large-scale genomic analysis in the UK indicates that this can imply very high absolute numbers of cross-country transmissions, as more than 2,800 independent introduction events were identified from the analysis of 26,181 genomes<sup>13</sup>. Although our sample is limited compared to this UK-focused analysis<sup>13</sup>, our reconstructions also recover major influx from Spain, France and Italy during the first wave in the UK (Extended Data Fig. 2). We estimate an increase in the proportion of introductions for the UK from mid-June, indicating an important viral import relative to export around this time. The phylogenetic entropy also peaked around this time. In Germany, the proportions increased slightly later in the summer with a concomitant rise in phylogenetic entropy.

## Introductions thrive in low incidence

To assess the effect of summer travel on the second wave in the different countries, we use our genomic–mobility reconstruction to estimate both the number of lineages persisting in each country and the number of newly introduced lineages, and how these proliferated early in the second wave. We focus on a 2-month time period between 15 June 2020—when many EU and Schengen-area countries opened their borders to

other countries—and 15 August 2020, before which the majority of holiday return travel is expected for many countries. We identify the number of lineages circulating in each country on 15 August and determine whether they result from a lineage that persisted since 15 June or from a unique introduction after this date (independent of the number of descendants for this lineage on 15 August; Extended Data Fig. 4). In Fig. 2, we plot (1) the ratio of these unique introductions over the total unique lineages (unique introductions and persisting lineages); (2) the proportion of descendant lineages on 15 August that resulted from the unique introductions over the total descendants circulating on this date; and (3) the proportion of descendant tips (sampled genomes) after 15 August that resulted from the unique introductions over the total number of descendant tips (see [Methods](#) and Extended Data Fig. 4). We estimate a posterior mean proportion of unique introductions that is close to or higher than 0.5 except for Spain and Portugal. This indicates that by 15 August, a relatively large fraction of circulating lineages in each country was produced by new introductions over summer. Because the B.1.177/20E (EU1) variant that was predominantly disseminated through summer travel does not appear to be particularly more transmissible<sup>2</sup>, this is unlikely to be due to strong intrinsic advantages of the newly introduced viruses.

**Fig. 2: Posterior estimates for the relative importance of lineage introduction events in 10 European countries.**

 [figure2](#)

We report three summaries (posterior mean and 95% HPD intervals) for each country: the ratio of unique introductions over the total number of unique persisting lineages and unique introductions between 15 June and 15 August 2020 (1), the ratio of descendant lineages from these unique introduction events over the total number of descendants circulating on 15 August 2020 (2) and the ratio of descendant taxa from these unique introductions over the total number of descendant taxa sampled after 15 August 2020 (3) (see Extended Data Fig. 4). The dots are numbered and the sizes are proportional to: (1) the total number of unique lineage introductions identified between 15 June and 15 August 2020; (2) the total number of lineages inferred on 15 August 2020; and (3) the total number of descendant tips after 15 August 2020.

[Full size image](#)

The two proportions of descendants from these introductions on 15 August and after this date measure the relative success of newly introduced lineages compared to persisting lineages, indicating considerable variation in onward transmission. In Fig. 2, the country estimates are ordered according to decreasing average incidence during the 15 June–15 August time period, suggesting that incidence may shape the outcome of the introductions. In countries that experienced relatively high summer incidence (for example, Spain, Portugal, Belgium and France), the introductions lead to comparatively fewer descendants on 15 August or after. We find a significant overall association between the incidence and the difference in the logit-scaled proportion of unique introductions and the logit-scaled proportion of their descendants on 15 August ( $P = 0.007$ ) as well as between the incidence and the difference in the logit-scaled proportion of unique introductions and the logit-scaled proportion of descendant tips after 15 August ( $P = 0.019$ ) (Extended Data Fig. 5). With comparatively few descendants from introductions (Fig. 2), Norway may to some extent be an outlier because lineages estimated as persisting in this country could in fact be introductions from other Scandinavian countries that are not represented in our genome sample. We recover qualitatively similar, but more variable and statistically unsupported associations between the success of introductions and incidence for the 2-month time periods before and after the 15 June–15 August time period (Extended Data Fig. 5). This indicates that the comparatively higher proportion of introductions as well as the more stable and lower incidence between 15 June and 15 August provided the ideal conditions for a process of genetic drift by which introductions were able to fuel transmission.

Our estimates show that introductions in the UK particularly benefited from the conditions for successful onward transmission (Fig. 2), with a considerable fraction of introductions originating from Spain (Extended Data Fig. 6), reflecting the spread of B.1.177/20E (EU1), which rapidly became the most dominant strain in the UK<sup>2</sup>. Our analysis captures the expansion of this variant as well as that of B.1.160/20A.EU2, which together account for more than 25% of the genomes in our dataset. Although Spain was indeed inferred to be the origin of B.1.177/20E (EU1), the UK also considerably contributed to its spread (Fig. 3). The earliest introduction from Spain to the UK was estimated around the time Spain opened most EU borders (21 June) (Fig. 3). Although introductions from Spain to other countries soon followed, we estimate a similar rate and amount of spread from the UK to other countries before these other countries also further disseminated the virus. Although inferred from a limited sample, this illustrates a dynamic pattern of spread and the importance of the early establishment of B.1.177/20E (EU1) in the UK that probably served as an important secondary centre of dissemination. We note, however, that this pattern may be affected by the intensive and continuous genomic surveillance in the UK, which may also be reflected in our subsample of the available data. Although the UK is also involved in the spread of B1.160/20A.EU2, this variant has been largely disseminated from

France (Fig. 3). The fact that this variant expanded later in France and subsequently also started to spread later compared to B.1.177/20E (EU1) (Extended Data Fig. 7) may explain why the latter spread more successfully.

**Fig. 3: Phylogeographical estimates of SARS-CoV-2 spread in 10 European countries.**



**a**, The maximum clade credibility tree summary of the Bayesian inference. Colours correspond to the countries in the legend. The two clades corresponding to B1.160/20A.EU2 and B1.177/20E (EU1) are highlighted in grey. **b, c**, Circular migration flow plots for B1.160/20A.EU2 (**b**) and B1.177/20E (EU1) (**c**) based on the posterior expectations of the Markov jumps. In these plots, migration flow out of a particular location starts close to the outer ring and ends with an arrowhead more

distant from the destination location. **d**, Posterior mean estimates with 95% HPD intervals over time for four types of Markov jumps for B1.177/20E (EU1): from Spain to the UK, from Spain to other countries, from the UK and from other countries.

[Full size image](#)

## Discussion

Our Bayesian phylogeographical approach builds on a rich history of identifying drivers of spatial spread, with applications to various pathogens at different spatial scales, ranging from air transportation for influenza at a global scale<sup>10</sup> to gravity model transmission for Ebola in West Africa<sup>14</sup>. Such studies use a relatively limited genomic sample to gain insights into viral transmission dynamics. This is also the case in our application to SARS-CoV-2 in Europe for which we further extend the phylodynamic data integration approach to confront the lack of resolution offered by SARS-CoV-2 genomic data. A concerted effort in containing international spread further sets apart the COVID-19 pandemic from these earlier events. For this reason, we have now incorporated variation in mobility over time to account for the effect of these measures. Our reconstructions show that the composition of lineages circulating towards the end of the summer was to an important extent shaped by introductions in most of the European countries. The relative success of onward transmission of the introduced lineages appears to be shaped by the local incidence of COVID-19 during summer.

Our results should be interpreted in light of several important limitations. In addition to a limited overall size, the genome data only cover a selection of European countries, suggesting that we are missing transmission events that involve unsampled countries. This may be important for Norway, for example, which according to our mobility data, is largely connected to other Scandinavian countries. We also lack sampling from eastern Europe, which was to a large extent spared by border controls and lockdowns during the first wave, but witnessed high excess mortality rates during the second wave. The emergence of more transmissible variants has led to more intensified genomic surveillance, so similar phylodynamic reconstructions may now be performed on a wider scale.

The pandemic exit strategy offered by vaccination programmes is a source of optimism that also sparked proposals by EU member states to issue vaccine passports in a bid to revive travel and rekindle the economy. In addition to implementation challenges and issues of fairness, there are risks associated with such strategies when immunization is incomplete, as probably will be the case for the European population this summer. A recent modelling study for the UK suggests that vaccination in adults alone is unlikely to completely halt the spread of cases of COVID-19 and that lifting

containment measures early and suddenly can lead to a large wave of infections<sup>15</sup>. A gradual release of restrictions was shown to be critical for minimizing the infection burden<sup>15</sup>. We believe that travel policies may be a key consideration in this respect because similar conditions may arise to the ones that we demonstrated to provide fertile ground for viral dissemination and resurgence in 2020. This may now also involve the spread of variants that are more transmissible and/or evade the immune responses triggered by vaccines and previous infections. Well-coordinated European strategies will therefore be required to manage the spread of SARS-CoV-2 and reduce future waves of infection, with hopefully a more unified implementation than hitherto observed.

## 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.

### Sequence data and subsampling

We used a two-step genome data collection procedure. We first evaluated the available genomes from European countries in GISAID<sup>16</sup> on 3 November 2020. We selected genomes from Belgium, France, Germany, Italy, the Netherlands, Norway, Portugal, Spain, Switzerland and the UK primarily based on the availability of genome data from both the first and second wave at that time but also because of their high ratio of genomes to positive cases. A total of 39,812 genomes were available for these countries on 3 November 2020; the available number of genomes per country is listed in Extended Data Table 1. Portugal represented an exception because data for this country were limited to the first wave at that time, but we included genomes from Portugal because of its potential importance as a summer travel location.

We aligned the genomes from each country using MAFFT v.7.453<sup>17</sup> and trimmed the 5' and 3' ends and only retained unique sequences from each location. To further mitigate the disparities in sampling, we subsampled each country proportionally to the cumulative number of cases on 21 October 2020 (the most recently sampled sequence at the time) by setting an arbitrary threshold of 6.5 sequences per 10,000 cases, with a minimum number of 100 sequences per country. To maximize the temporal and spatial coverage in each country, we binned genomes by epi-week and sampled as evenly as possible, sampling from a different region within the country when available. Only sequences from the B.1 lineage with the D614G substitution and exact sampling dates

were selected for the analyses. From the final aligned sequence set, we removed 12 potential outliers, based on a root-to-tip regression applying TempEst v.1.5.3<sup>18</sup> to a maximum-likelihood tree inferred with IQTREE v.2.0.3<sup>19</sup>, yielding a dataset of 2,909 genomes (Extended Data Table 1).

Because of the nature of genome sequence accumulation, fewer recently sampled genomes were available for most countries on 3 November 2020 (relative to the case counts at this time). Because our primary goal was to assess the persistence and introduction of lineages leading up to the second wave, we sought to augment our dataset with more recent genomes, having already performed analyses on the initial dataset. In the section on Bayesian evolutionary reconstructions, we outline how we updated these analyses accordingly. On 5 January 2021, we updated our dataset by adding more than 1,000 non-identical sequences collected between 1 August and 31 October (out of a total of 56,395 available genomes; the available and selected numbers of genomes per country are listed in Extended Data Table 1). For Portugal, we extended this period back to 22 June (the most recent sampling date for the previous Portuguese selection). We downloaded all new B.1 sequences with the D614G substitution collected during the selected time period from GISAID and performed the following subsampling. The number of genomes to add per country was obtained by raising the threshold ratio of sequences/cases to 8.5 and increasing the minimum number of sequences to 200. To bias the temporal coverage towards more recent samples, the genomes from each country were binned by week and sampled such that the number of sequences added per week was proportional to an exponential function of the form  $e^{t/4}$ , where  $t = 0$  represents 1 August and  $t = 13$  is 31 October. For Portugal, we did not use this preferential sampling as we needed to include close to all available genomes to increase the number of genomes to 200. The selected sequences were deduplicated and outliers were removed as described in the previous paragraph. With the additional selection of 1,050 genomes, we obtained a dataset of 3,959 genomes (Extended Data Table 1).

## Mobility data

We analysed four different mobility and connectivity measures: air traffic flows, a social connectedness index provided by Facebook, as well as aggregated Facebook<sup>20</sup> and Google international mobility data. Air traffic flow data were obtained from the International Air Transport Association (<http://www.iata.org>) and based on the number of origin–destination tickets while also taking into account connections at intermediate airports<sup>21</sup>. We used monthly air traffic data between the 10 European countries under investigation for the time period between January 2020 and October 2020. The social connectedness index (SCI) is an anonymized snapshot of active Facebook users and their friendship networks to measure the intensity of social connectedness between countries (<https://data.humdata.org/>)<sup>22</sup>. In practice, the SCI measures the relative

probability of a Facebook friendship link between two users of the application in different countries. We used the SCI calculated for the 10 European countries represented in our genomic sample as of August 2020.

The Google COVID-19 Aggregated Mobility Research Dataset contains anonymized mobility flows aggregated over users who have turned on the location history setting (on a range of platforms<sup>23</sup>), which is off by default. To produce this dataset, machine learning is applied to logs data to automatically segment it into semantic trips<sup>24</sup>. To provide strong privacy guarantees, all trips were anonymized and aggregated using a differentially private mechanism<sup>25</sup> to aggregate flows over time (see <https://policies.google.com/technologies/anonymization>). This research was done on the resulting heavily aggregated and differentially private data. No individual user data was ever manually inspected, only heavily aggregated flows of large populations were handled. All anonymized trips were processed in aggregate to extract their origin and destination location and time. For example, if users travelled from location  $a$  to location  $b$  within time interval  $t$ , the corresponding cell  $(a, b, t)$  in the tensor would be  $n \pm \eta$ , where  $\eta$  is Laplacian noise. The automated Laplace mechanism adds random noise drawn from a zero-mean Laplace distribution and yields  $(\varepsilon, \delta)$ -differential privacy guarantee of  $\varepsilon = 0.66$  and  $\delta = 2.1 \times 10^{-29}$  per metric. Specifically, for each week  $W$  and each location pair  $(A, B)$ , we compute the number of unique users who took a trip from location  $A$  to location  $B$  during week  $W$ . To each of these metrics, we add Laplace noise from a zero-mean distribution of scale  $1/0.66$ . The parameter  $\varepsilon$  controls the noise intensity in terms of its variance and  $\delta$  represents the deviation from pure  $\varepsilon$  privacy. The closer these parameters are to zero, the stronger the privacy guarantees. We used aggregated mobility flows between the 10 European countries and summarized them by 2-week or monthly time periods between January 2020 and October 2020.

Finally, we also considered international mobility data from Facebook mobility data as an alternative to Google mobility data. These data are based on the numbers of Facebook users moving over large distances, such as air or train travel. Counts of international travel patterns are updated daily based only on users who have opted to share precise location data from their device with the Facebook mobile app through location services. Also in this case, we used aggregated mobility flows between the 10 European countries and summarized them by month between January 2020 and October 2020. Because international aggregate mobility data obtained from Google and Facebook are highly correlated (monthly Spearman correlation ranging from 0.84 to 0.92) (Supplementary Fig. 1), we only included the Google aggregate mobility data as a covariate in the phylogeographical analyses. We note that the mobility data are subject to limitations as these may not be representative of the population as whole and their representativeness may vary by location.

## Bayesian evolutionary reconstructions

### Joint sequence–trait inference with a time-homogeneous generalized linear model of discrete trait diffusion

We performed a Bayesian evolutionary reconstruction of timed phylogeographical history using BEAST 1.10<sup>26</sup>, incorporating the genome sequences, their country and date of sampling, epidemiological and mobility and/or connectivity data. Because of the relatively low degree of resolution offered by the sequence data, our full probabilistic model specification focuses on (1) relatively simple model specifications and (2) informing parameters by additional non-genetic data sources. We modelled sequence evolution using an HKY85 nucleotide substitution model with a gamma-distributed rate variation among sites and a strict molecular clock model. Our genome set includes three genomes from an early outbreak in Bavaria, which was caused by an independent introduction from China<sup>27,28</sup>. We therefore constrained these genomes as an outgroup in the analysis, which according to root-to-tip regression plots as a function of sampling time resulted in a better correlation coefficient and  $R^2$  compared to the best-fitting root under the heuristic mean residual squared criterion<sup>18</sup> (Supplementary Fig. 2).

As a coalescent tree prior, we modelled the effective population size trajectory as a piecewise constant function that changes values at pre-specified times (following a previously published study<sup>29</sup>), with log-transformed population sizes modelled as a deterministic function of log-transformed counts of cases of COVID-19 (following a previous publication<sup>30</sup>). This reduces the nonparametric skygrid parameterization to a generalized linear model (GLM) formulation with an estimable regression intercept ( $\alpha$ ) and coefficient ( $\beta$ ). In this parameterization, a coefficient estimate centred around 0 would imply constant population size dynamics through time. We specified 2-week intervals and summarized as a covariate the total case counts over these time intervals for the 10 countries of sampling (obtained from <https://www.ecdc.europa.eu/en/covid-19/data>). The earliest interval with non-zero cases counts was from 14 January 2020 to 28 January 2020; before 14 January 2020, the log-transformed and standardized case count covariate was set to the equivalent of 1 case. We also tested whether a lag time was required for the case count covariate using marginal likelihood estimation<sup>31</sup>. Specifically, we shifted the case counts by 1, 2, 3 and 4 weeks before summarizing them according to 2-week intervals and estimated the model fit of these covariates against case counts without lag time (Supplementary Table 1). To mitigate the computational burden associated with the marginal likelihood estimation procedure, we performed these analyses on a subset of 1,000 genomes (obtained using the Phylogenetic Diversity Analyzer tool<sup>32</sup>). We estimated the highest (log-transformed)

marginal likelihood for a two-week lag time (Supplementary Table 1) and used this for the case count covariate in our analyses.

Similar to sequence evolution, we modelled the process of transitioning through discrete location states (countries of sampling) according to a continuous-time Markov chain (CTMC)<sup>33</sup>. We used a parameterization that models the log-transformed transition rates as a log-linear function of mobility and connectivity covariates<sup>10</sup>. The Bayesian implementation of this model simultaneously estimates the phylogenetic history, ancestral movement and the contribution of covariates to the movement patterns<sup>10</sup>. Although we mainly use this approach to obtain well-informed phylodynamic estimates, we also make use of its capacity to identify the most-relevant mobility measure to inform our reconstructions. As covariates we considered the SCI of Facebook, air transportation data and mobility data. For the two time-variable mobility measures, we used the average of the log-transformed and standardized monthly mobility measures as a single covariate in our time-homogeneous phylogeographical GLM model. In this GLM formulation, we estimate the positive effect sizes for each covariate as well as their inclusion probability through a spike-and-slab procedure<sup>10</sup>. Although we subsampled the number of SARS-CoV-2 genomes per country in proportion to case counts, they do not fully correspond because we used a minimum number of genomes for countries with low case counts. We therefore evaluated whether this resulted in signal for sampling bias by including an origin and destination covariate in the GLM based on the residuals for a regression analysis between genomes and case counts (following a previously published study<sup>14</sup>). We performed this analysis using a set of empirical trees (see ‘Time-inhomogeneous reconstructions’) applying both a time-homogeneous and time-inhomogeneous model, but found no support for these additional covariates (Supplementary Table 2).

We performed inference under the full model specification using Markov chain Monte Carlo (MCMC) sampling and used the BEAGLE library v.3<sup>34</sup> to increase computational performance. We specified standard transition kernels on all parameters, except for the regression coefficients of the piecewise-constant coalescent GLM model. For these parameters, we implemented new Hamiltonian Monte Carlo transition kernels to improve sampling efficiency. These kernels use principles from Hamiltonian dynamics and their approximate energy conserving properties to reduce correlation between successive sampled states, but require computation of the gradient of the model log-posterior with respect to the parameters of interest, in addition to efficient evaluation of the log-posterior that BEAGLE provides. To accomplish this, we extended our previous analytic derivation of the gradient of the log-transformed density from the skygrid coalescent model with respect to the log-transformed population sizes<sup>35</sup> to now be with respect to the regression coefficients using the chain rule and their regression design matrix.

Owing to the dataset size, MCMC burn-in takes up considerable computational time. We therefore iterated through a series of BEAST inferences, initially only considering sequence evolution and subsequently adding the location data, to arrive at a tree distribution from which trees were taken as starting trees in our final analyses. The latter was composed of multiple independent MCMC runs that were run sufficiently long to ensure that their combined posterior samples achieved effective sample sizes larger than 100 for all continuous parameters.

## Data augmentation through online BEAST

As we updated our dataset after the initial analyses of the 2,909 genome collection using the approach discussed (see ‘Bayesian evolutionary reconstructions’), we sought to capitalize on these efforts to limit the burn-in for subsequent analyses of the 3,959 dataset. Specifically, we adopted the distance-based procedure to insert new taxa into a time-measured phylogenetic tree sample as implemented in the BEAST framework for online inference<sup>36</sup>. We subsequently use the augmented tree as the starting tree for the analyses of the updated dataset.

## Time-inhomogeneous reconstructions

To accommodate the time variability of the mobility measures, we constructed epoch model extensions of the discrete phylogeography approach that allow specifying arbitrary intervals over the evolutionary history and associating them with different model parameterizations<sup>37</sup>. As a complement to testing covariates of spatial diffusion using a time-homogeneous model, we used the epoch extension to specify monthly intervals, enabling us to incorporate monthly mobility matrices (air transportation data were available only as monthly numbers), but assuming time-homogeneous effect sizes and inclusion probabilities. Monthly covariates were again log-transformed and standardized after adding a pseudo-count to each entry in the monthly matrices.

In addition, we performed another analysis in which we relaxed the constant-through-time inclusion probability of the covariates. In this model specification, each interval is associated with a specific set of indicator variables to represent the inclusion or exclusion of covariates, but we pool information about predictor inclusion across the intervals using hierarchical graph modelling<sup>38</sup>. This approach uses a set of indicator variables to model covariate inclusion at the hierarchical level but enables interval-specific inclusion or predictors to diverge from the hierarchical level with a non-zero probability (with the number of differences modelled as a binomial distribution<sup>38</sup>), which was set to 0.10 in our study. We estimated hierarchical and interval-level inclusion using the spike-and-slab procedure<sup>38</sup>.

Finally, we performed an analysis using the time-inhomogeneous model in which the interval-specific transition rates are modelled as a function of the single covariate that is supported by the analyses above leveraging aggregate mobility. We incorporated more variability through time by specifying 2-week intervals (similar to the coalescent GLM interval specification). In addition, we add time-homogeneous random effects to the phylogeographical transition rate parameterization to account for potential biases in the ability of mobility to predict phylogeographical spread. Although the posterior mean estimates for these random effects vary, only very few indicate that individual phylogeographical transition rates significantly deviate from the mobility data (Supplementary Fig. 3). The time-inhomogeneous GLM approach that we use enables the modelling of relative differences in transition rates, but also the overall rate of migration between countries varies through time, and importantly, this is strongly affected by intervention strategies. To accommodate these dynamics, we further extend this model by incorporating a time-inhomogeneous overall CTMC rate scalar and parameterize it as a log-linear function of the total monthly between-country log-transformed and standardized mobility (time-variable rate scalar GLM in Extended Data Table 2). To generate realizations of the discrete location CTMC process and obtain estimates of the transitions (Markov jumps) between states under this model, we used posterior inference of the complete Markov jump history through time<sup>10,39</sup>.

Although the epoch model enables us to flexibly accommodate time-variable spatial dynamics, it considerably increases the computational burden associated with likelihood evaluations. To efficiently draw inferences under this model for our large dataset, we fit the time-inhomogeneous spatial diffusion process to a set of trees inferred under the time-homogeneous GLM diffusion model described above.

Although likelihood evaluations remain computationally expensive, even with the speed-up offered by GPU computation with BEAGLE, eliminating simultaneous tree estimation tremendously reduces the parameter space, requiring only modest MCMC chain lengths to adequately explore it. Model and inference specifications for the different analyses are available as BEAST XML input files on GitHub ([https://github.com/phylogeography/SARS-CoV-2\\_EUR\\_PHYLOGEOGRAPHY](https://github.com/phylogeography/SARS-CoV-2_EUR_PHYLOGEOGRAPHY)) and Zenodo (<https://doi.org/10.5281/zenodo.4876442>).

## Posterior summaries

We assessed MCMC mixing (for example, using effective sample sizes) and summarized continuous parameter estimates using Tracer v.1.7.1<sup>40</sup>. Credible intervals were computed as 95% HPD intervals. Trees were visualized using FigTree v.1.4.4 (available at <https://github.com/rambaut/figtree/releases>). In terms of phylogeographical estimates, we mainly focused on (1) transitions to each location and from each location (based on Markov jump estimates) instead of pairwise transitions; (2) ratios of these transitions and (3) how these transitions structured transmission

chains in individual countries. Transitions to and from each location avoid drawing conclusions about direct migration between countries, which can be tenuous given the incomplete genome coverage of Europe, while their ratios avoid using absolute numbers of transitions, which are highly sample-dependent. Phylogeographical inference is limited to reconstructing the transitions in the ancestral history of a sample of sequences, which will only be a small fraction of the actual migration events especially when these events result in insufficient onward transmission to be captured in our limited sample. In addition, SARS-CoV-2 genome data can be poorly resolved and identical genomes in different locations are consistent with hypotheses that involve both a sparse and a rich number of virus flows between these locations. As the data hold little information to distinguish these hypotheses, we only consider sparse scenarios by including only unique sequences for each location. A joint inference of sequence evolution and discrete spatial diffusion would err on the side of sparse hypotheses anyway because it will tend to cluster identical sequences that share a location. Despite the general underestimation of spatial dispersal, a phylogeographical inference is still likely to capture the transition events with important onward transmission, and evaluating the importance of such events relative to persistence is a major focus of this study. Cryptic transmission also complicates the ability to reconstruct spatial dispersal, but we expect this to be equally likely for introductions and persistence and therefore focus on their ratio for each location.

We provide three new tree sample tools in the BEAST codebase on GitHub (<https://github.com/beast-dev/beast-mcmc>) to obtain posterior summaries of location transition histories using posterior tree distributions annotated with Markov jumps:

- (1) TreeMarkovJumpHistoryAnalyzer enables the collection of Markov jumps and their timings from a posterior tree distribution annotated with Markov jumps histories in a .csv file for further analyses.
- (2) TreeStateTimeSummarizer decomposes the total tree time into the times associated with contiguous partitions of a tree estimated to be in a particular location state, with the partitions determined by the Markov jumps. An arbitrary lower- and upper-time boundary can be specified to restrict the summary to a particular time interval in the evolutionary history. We use the time estimates for the separate partitions associated with each state to calculate an entropy measure that summarizes the genetic make-up of country-specific transmission chains. Specifically, we use for each location a normalized Shannon entropy:

$$\text{Equation (1)}: \text{Shannon Entropy} = -\frac{1}{\ln(n)} \sum_{i=1}^n p_i \ln(p_i)$$

where  $p_i$  is the proportion of time associated with that location for partition  $i$  of a phylogeographical tree and  $n$  represents the number of partitions for that location in the tree.

(3) PersistenceSummarizer also uses posterior tree distributions annotated with Markov jumps to summarize the number of lineages at a particular point in time (evaluation time ( $T_e$ ); see Extended Data Fig. 5), which location states they are associated with, since what time point in the past they have maintained that state and how many sampled descendants they have after time  $T_e$  (Extended Data Fig. 5). In addition, it enables summarizing how long these lineages have circulated independently before  $T_e$ , so before sharing common ancestry with other lineages that maintained the same location state. This information allows us to determine how many lineages are circulating at  $T_e$  that stem either from a unique persistent lineage (maintaining the same location states) or unique introduction event since a particular time before  $T_e$  (ancestral time ( $T_a$ ) in Extended Data Fig. 5). The association between incidence and the difference in the logit proportion of unique introductions and the logit proportion of their descendants on 15 August was evaluated using a  $P$  value obtained by a linear regression analysis.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

BEAST XML input files are available on GitHub ([https://github.com/phylogeography/SARS-CoV-2\\_EUR\\_PHYLOGEOGRAPHY](https://github.com/phylogeography/SARS-CoV-2_EUR_PHYLOGEOGRAPHY)) and Zenodo (<https://doi.org/10.5281/zenodo.4876442>). The SARS-CoV-2 genome data required to run these XML files can be downloaded from <https://www.gisaid.org>; all GISAID accession numbers are listed in the GISAID acknowledgements table (Supplementary Table 3).

The Google COVID-19 Aggregated Mobility Research Dataset and the Facebook mobility data are not publicly available owing to stringent licensing agreements. Information on the process of requesting access to the Google mobility data are available from A.S. (sadilekadam@google.com) and the COVID-19 Community Mobility Reports that were derived from the Google data are publicly available at <https://www.google.com/covid19/mobility/>. The Facebook mobility data are made available through the Data for Good programme (<https://dataforgood.fb.com>) under the terms of a data license agreement that defines the allowed terms of use by partners

(contact: [disastermaps@fb.com](mailto:disastermaps@fb.com)). Once a request for access from a partner institution is vetted and an appropriate data license agreement is signed, then access is granted through Facebook's web-based spatial visualization tool called GeoInsight. Air travel data were obtained from the International Air Transport Association (<http://www.iata.org>).

log-transformed and standardized among-country mobility and air travel data are specified in the available BEAST XML files ([https://github.com/phylogeography/SARS-CoV-2\\_EUR\\_PHYLOGEOGRAPHY](https://github.com/phylogeography/SARS-CoV-2_EUR_PHYLOGEOGRAPHY) and <https://doi.org/10.5281/zenodo.4876442>). COVID-19 incidence data were obtained from <https://www.ecdc.europa.eu/en/covid-19/data>.

## Code availability

The code to run the BEAST analyses is available in the hmc\_develop branch of the BEAST codebase on GitHub (<https://github.com/beast-dev/beast-mcmc>) and Zenodo (<https://doi.org/10.5281/zenodo.4895235>). The tools TreeMarkovJumpHistoryAnalyzer, TreeStateTimeSummarizer and PersistenceSummarizer are available from the master branch in the same codebase.

## References

1. 1.  
European Centre for Disease Prevention and Control. *Data on 14-Day Notification Rate of New COVID-19 Cases and Deaths* (2021); <https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19>
2. 2.  
Hodcroft, E. B. et al. Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* <https://doi.org/10.1038/s41586-021-03677-y>. (2021).
3. 3.  
European Centre for Disease Prevention and Control. *COVID-19 Situation Update for the EU/EEA, as of Week 3, Updated 28 January 2021* (2021); <https://www.ecdc.europa.eu/en/cases-2019-ncov-eueea>
4. 4.

Rambaut, A. et al. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. *Virological* <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563> (2020).

5. 5.

Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BMC Med.* **18**, 240 (2020).

[Article](#) [Google Scholar](#)

6. 6.

Neher, R. A., Dyrdak, R., Druelle, V., Hodcroft, E. B. & Albert, J. Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. *Swiss Med. Wkly.* **150**, w20224 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

McKee, M. A European roadmap out of the COVID-19 pandemic. *Br. Med. J.* **369**, m1556 (2020).

[Article](#) [Google Scholar](#)

8. 8.

Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Morel, B. et al. Phylogenetic analysis of SARS-CoV-2 data is difficult. *Mol. Biol. Evol.* **38**, 1777–1791 (2020).

[Article](#) [Google Scholar](#)

10. 10.

Lemey, P. et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog.* **10**, e1003932 (2014).

[Article](#) [Google Scholar](#)

11. 11.

Kuchler, T., Russel, D. & Stroebel, J. *The Geographic Spread of COVID-19 Correlates with the Structure of Social Networks as Measured by Facebook*, NBER Working Paper 26990 (National Bureau of Economic Research, 2020).

12. 12.

Dellicour, S. et al. A phylodynamic workflow to rapidly gain insights into the dispersal history and dynamics of SARS-CoV-2 lineages. *Mol. Biol. Evol.* **38**, 1608–1613 (2021).

[Article](#) [Google Scholar](#)

13. 13.

du Plessis, L. et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **371**, 708–712 (2021).

[ADS](#) [Article](#) [Google Scholar](#)

14. 14.

Dudas, G. et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**, 309–315 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Moore, S., Hill, E. M., Tildesley, M. J., Dyson, L. & Keeling, M. J. Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **21**, 793–802 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Shu, Y. & McCauley, J. GISAID: global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* **22**, 30494 (2017).

[Article](#) [Google Scholar](#)

17. 17.

Katoh, K., Asimenos, G. & Toh, H. Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* **537**, 39–64 (2009).

[CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016).

19. 19.

Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Maas, P. Facebook disaster maps: aggregate insights for crisis response & recovery. In *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (ACM, 2019).

21. 21.

Gilbert, M. et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *Lancet* **395**, 871–877 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J. & Wong, A. Social connectedness: measurement, determinants, and effects. *J. Econ. Perspect.* **32**, 259–280 (2018).

[Article](#) [Google Scholar](#)

23. 23.

Kraemer, M. U. G. et al. Mapping global variation in human mobility. *Nat. Hum. Behav.* **4**, 800–810 (2020).

[Article](#) [Google Scholar](#)

24. 24.

Bassolas, A. et al. Hierarchical organization of urban mobility and its connection with city livability. *Nat. Commun.* **10**, 4817 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

25. 25.

Wilson, R. J. et al. Differentially private SQL with bounded user contribution. *Proc. Priv. Enhanc. Technol.* **2020**, 230–250 (2020).

[Google Scholar](#)

26. 26.

Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).

[Article](#) [Google Scholar](#)

27. 27.

Böhmer, M. M. et al. Investigation of a COVID-19 outbreak in Germany resulting from a single travel-associated primary case: a case series. *Lancet Infect. Dis.* **20**, 920–928 (2020).

[Article](#) [Google Scholar](#)

28. 28.

Worobey, M. et al. The emergence of SARS-CoV-2 in Europe and North America. *Science* **370**, 564–570 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Gill, M. S. et al. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* **30**, 713–724 (2013).

[CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Faria, N. R. et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. *Science* **361**, 894–899 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Baele, G., Lemey, P. & Suchard, M. A. Genealogical working distributions for Bayesian model testing with phylogenetic uncertainty. *Syst. Biol.* **65**, 250–264 (2016).

[Article](#) [Google Scholar](#)

32. 32.

Chernomor, O. et al. Split diversity in constrained conservation prioritization using integer linear programming. *Methods Ecol. Evol.* **6**, 83–91 (2015).

[Article](#) [Google Scholar](#)

33. 33.

Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots. *PLOS Comput. Biol.* **5**, e1000520 (2009).

[ADS](#) [MathSciNet](#) [Article](#) [Google Scholar](#)

34. 34.

Ayres, D. L. et al. BEAGLE 3: improved performance, scaling, and usability for a high-performance computing library for statistical phylogenetics. *Syst. Biol.* **68**, 1052–1061 (2019).

[Article](#) [Google Scholar](#)

35. 35.

Baele, G., Gill, M. S., Lemey, P. & Suchard, M. A. Hamiltonian Monte Carlo sampling to estimate past population dynamics using the skygrid coalescent model in a Bayesian phylogenetics framework. *Wellcome Open Res.* **5**, 53 (2020).

[Article](#) [Google Scholar](#)

36. 36.

Gill, M. S., Lemey, P., Suchard, M. A., Rambaut, A. & Baele, G. Online Bayesian phylodynamic inference in BEAST with application to epidemic reconstruction. *Mol. Biol. Evol.* **37**, 1832–1842 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

37. 37.

Bielejec, F., Lemey, P., Baele, G., Rambaut, A. & Suchard, M. A. Inferring heterogeneous evolutionary processes through time: from sequence substitution to phylogeography. *Syst. Biol.* **63**, 493–504 (2014).

[Article](#) [Google Scholar](#)

38. 38.

Cybis, G. B., Sinsheimer, J. S., Lemey, P. & Suchard, M. A. Graph hierarchies for phylogeography. *Phil. Trans. R. Soc. Lond. B* **368**, 20120206 (2013).

[Article](#) [Google Scholar](#)

39. 39.

Minin, V. N. & Suchard, M. A. Fast, accurate and simulation-free stochastic mapping. *Phil. Trans. R. Soc. Lond. B* **363**, 3985–3995 (2008).

[Article](#) [Google Scholar](#)

40. 40.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank all of the authors who have shared genome data on GISAID, and we have included a table (Supplementary Table 3) acknowledging the authors and institutes involved. The research leading to these results has received funding from the European Research Council under the European Union’s Horizon 2020 research and innovation programme (725422-ReservoirDOCS) and the Bill & Melinda Gates Foundation (OPP1094793 and INV-024911). This study was partially funded by EU grant 874850 MOOD and is catalogued as MOOD 005. The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. The Artic Network receives funding from the Wellcome Trust through project 206298/Z/17/Z. P.L. acknowledges support by the Research Foundation–Flanders ('Fonds voor Wetenschappelijk Onderzoek–Vlaanderen', G066215N, G0D5117N and G0B9317N). G.B. acknowledges support from the 'Interne Fondsen KU Leuven'/Internal Funds KU Leuven under grant agreement no. C14/18/094, and the Research Foundation–Flanders ('Fonds voor Wetenschappelijk Onderzoek–Vlaanderen', G0E1420N and G098321N). M.A.S. acknowledges support from National Institutes of Health grants U19 AI135995 and R01 AI153044. S.D. is supported by the Fonds National de la Recherche Scientifique (FNRS, Belgium). We acknowledge support from NVIDIA Corporation through the donation of parallel computing resources used for this research and thank AMD for the donation of critical hardware and support resources from its HPC Fund that made this work possible.

## Author information

### Affiliations

1. Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium

Philippe Lemey, Samuel L. Hong, Frederik Van den Broeck, Mandev S. Gill, Guy Baele & Simon Dellicour

2. Global Virus Network (GVN), Baltimore, MD, USA

Philippe Lemey

3. WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton, UK

Nick Ruktanonchai, Shengjie Lai & Andrew J. Tatem

4. Population Health Sciences, Virginia Tech, Blacksburg, VA, USA

Nick Ruktanonchai

5. INSERM, Sorbonne Université, Institut Pierre Louis d'Epidémiologie et de Santé Publique IPLESP, Paris, France

Vittoria Colizza & Chiara Poletto

6. Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

Frederik Van den Broeck

7. Department of Mathematics, School of Science & Engineering, Tulane University, New Orleans, LA, USA

Xiang Ji

8. UMR MEGH (Microbes, Evolution, Phylogeny and Infections), Aix-Marseille Université (AMU) and Institut Universitaire de France (IUF), Marseille, France

Anthony Levasseur

9. Department of Viroscience, WHO Collaborating Centre for Arbovirus and Viral Hemorrhagic Fever Reference and Research, Erasmus MC, Rotterdam, The Netherlands

Bas B. Oude Munnink & Marion Koopmans

10. Google, Mountain View, CA, USA

Adam Sadilek

11. Department of Biomathematics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

Marc A. Suchard

12. Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA, USA

Marc A. Suchard

13. Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

Marc A. Suchard

14. Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, Bruxelles, Belgium

Simon Dellicour

#### Authors

1. Philippe Lemey

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Nick Ruktanonchai

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Samuel L. Hong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Vittoria Colizza

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Chiara Poletto

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Frederik Van den Broeck

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Mandev S. Gill

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Xiang Ji

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Anthony Levasseur

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Bas B. Oude Munnink

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Marion Koopmans

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Adam Sadilek

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Shengjie Lai

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Andrew J. Tatem

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Guy Baele

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Marc A. Suchard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Simon Dellicour

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

P.L. and S.D. designed the study, performed analyses and drafted the manuscript. V.C., C.P. and A.S. provided and analysed data. S.L.H., F.V.d.B., N.R., S.L. and A.J.T. compiled and analysed data. A.L., B.B.O.M. and M.K. contributed data. G.B. performed data analyses. M.S.G., X.J. and M.A.S. developed statistical inference methodology. All authors contributed to interpreting and reviewing the manuscript.

## Corresponding authors

Correspondence to [Philippe Lemey](#) or [Simon Dellicour](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Jason Thomas Ladner, Matthew Scotch and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review 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 Monthly international mobility data matrices for air traffic and Google and Facebook mobility data.](#)

**a–c**, International air traffic data (**a**), international Facebook mobility data (**b**), and international Google mobility data (**c**). For Facebook data (**b**), we also report the single SCI matrix.

**Extended Data Fig. 2 Estimated introductions through time in the 10 European countries and circular migration flow plots summarizing the estimated transitions between the countries for different time intervals throughout the evolutionary history of SARS-CoV-2.**

**a**, The introductions through time serve as an illustration and are based on the Markov jump history in the maximum clade credibility tree. We note that the posterior distribution of trees is accompanied by considerable uncertainty about the location of origin, destination and timing of the transitions that is difficult to appropriately visualize. The grey box represents the time period from 15 June to 15 August 2020. **b**, The circular migration flow plots are based on the posterior expectations of the Markov jumps. The sizes of the plots reflect the total number of transitions for each period. In these plots, migration flow out of a particular location starts close to the outer ring and ends with an arrowhead more distant from the destination location.

**Extended Data Fig. 3 Pairwise mobility data among the 10 countries included in the phylogeographical analysis and other European countries.**

Heat-map cells are coloured according to international Google mobility data for the time period between January and October 2020.

**Extended Data Fig. 4 Conceptual representation of persistent lineages and introductions during the time interval delineated by the evaluation time and the ancestral time.**

At evaluation time ( $T_e$ ), we evaluate how many lineages are circulating in the location of interest; in this case, 12 lineages (lineages in other locations are represented by thick grey branches). We subsequently identify whether these lineages maintained this location up to ancestral time ( $T_a$ ) in their ancestry or whether they result from an introduction event in the time interval of interest. By determining whether other lineages circulating in the location of interest at  $T_e$  are descendants of the same persistent lineage or whether they share an introduction event, we identify the unique persistent lineages or introductions, in this case 2 and 4 lineages, respectively. In addition to the proportion of unique introductions ( $p_1 = 4/6$ ), we also summarize the proportion of their descendants at  $T_e$  ( $p_2 = 9/(9 + 3)$  in this case) and the proportion of

their descendants in terms of sampled tips after  $T_e$  ( $p_3$ ). Those tips are not shown here but are conceptually represented for both introductions and persistent lineages by ovals.

**Extended Data Fig. 5 Scatter plots of the difference in the logit proportion of unique introductions and the logit proportion of their descendants on 15 August against the incidence and the difference in the logit proportion of unique introductions and the logit proportion of descendant tips after 15 August against incidence.**

Left, the difference in logit proportions of unique introductions ( $p_1$ ) and their descendants ( $p_2$ ). Right, the difference in logit proportions of unique introductions ( $p_1$ ) and descendant tips ( $p_3$ ). Data are shown for the periods between 15 April and 15 June, between 15 June and 15 August, and between 15 August and 15 October. The  $P$  values in the lower right corner of the plots are the p-values for the hypothesis tests based on the t-statistic evaluating whether the regression coefficient in a linear regression model is different from 0.

**Extended Data Fig. 6 Estimated geographical origin of viral influx over the summer (15 June–15 August 2020) in each country.**

Each bar plot summarizes the posterior Markov jump estimates into a specific country. For the bar representing a low number of introductions into Portugal, a magnified view is provided.

**Extended Data Fig. 7 Phylogeographical transitions for lineages B1.177/20E (EU1) and B1.160/20A.EU2.**

Cumulative phylogeographical transitions are summarized as posterior mean estimates with 95% HPD intervals over time for four types of Markov jumps. For B1.177/20E (EU1), jumps from Spain to the UK, from Spain to other countries, from the UK and from other countries are included. For B1.160/20A.EU2 jumps from France to the UK, from France to other countries, from the UK and from other countries are included.

**Extended Data Table 1 Genome sampling by country, collected on 3 November 2020 and updated on 5 January 2021**

[Full size table](#)

**Extended Data Table 2 Parameter estimates for the various Bayesian time-measured phylogeographical models**

[Full size table](#)

**Extended Data Table 3 Mobility percentage to or from each country within our 10-country sample**

[Full size table](#)

## Supplementary information

### Supplementary Information

This file contains Supplementary Figures 1-3 and Supplementary Tables 1-2.

### Reporting Summary

### Supplementary Table 3

GISAID acknowledgements table including accession ID, originating laboratory, submitting laboratory and authors.

### Peer Review File

## Rights and permissions

### Reprints and Permissions

## About this article



Check for updates

## Cite this article

Lemey, P., Ruktanonchai, N., Hong, S.L. *et al.* Untangling introductions and persistence in COVID-19 resurgence in Europe. *Nature* **595**, 713–717 (2021).  
<https://doi.org/10.1038/s41586-021-03754-2>

### Download citation

- Received: 04 February 2021
- Accepted: 22 June 2021

- Published: 30 June 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03754-2>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03754-2>

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

---

Nasal delivery of an IgM offers broad protection from SARS-CoV-2 variants

[Download PDF](#)

- Article
- [Published: 03 June 2021](#)

# Nasal delivery of an IgM offers broad protection from SARS-CoV-2 variants

- [Zhiqiang Ku](#) ORCID: [orcid.org/0000-0001-8347-0013](#)<sup>1 na1</sup>,
- [Xuping Xie](#) ORCID: [orcid.org/0000-0003-0918-016X](#)<sup>2 na1</sup>,
- [Paul R. Hinton](#) ORCID: [orcid.org/0000-0002-9554-4144](#)<sup>3 na1</sup>,
- [Xinli Liu](#)<sup>4 na1</sup>,
- [Xiaohua Ye](#)<sup>1</sup>,
- [Antonio E. Muruato](#)<sup>5,6</sup>,
- [Dean C. Ng](#)<sup>3</sup>,
- [Sujit Biswas](#)<sup>4</sup>,
- [Jing Zou](#)<sup>2</sup>,
- [Yang Liu](#)<sup>2</sup>,
- [Deepal Pandya](#)<sup>3</sup>,
- [Vineet D. Menachery](#) ORCID: [orcid.org/0000-0001-8803-7606](#)<sup>6</sup>,
- [Sachi Rahman](#)<sup>3</sup>,
- [Yu-An Cao](#)<sup>3</sup>,
- [Hui Deng](#)<sup>1</sup>,
- [Wei Xiong](#)<sup>1</sup>,
- [Kevin B. Carlin](#)<sup>3</sup>,
- [Junquan Liu](#)<sup>1</sup>,
- [Hang Su](#)<sup>1</sup>,
- [Elizabeth J. Haanes](#)<sup>3</sup>,
- [Bruce A. Keyt](#) ORCID: [orcid.org/0000-0002-3508-9223](#)<sup>3</sup>,
- [Ningyan Zhang](#) ORCID: [orcid.org/0000-0002-4348-2180](#)<sup>1</sup>,
- [Stephen F. Carroll](#) ORCID: [orcid.org/0000-0001-7366-9918](#)<sup>3</sup>,
- [Pei-Yong Shi](#) ORCID: [orcid.org/0000-0001-5553-1616](#)<sup>2,5,7,8,9</sup> &
- [Zhiqiang An](#) ORCID: [orcid.org/0000-0001-9309-2335](#)<sup>1</sup>

[Nature](#) volume 595, pages 718–723 (2021) [Cite this article](#)

- 36k Accesses
- 783 Altmetric
- [Metrics details](#)

## Subjects

- [Antibodies](#)
- [Antibody therapy](#)
- [Mucosal immunology](#)
- [SARS-CoV-2](#)

## Abstract

Resistance represents a major challenge for antibody-based therapy for COVID-19<sup>1,2,3,4</sup>. Here we engineered an immunoglobulin M (IgM) neutralizing antibody (IgM-14) to overcome the resistance encountered by immunoglobulin G (IgG)-based therapeutics. IgM-14 is over 230-fold more potent than its parental IgG-14 in neutralizing SARS-CoV-2. IgM-14 potently neutralizes the resistant virus raised by its corresponding IgG-14, three variants of concern—B.1.1.7 (Alpha, which first emerged in the UK), P.1 (Gamma, which first emerged in Brazil) and B.1.351 (Beta, which first emerged in South Africa)—and 21 other receptor-binding domain mutants, many of which are resistant to the IgG antibodies that have been authorized for emergency use. Although engineering IgG into IgM enhances antibody potency in general, selection of an optimal epitope is critical for identifying the most effective IgM that can overcome resistance. In mice, a single intranasal dose of IgM-14 at 0.044 mg per kg body weight confers prophylactic efficacy and a single dose at 0.4 mg per kg confers therapeutic efficacy against SARS-CoV-2. IgM-14, but not IgG-14, also confers potent therapeutic protection against the P.1 and B.1.351 variants. IgM-14 exhibits desirable pharmacokinetics and safety profiles when administered intranasally in rodents. Our results show that intranasal administration of an engineered IgM can improve efficacy, reduce resistance and simplify the prophylactic and therapeutic treatment of COVID-19.

[Download PDF](#)

## Main

The respiratory tract is the major target for SARS-CoV-2 infection<sup>5</sup>. High viral load in the respiratory tract correlates with severe disease in patients with COVID-19<sup>6</sup>. So far,

almost all neutralizing monoclonal antibodies (monoclonal antibodies) that have been tested in clinical trials are the IgG1 isotype and are administered through intravenous infusion. Circulating IgG antibodies lack efficient access to mucosal compartments<sup>7</sup>. After intravenous infusion, antibody levels in the lung are around 200–500 times lower than those in the serum<sup>8</sup>. Potent neutralizing IgG1 monoclonal antibodies need to be administered at high doses (up to 8 g); even so, the antiviral effect is small in the respiratory tract in patients treated with drug compared to those treated with placebo<sup>9</sup>. Emerging SARS-CoV-2 variants of concern (VOCs) are resistant to many potent neutralizing IgG1 monoclonal antibodies, including those that are in clinical trials and that have been authorized for emergency use<sup>2,3,4</sup>. Therefore, the development of new antibody therapies that can overcome these challenges is urgently needed.

IgM and IgA are mucosal antibodies that constitute the first line of defence against mucosal pathogens. Typically, IgM assembles into pentamers and IgA1 into dimers in the presence of the joining chain (J-chain), which facilitates efficient mucosal transcytosis of antibodies<sup>10,11</sup>. IgM and IgA1 can also be nebulized and reach airway tissues after inhalation<sup>12</sup>. Owing to avidity effects, multivalent antibodies can exhibit enhanced neutralization of SARS-CoV-2<sup>13</sup> and reduce the evasion of antibodies by the virus<sup>14</sup>. An IgM pentamer is naturally decavalent owing to the repetitive antigen-binding variable fragments (Fvs)<sup>11</sup>. These unique features make the intranasal delivery of IgM neutralizing monoclonal antibodies appealing for the prevention and treatment of COVID-19.

## Engineering of neutralizing IgM and IgA1

To develop human IgM and IgA1 neutralizing monoclonal antibodies, we performed antibody engineering based on the CR3022 monoclonal antibody<sup>15</sup> and five IgG1 monoclonal antibodies (CoV2-06, CoV2-09, CoV2-12, CoV2-14 and CoV2-16) that were previously isolated from a phage-displayed antibody library<sup>1</sup>. These six monoclonal antibodies recognize different epitopes on the receptor-binding domain (RBD) (Fig. 1a, Extended Data Fig. 1a, b). The Fv of IgG1 was engineered into human IgM or IgA1 scaffolds for co-expression with the J-chain. The engineered IgM is a pentamer and IgA1 is a dimer (Fig. 1b). After production, the monoclonal antibodies of different isotypes assembled correctly and exhibited a purity of greater than 95% (Fig. 1c, d). IgM and IgA1 bound to the RBD more strongly than did IgG1 (Fig. 1e, Extended Data Fig. 1c). The higher binding activities, as evidenced by smaller values of the half-maximal effective concentration ( $EC_{50}$ ), were consistent with the higher valencies of IgA1 and IgM (Fig. 1f). Neutralization of live SARS-CoV-2 at monoclonal antibody concentrations of 0.1 and 1  $\mu\text{g ml}^{-1}$  showed stronger inhibition of virus infection by IgM and IgA1, as compared to IgG1, with IgM CoV2-14 being the most potent monoclonal antibody (Fig. 1g, Extended Data Fig. 1d). These results

demonstrate successful engineering of IgM and IgA1 monoclonal antibodies. The enhancement in binding and neutralization for IgM and IgA1 antibodies derived from different IgG1 antibodies exhibited different patterns, as exemplified by CoV2-06 and CoV2-14 (Fig. 1h), which suggests that, in addition to valency, IgG1 epitopes affect the neutralizing potency of the corresponding IgM and IgA1 antibodies.

**Fig. 1: Engineering of IgM and IgA1 neutralizing monoclonal antibodies.**

 [figure1](#)

**a**, The IgG1 epitope residues are shown as spheres on the RBD. **b**, Illustration of antibody engineering from IgG1 into IgM and IgA1. **c, d**, SDS-PAGE (left) and native PAGE (right) (**c**) and size-exclusion chromatography (**d**) analysis showing monoclonal antibody assembly and purity. The gel images were from one experiment. HC, heavy chain; LC, light chain. **e**, ELISA binding to the RBD by the indicated isotypes of CoV2-14. Data are mean of duplicate wells. **f**, Antibody valency versus EC<sub>50</sub> for the indicated isotypes. **g**, Neutralization of SARS-CoV-2 by the indicated monoclonal antibodies at 0.1 µg ml<sup>-1</sup>. Data are mean ± s.d. of triplicates. **h**, The EC<sub>50</sub> and per cent neutralization of the three isotypes for CoV2-06 and CoV2-14 are plotted to illustrate different correlation patterns between binding and neutralization.

[Source data](#)

[Full size image](#)

## Enhanced potency of IgM over IgG

We focused on IgM CoV2-14 (IgM-14) for detailed characterization in parallel with IgG-14. To examine how epitopes affect the potency of engineered IgM antibodies, we also characterized the IgM and IgG1 pairs of CoV2-06 (IgM-06 and IgG-06). In an enzyme-linked immunosorbent assay (ELISA), IgM-14 showed much stronger binding to the spike protein than did IgG-14 (Fig. 2a). In kinetic binding, IgM-14 exhibited faster association with and slower dissociation from the spike protein than did IgG-14

(Fig. 2b, c). In neutralization titrations in Vero and human ACE2-overexpressing A549 (A549-ACE2) cells, IgM-14 markedly shifted the curves towards a higher potency relative to IgG-14 (Fig. 2d, e). Figure 2f summarizes the binding EC<sub>50</sub>, the avidity-related parameters (association  $K_{\text{on}}$ , dissociation  $K_{\text{dis}}$  and avidity  $K_D$ ), the half-maximal neutralization titre (NT<sub>50</sub>) and the fold changes of these values between IgM-14 and IgG-14. The change was 135-fold for EC<sub>50</sub>, 13.6-fold for  $K_D$ , 39-fold for NT<sub>50</sub> on a weight basis and 236-fold for NT<sub>50</sub> on a molar basis. For IgM-06 and IgG-06, there was a 17.6-fold change in EC<sub>50</sub> and a change of more than 750-fold in  $K_D$ , but the change in NT<sub>50</sub> was only 2.5-fold on a weight basis and 14.9-fold on a molar basis (Extended Data Fig. 2a–d). A substantial enhancement of neutralizing activity compared to the corresponding IgG1 was observed for IgA1-14 (26-fold) but not for IgA1-06 (2.3-fold) (Extended Data Fig. 2e). The contrast between CoV2-06 and CoV2-14 supports the conclusion that epitope selection is critical for identifying the most potent neutralizing IgM.

**Fig. 2: Enhanced binding, neutralization and ACE2-blocking by IgM-14 over IgG-14.**

 figure2

**a**, ELISA for binding to the spike protein. Data are mean of duplicate wells. OD<sub>450 nm</sub>, optical density at 450 nm. **b, c**, Binding kinetics of IgM-14 (b) and IgG-14 (c) to the spike protein. **d, e**, Neutralization of SARS-CoV-2 in Vero (d) and A549-ACE2 (e) cells. Data are mean of duplicate wells. **f**, Summary of binding EC<sub>50</sub>, association ( $K_{\text{on}}$ ), dissociation ( $K_{\text{dis}}$ ), avidity ( $K_D$ ) and neutralization NT<sub>50</sub> values. The fold change of NT<sub>50</sub> values between IgM-14 and IgG-14 is highlighted in red. **g**, Superposition of the IgFv-14–RBD and ACE2–RBD complexes. IgFv-14 is shown as a surface with the variable heavy chain (Vh) coloured in cyan and the variable light chain (VL) in magenta. The RBD–ACE2 complex is shown as a cartoon with the RBD coloured in grey and ACE2 in green. The dashed box indicates steric clash. **h**, Blocking of RBD and ACE2 interaction. Data for each antibody concentration are from one biosensor.

[Source data](#)

[Full size image](#)

To understand the structural mechanism of this observation, we performed antibody docking to simulate the Fv and RBD complex using Rosetta-based protocols<sup>16</sup>. The Fv of IgG-14 (IgFv-14) targets the back side of the RBD, whereas IgFv-06 targets the front side of the RBD (Extended Data Fig. 3a, b). Therefore, IgFv-14 has a larger steric clash with ACE2 (Fig. 2g). To support the structural analysis, we used a bio-layer interferometry (BLI)-based assay to assess the ACE2-blocking activities of IgM and IgG1 (Extended Data Fig. 4a–d). IgM-14 inhibited RBD–ACE2 interaction more strongly than did IgG-14 (Fig. 2h). By contrast, IgFv-06 has a smaller steric clash with ACE2 (Extended Data Fig. 4e). IgM-06 also blocked RBD–ACE2 interaction more strongly than did IgG-06, however, neither IgM-06 nor IgG-06 achieved full blocking even at the highest concentration tested (Extended Data Fig. 4f). These results demonstrate that epitope-dependent steric hindrance is an important mechanism for IgM-14 to exhibit potent neutralization.

## Broad coverage of variants by IgM-14

SARS-CoV-2 escapes from antibody neutralization by acquiring mutations in resistance-selection experiments and in natural circulation<sup>1,2,3,4,17,18</sup>. We previously identified neutralization-resistant RBD mutations K444R for IgG-06 and E484A for IgG-14 (ref. 1). To test whether IgM can neutralize these IgG escape mutants, we constructed three SARS-CoV-2 variants that contain K444R, E484A or both K444R and E484A mutations (K444R + E484A) (Extended Data Fig. 5a, b). IgG-14 effectively neutralized the K444R variant and marginally neutralized the E484A and K444R + E484A variants. Notably, IgM-14 potently neutralized all three variants including the K444R + E484A variant, which is resistant to the cocktail of IgG-06 + IgG-14 (Fig. 3a). The NT<sub>50</sub> values of IgM-14 against the E484A and the K444R + E484A variants were 0.064 µg ml<sup>-1</sup> and 0.055 µg ml<sup>-1</sup>, respectively (Fig. 3b), which are comparable to the NT<sub>50</sub> (0.01 µg ml<sup>-1</sup>) value of IgM-14 against the wild-type virus (Fig. 2f). IgM-14 is 2,343-fold and 1,949-fold more potent than IgG-14 in neutralizing the E484A and K444R + E484A variants, respectively (Fig. 3b). Similarly, in A549-ACE2 cells, IgM-14 is 714-fold and 2,217-fold more potent than IgG-14 in neutralizing the E484A and K444R + E484A variants, respectively (Extended Data Fig. 5c, d). By contrast, IgM-06 only neutralized the E484A variant and not the K444R or K444R + E484A variants (Extended Data Fig. 5e, f). These data show that IgM-14 can effectively neutralize IgG-14 escape variants. The contrast between CoV2-06 and CoV2-14 further underscores that epitope selection is critical for identifying the IgM antibody that can overcome IgG escape mutations.

**Fig. 3: Broader coverage of escape variants by IgM-14 over IgG-14.**

---

 **figure3**

**a**, Neutralization of the K444R (IgG-06-resistant), E484A (IgG-14-resistant) and K444R + E484A (IgG-06 + IgG-14-resistant) SARS-CoV-2 variants. **b**, Summary of NT<sub>50</sub> values against the indicated SARS-CoV-2 variants. **c**, PRNT assay using the US-WA1 strain and the recombinant B.1.1.7, P.1 and B.1.351 variants. **d**, Summary of PRNT<sub>50</sub> values against the indicated viruses. Data are mean of duplicate wells for **a**, **c**. The fold changes in NT<sub>50</sub> and PRNT<sub>50</sub> values between IgM-14 and IgG-14 against the indicated resistant viruses are highlighted in red. **e**, Summary of binding K<sub>D</sub> and ACE2-blocking IC<sub>50</sub> values to a panel of 21 RBD mutants. The hash symbol indicates that E484K is an escape mutation for both LY-CoV555 and REGN-10933. Asterisks indicate that the half-maximal blocking was not achieved at the highest monoclonal antibody concentration (30 nM) and the IC<sub>50</sub> values are defined as ≥90 nM.

[Source data](#)

[Full size image](#)

To assess the neutralizing activities of IgM-14 and IgG-14 against the recently emerged SARS-CoV-2 VOCs, we constructed recombinant viruses based on the US-WA1 strain and replaced its full spike gene with that of the B.1.1.7, P.1 or B.1.351 variants. In the plaque-reduction neutralization test (PRNT), IgG-14 potently neutralized the B.1.1.7 variant but weakly neutralized the P.1 and B.1.351 variants (Fig. 3c, d). These results are consistent with E484 being a critical epitope residue for IgG-14, as both the P.1 and B.1.351 variants contain the E484K neutralization-resistant mutation. Notably, IgM-14 potently neutralized all three variants (Fig. 3c). The half-maximal concentration values in PRNTs (PRNT<sub>50</sub> values) of IgM-14 against the B.1.1.7, P.1 and B.1.351 variants were 0.006, 0.023 and 0.031 µg ml<sup>-1</sup>,

respectively, which are comparable to the PRNT<sub>50</sub> (0.011 µg ml<sup>-1</sup>) value of IgM-14 against the US-WA1 strain. IgM-14 is 45-fold, 547-fold and 374-fold more potent than IgG-14 in neutralizing the B.1.1.7, P.1 and B.1.351 variants, respectively (Fig. 3d). These data show that IgM-14, but not IgG-14, can tolerate the RBD mutations in VOCs.

We used the wild-type, K444R, E484A and K444R + E484A RBD proteins to assess the binding and ACE2-blocking activities of IgM-14 versus IgG-14 and IgM-06 versus IgG-06 (Extended Data Fig. 6a–j). Consistent with the neutralization results, IgM-14 and IgM-06 bound to these RBD proteins and blocked ACE2 interactions more effectively than did their parental IgG1 antibodies (Extended Data Fig. 6k). Similar results were also observed using the F486S (another IgG-14-resistant mutation) and K444S (IgG-06-resistant) RBD proteins (Extended Data Fig. 6l–o).

We generated another 19 RBD mutants to characterize the binding and ACE2-blocking activities of IgM-14 and IgG-14 (Extended Data Figs. 7a–u, 8a–s). The N439K and S477N are resistant mutations for several neutralizing IgG1 monoclonal antibodies<sup>17,18</sup>. The K417N, N501Y, E484K + N501Y and K417N + E484K + N501Y mutations represent the B.1.351 variant<sup>2</sup>. The remaining 13 mutations are associated with resistance to three IgG1 monoclonal antibodies that have been authorized for emergency use<sup>19,20</sup>. Figure 3e summarizes the K<sub>D</sub> and half-maximal inhibitory concentration (IC<sub>50</sub>) values. Across all RBD mutants, IgM-14 exhibited much higher binding to RBD and higher ACE2-blocking activity compared to IgG-14. This superior activity is probably attributable to avidity effects, which result in much slower dissociation kinetics from the RBD mutants for IgM-14 than that for IgG-14 (Extended Data Fig. 7v, Supplementary Table 1). A good correlation between IC<sub>50</sub> and NT<sub>50</sub> values was observed for the VOCs (Extended Data Fig. 8t). Globally, these mutations emerged with varying frequencies in the 373,387 viral sequences analysed (Extended Data Fig. 8u). These mutations affect RBD functionality to various degrees (Extended Data Fig. 8v), indicating that RBD mutational effects on antibody binding and virus fitness together determine the neutralizing activities of monoclonal antibodies against these mutants. Together, these data demonstrate that IgM-14 is superior to IgG-14 in covering viral escape mutations.

## IgM-14 delivered intranasally targets the airways

We evaluated the feasibility of IgM-14 for intranasal administration by tracking antibody bio-distribution in mice (Fig. 4a). After a single intranasal dose, IgM-14 (labelled with Alexa Fluor 750) was enriched in the nasal cavity and lasted for at least 96 h in whole-body imaging (Fig. 4b, Extended Data Fig. 9a). Ex vivo organ imaging showed that IgM-14 enriched in the nasal cavity and lung at various time points and

was still evident in the nasal cavity at 168 h. The blood and other organs had minimal antibody exposure (Fig. 4c, d, Extended Data Fig. 9b, c). These results indicate that intranasally administered IgM-14 mainly targets the respiratory tract, with long-term retention in the nasal cavity and lung. Nasal epithelium is at first the dominant site for SARS-CoV-2 respiratory tract infection, followed by aspiration of virus into the lung<sup>21</sup>. Therefore, intranasal administration can efficiently load IgM-14 to the airways, which should confer protection against respiratory infection.

**Fig. 4: Intranasally delivered IgM-14 confers protection against SARS-CoV-2 VOCs.**

 figure4

**a**, Experimental design for the evaluation of antibody bio-distribution. **b**, Representative whole-body images. **c**, Representative ex vivo images. Bl, blood (20 µl); Br, brain; H, heart; K, kidney; Lu, lung; Lv, liver; N, nasal cavity; S, spleen. **d**, Quantification of fluorescence signals. Data are mean ± s.d. of four mice. Dashed line shows the average autofluorescence of organs. **e**, Experimental design of dose-range evaluations of IgM-14 against a mouse-adapted SARS-CoV-2 (CMA4 strain). For prophylactic treatment, IgM-14 was evaluated at two sets of dose ranges to determine the minimal effective dose. **f**, **g**, Viral loads in the lungs of CMA4-infected mice after prophylactic (**f**) or therapeutic (**g**) treatment with IgM-14.  $n = 10$  biologically independent mice for all groups, except for the 1.2 mg kg<sup>-1</sup> group in **g** ( $n = 15$ ). **h**, **i**, Viral loads in the lungs of mice infected with the P.1 (**h**) and B.1.351 (**i**) variant, after therapeutic treatment as indicated.  $n = 5$  biologically independent mice for all groups. The solid lines in **f–i** indicate median viral loads in the lung. An ordinary one-way ANOVA with Sidak's multiple comparisons was used in the statistical analysis for **f**, **g**. A two-sided Mann–Whitney test was used in the statistical analysis for **h**, **i**.

[Source data](#)

[Full size image](#)

## In vivo protection against VOCs by IgM-14

We evaluated the protective efficacy of IgM-14 using intranasal administration with a mouse-adapted virus that contains an N501Y mutation<sup>26</sup>, thus representing the RBD of the B.1.1.7 variant. First, we tested the three isotypes of CoV2-06 and CoV2-14 at a prophylactic dose of 3.5 mg per kg body weight (hereafter, mg kg<sup>-1</sup>) (Extended Data Fig. 10a). Peak lung viral loads on day 2 after infection were reduced to undetectable levels in all four mice of the IgG-06 and IgM-06 groups and in three of the four mice of the IgA1-06, IgG-14, IgA1-14 and IgM-14 groups (Extended Data Fig. 10b). We then focused on IgM-14 and performed dose range evaluations. Five dose levels (3.5, 1.2, 0.4, 0.13, and 0.044 mg kg<sup>-1</sup>) were tested for prophylactic treatment and three dose levels (3.5, 1.2 and 0.4 mg kg<sup>-1</sup>) were tested for therapeutic treatment (Fig. 4e). For prophylactic treatment, lung viral loads were reduced to undetectable levels in 90% (9/10), 70% (7/10), 90% (9/10) and 60% (6/10) of mice in the 3.5, 1.2, 0.4 and 0.13 mg kg<sup>-1</sup> groups, respectively. Even with the 0.044 mg kg<sup>-1</sup> dose, the median viral load (5.56-log) was significantly reduced by fivefold as compared to the isotype group (6.29-log; Fig. 4f). For therapeutic treatment, the median viral loads (excluding the mice with undetectable virus) were reduced by 13,667-fold, 13,667-fold and 56-fold in the 3.5, 1.2, and 0.4 mg kg<sup>-1</sup> groups, respectively (Fig. 4g). Quantification of viral RNA in lung samples confirmed the anti-viral effects of IgM-14 (Extended Data Fig. 10c). IgM-14 showed more potent efficacy than IgG-14 at a therapeutic dose of 1.2 mg kg<sup>-1</sup> (Extended Data Fig. 10d). No mutations were observed in the RBD of recovered virus after treatment with IgM-14 (Extended Data Fig. 10e). These data show that IgM-14 confers protection with an effective dose as low as 0.044 mg kg<sup>-1</sup> for prophylactic treatment and 0.4 mg kg<sup>-1</sup> for therapeutic treatment.

We also compared the therapeutic efficacy of IgM-14 and IgG-14 against the P.1 and B.1.351 VOCs. For the P.1 variant, IgM-14 reduced lung viral loads by 16,429-fold and 144-fold in the 3.5 and 1.2 mg kg<sup>-1</sup> groups, respectively. In comparison, IgG-14 had a marginal anti-viral effect, reducing viral loads by around fourfold at either dose (Fig. 4h). Similarly, IgM-14 exhibited a significantly higher potency than IgG-14 against the B.1.351 variant (Fig. 4i). Together, our data show that intranasal administration of IgM-14 confers highly efficacious respiratory protection and that IgM-14 is superior to IgG-14 for protection against the tested VOCs.

## Preclinical pharmacokinetics and safety

To further evaluate the translational potential of IgM-14, we conducted an intranasal pharmacokinetic study in mice. A single intranasal administration of  $5 \text{ mg kg}^{-1}$  of IgM-14 resulted in low levels of antibody in the blood (around  $100 \text{ ng ml}^{-1}$ ) that persisted for several hours (Extended Data Fig. 10f). In patients with COVID-19, the positive rate of viral RNA is high (93%) in bronchoalveolar lavage samples but very low (less than 1%) in blood samples<sup>22</sup>. Therefore, intranasally administered IgM-14 is better focused on targeting the site of virus replication in the respiratory tract, compared to intravenously infused IgG1. To further assess the tolerability of IgM-14, a pilot safety study was conducted in rats. Groups of rats were dosed intranasally with IgM-14 twice daily for five consecutive days. All rats survived to the end of the study with no change in body weight (Extended Data Fig. 10g).

## Discussion

In this study, we have shown that six engineered IgM antibodies exhibit higher binding and neutralizing activities than their parental IgG1 antibodies, indicating a general mechanism of gaining potency through avidity. However, epitope selection is required to identify the most appropriate engineered IgM. Epitope-dependent steric hindrance is one mechanism for IgM to exhibit enhanced neutralization compared to IgG1. IgM may also engage more spike proteins than can an IgG1, and such a mechanism is related to antibody binding angles<sup>23</sup>. Epitope selection is also critical for identifying the IgM that can overcome resistance. The critical epitope residues for IgG-14, but not for IgG-06, make direct contact with ACE2. Investigating resistance for other monoclonal antibody pairs may provide additional insights into the precise mechanism. Nevertheless, the finding that IgM-14 has a broader coverage of variants than IgG-14 points to the potential for leveraging IgM to combat other dynamically evolving pathogens.

Intranasally delivered IgM mainly targets the airways, which allows dose-sparing and offers highly effective protection. Previously, intrarectal delivery of an IgM has been shown to reduce viral rectal infection in macaques<sup>24</sup>, and intranasal delivery of IgG antibodies has been studied. Together, these data indicate that mucosal delivery is a promising alternative to intravenous infusion for treating mucosal viral infections. Notably, intranasal dosing of IgM simplifies administration and also provides broader coverage and better protection against variants.

Traditionally, IgM antibodies have been considered difficult to produce. With recent advances in manufacturing, engineered IgM antibodies—including IgM-14—can be produced at high titres (greater than  $1 \text{ g l}^{-1}$ ) with good purity and stability. Several engineered IgM antibodies have successfully entered human clinical trials in

oncology<sup>25</sup>. IgM-14 holds potential for development as an effective and flexible therapy for outpatients and high-risk populations.

In summary, we show that engineered IgM is a promising drug modality with potent neutralization, broad coverage of variants, desirable pharmacokinetics and safety profiles, and effective respiratory protection. We envision that engineered IgM administered in an intranasal manner can serve as a therapeutic platform for COVID-19 as well as for other respiratory viral diseases.

## 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.

### Cells and viruses

Vero (ATCC, CCL-81) and Vero-E6 (ATCC, CRL-1586) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Expi293F cells (Gibco, 100044202) were maintained in Expi293 expression medium without FBS. ExpiCHO-S cells (Thermo Fisher Scientific, A29127) were maintained in ExpiCHO expression medium without FBS. CHO cells (Horizon Discovery) were maintained in Ex-Cell Advanced CHO fed-batch medium (Millipore-Sigma, 14366C) supplemented with 4 mM l-glutamine. The human ACE2-overexpressing A549 (A549-ACE2) cells were generated in-house using lentivirus transduction. All cells have been tested negative for mycoplasma contamination. The SARS-CoV-2 (US-WA1 strain) mNeonGreen (mNG) reporter virus was generated in a previous study using a SARS-CoV-2 infectious clone<sup>27</sup>. The SARS-CoV-2 variants with the K444R, the E484A or the K444R + E484A mutations were also generated based on this infectious clone. A mouse-adapted SARS-CoV-2 (CMA4 strain) used for in vivo studies was recombinantly generated. In brief, a previously used mouse-adapted SARS-CoV-2 (CMA3 strain)<sup>1</sup> was passed in BALB/c mice for 20 passages. To generate the CMA4 strain, the adaptive mutations were identified and reintroduced into the backbone of the CMA3 strain. The generation and characterization of these mouse-adapted viruses are reported in a separate preprint<sup>26</sup>. The recombinant SARS-CoV-2 viruses with spike mutations from the B.1.1.7, P.1 and B.1.351 lineages were prepared on the genetic background of an infectious cDNA clone derived from clinical strain USA-WA1/2020 (ref. <sup>27</sup>) using a previously described protocol<sup>28</sup>. The spike mutations for the B.1.1.7 variant included 69–70 deletion, Y145 deletion, N501Y,

A570D, D614G, P681H, T716I, S982A and D1118H. The spike mutations for the P.1 variant included L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I and V1176F. The spike mutations for the B.1.351 variant included D80A, D215G, 242–244 deletion, K417N, E484K, N501Y, D614G and A701V. The recombinant viruses were sequenced to confirm the presence of the engineered mutations and to rule out undesired mutations.

## Engineering and production of IgM and IgA1 antibodies

The VH and VL regions of five human anti-SARS-CoV-2 IgG1 antibodies—CoV2-06, CoV2-09, CoV2-12, CoV2-14, and CoV2-16—were incorporated into expression vectors encoding IgM and IgA1 constant regions and a human J-chain according to standard cloning protocols. The original isolated CoV2-06, CoV2-09, CoV2-12, CoV2-14, and CoV2-16 IgG1 antibodies had lambda light chains. For initial in vitro characterization, the light chains for these antibodies were constructed as lambda-VL-kappa-Cl hybrid light chains and the antibodies were transiently expressed in Expi293 cells. For subsequent in vitro and in vivo experiments, complete lambda light chains were used and the antibodies were expressed in Expi293F, ExpiCHO-S or CHO cells. The control antibody-binding domain from anti-SARS-CoV antibody CR3022 was similarly constructed. The IgG1 and IgA1 antibody constructs were purified by affinity chromatography using protein A and Capture Select LC-lambda resins (Thermo Fisher Scientific), respectively. The IgM antibody constructs were purified by mixed-mode chromatography and anion-exchange chromatography<sup>25</sup>. The degree of assembly and purity were assessed by SDS-PAGE and native PAGE. For SDS-PAGE, all antibodies were run on 10% Mini-PROTEAN TGX gels (Bio-Rad). The gels were stained with Coomassie blue R-250 (Bio-Rad). For native PAGE, IgG1 and IgA1 antibodies were run on native 4–12% Bis-Tris gels (Life Technologies), whereas IgM antibodies were run on native 3–12% Bis-Tris gels (Life Technologies) by referring to a previously described electrophoretic method<sup>29</sup>. The gels were stained with Colloidal Blue Stain (Life Technologies). For gel source data, see Supplementary Fig. 1. The IgG1, IgA1 and IgM antibodies were also assessed by size-exclusion chromatography and the data were analysed using the UNICORN v.7.0 software.

## Recombinant proteins

The spike protein (S1 + S2 ECD, His tag) (40589-V08B1) and the human ACE2 protein (10108-H08H) were purchased from Sino Biological. The His-tagged RBD protein (CoV2-RBD-His) of SARS-CoV-2 spike protein was purchased from ATUM (65639.1.a). The Fc-tagged wild type RBD protein and seven RBD proteins that contain amino acid mutations, including K444R, E484A, K444R + E484A, K444S and F486S, were generated in a previous study<sup>1</sup>. The expression constructs of 19 Fc-tagged RBD proteins that contain amino acid mutations, including N439K, S477N,

N501Y, K417N, E484K + N501Y, K417N + E484K + N501Y, E484K, F490S, Q493R, S494P, K417E, Y453F, L455F, G476S, F486V, Q493K, K444Q, V445A and G446V, were generated by overlap PCR using specific primers (Supplementary Table 2) in this study. The proteins were produced from Expi293F cells and purified using the CaptivA protein A affinity resin (Repligen, CA-PRI-0100).

## ELISA binding assay

ELISA titration of monoclonal antibody binding to the RBD of the SARS-CoV-2 spike protein was performed as follows. The 96-well white polystyrene ELISA plates (Pierce 15042) were coated with 100 µl per well of 0.5 µg ml<sup>-1</sup> recombinant His-tagged RBD protein overnight at 4 °C. Plates were then washed five times with phosphate-buffered saline (PBS) with 0.05% Tween and blocked with 2% bovine serum albumin (BSA)-PBS. After blocking, 100-µl serial dilutions of monoclonal antibodies were added to the wells and incubated at room temperature for 2 h. The plates were then washed 10 times and incubated with HRP-conjugated mouse anti-human kappa (Southern Biotech, 9230-05, 1:6,000 diluted in 2% BSA-PBS) for 30 min. After 10 final washes using 0.05% PBS-Tween, the plates were read using Super Signal chemiluminescent substrate (Thermo Fisher Scientific, 37070). Luminescent data were collected on an EnVision plate reader (Perkin Elmer) and analysed with GraphPad Prism 8.

For ELISA titration of monoclonal antibody binding to SARS-CoV-2 spike protein, high-binding ELISA plates were coated with recombinant spike protein (1 µg ml<sup>-1</sup>) at 4 °C overnight and blocked with 5% skimmed milk at 37 °C for 2 h. Antibodies were serially diluted in 1% skimmed milk and added at a volume of 100 µl per well for incubation at 37 °C for 2 h. The HRP-conjugated F(ab')<sub>2</sub> fragment goat anti-human IgA + IgG + IgM (H+L) antibody (Jackson ImmunoResearch, 109-036-064) was diluted 1:5,000 and added at a volume of 100 µl per well for incubation at 37 °C for 1 h. The plates were washed around three to five times with PBST (0.05% Tween-20) between incubation steps. TMB (3,3',5,5'-tetramethylbenzidine) substrate was added at 100 µl per well for colour development. The reaction was stopped by adding 50 µl per well 2M H<sub>2</sub>SO<sub>4</sub>. The OD<sub>450</sub> nm was read by a SpectraMax microplate reader and analysed with GraphPad Prism 8.

## Antibody avidity

The measurement of antibody avidity (apparent affinity) to the spike protein and the wild-type or mutant RBD proteins were performed on the ForteBio Octet RED96 system. The His-tagged spike protein (15 µg ml<sup>-1</sup>) was captured on the Ni-NTA biosensor and the Fc-tagged RBD proteins (20 µg ml<sup>-1</sup>) were captured on the protein

A biosensor. Following 10 s of baseline run in kinetics buffer, the sensors were dipped in threefold serially diluted antibodies (0.12 nM to 90 nM) for 200 s to record association kinetics. Then, the sensors were dipped into kinetics buffer for 400 s to record dissociation kinetics. For avidity measurement with Fc-tagged RBD proteins, the biosensors were blocked with a control Fc protein (150 µg ml<sup>-1</sup>) for 200 s to occupy the free protein A. ForteBio Octet Data Analysis software was used to fit the  $K_D$  data using the global fitting method.

## Neutralization assays

The neutralization of SARS-CoV-2 viruses with wild-type RBD, K444R RBD, E484A RBD and K444R + E484A RBD was performed using the mNG reporter viruses. In brief, a total of  $1.5 \times 10^4$  Vero cells or A549-ACE2 cells were plated into each well of a black transparent flat-bottom 96-well plate (Greiner Bio-One, 655090). The next day, antibodies (serial dilutions) were mixed with an equal volume of SARS-CoV-2-mNG virus (multiplicity of infection = 0.5). After 1 h incubation at 37 °C, the antibody–virus complexes were inoculated with Vero cells. At 20 h after infection, nuclei were stained by the addition of Hoechst 33342 (Thermo Fisher Scientific) to a final concentration of 10 nM. Fluorescent images were acquired using a Cytation 7 multimode reader (BioTek). Total cells (in blue) and mNG-positive cells (in green) were counted, and the infection rate was calculated. The relative infection rates were calculated by normalizing the infection rate of each well to that of control wells (no antibody treatment).

The neutralization of the SARS-CoV-2 US-WA1 strain and the SARS-CoV-2 viruses containing the spike mutations from the B.1.1.7, P.1 and B.1.351 lineages were performed using the PRNT. In brief, antibodies were serially diluted in culture medium and incubated with 100 PFU of wild-type or mutant viruses at 37 °C for 1 h, after which the antibody–virus mixtures were inoculated onto Vero E6 cell monolayer in six-well plates. After 1 h of infection at 37 °C, 2 ml of 2% SeaPlaque agar (Lonza) in DMEM containing 2% FBS and 1% penicillin–streptomycin was added to the cells. After 2 d of incubation, 2 ml of 2% SeaPlaque agar in DMEM containing 2% FBS, 1% penicillin–streptomycin and 0.01% neutral red (Sigma-Aldrich) were added on top of the first layer. After another 16 h of incubation at 37 °C, plaque numbers were counted. The relative infection was obtained by the plaque counts from the antibody-treated groups to the untreated groups. The relative infection versus the concentration of the antibody (in log<sub>10</sub> scale) was plotted. PRNT<sub>50</sub> titres were calculated using a nonlinear regression model. All SARS-CoV-2 manipulations were conducted at the Biosafety Level-3 facility with approval from the Institutional Biosafety Committee at the University of Texas Medical Branch.

## Antibody structure modelling and molecular docking

The three-dimensional (3D) structures of antibody-RBD complexes were built using the Rosetta-based computational protocols<sup>16</sup>. In brief, the 3D structures of the antibody Fv region were first predicted from their amino acid sequences by the antibody function of Rosetta. Then, the structure of the Fv region was docked to the published RBD structure (Protein Data Bank (PDB) code: 6M0J) by using the SnugDock function of Rosetta. The binding interface of the antibody Fv relative to the RBD was positioned using PyMol software. The antibody epitope residues were previously identified using an alanine-scanning RBD mutant library<sup>1</sup> and were incorporated into the analysis to inform docking studies. All the modelling and docking tasks were performed by querying the online server (<https://rosie.graylab.jhu.edu/>). The Fv–RBD structures with the lowest interface energy and having good coverage of epitope residues in the docking were selected for further analysis. To visualize the binding model of the antibody relative to the ACE2, the Fv–RBD structures were aligned with the ACE2–RBD complex structure (PDB: 6M0J) based on the RBD using PyMol software.

### **Antibody blocking of RBD and ACE2 interaction**

The wild-type RBD or the mutant RBD proteins ( $4 \mu\text{g ml}^{-1}$ ) were captured on the protein A biosensor for 300 s. The sensors were then blocked by a control Fc protein ( $150 \mu\text{g ml}^{-1}$ ) for 200 s to occupy the free protein A on the sensor. The serially diluted antibodies (0.041 nM to 30 nM) were incubated with the sensors for 200 s to allow antibody and RBD binding. The irrelevant isotype antibodies (30 nM) were used as controls. After 10 s of baseline run in kinetics buffer, the sensors were dipped in the ACE2 solution ( $8 \mu\text{g ml}^{-1}$ ) for 200 s to record the response signal. For analysis of the  $\text{IC}_{50}$ , the ACE2 response values were normalized to the starting points. The blocking percentages at each concentration were calculated as:  $((\text{normalized ACE2 response of isotype antibody} - \text{normalized ACE2 response of tested antibody}) / \text{normalized ACE2 response of isotype antibody}) \times 100$ . The dose-blocking curves were plotted and the blocking  $\text{IC}_{50}$  values were calculated by nonlinear fit using GraphPad Prism 8.

### **Bioinformatics analysis of RBD mutations in circulating virus**

As of 8 February 2021, 373,387 human SARS-CoV-2 genomic sequences were analysed for the occurring frequencies of individual RBD mutations in the global circulating virus variants. The database of the COVID-19 Viral Genome Analysis Pipeline was queried and the Tracking Mutations tool ([https://cov.lanl.gov/content/sequence/TRACK\\_MUT/trackmut.html](https://cov.lanl.gov/content/sequence/TRACK_MUT/trackmut.html)) was used to obtain the numbers of virus variants that contain amino acid mutations at individual RBD sites. The frequencies of individual RBD mutations were calculated by dividing

the numbers of virus variants by a total of 373,387 virus sequences. The frequencies were expressed as number of variants per 10,000 sequences.

## Tracking antibody bio-distribution in mice

Antibodies were conjugated with Alexa Fluor 750 dye for in vivo imaging studies. In brief, 1.78 mg ml<sup>-1</sup> solution of IgM-14 in PBS (pH 7.4) was reacted with Alexa Fluor 750 succinimidyl ester (AF750-NHS, Thermo Fisher Scientific) in the presence of 3% DMSO and 10% sodium bicarbonate buffer (v/v, pH 8.3) using the molar ratio of 1:10 protein to fluorescent probe at room temperature for 1 h. Unreacted dye was removed by dialysis, the labelled antibody was washed in PBS and concentrated with an Amicon ultra centrifugal filter unit (MWKO 10 kDa). All procedures were performed under dimmed light. The CD-1 mice (6–8 weeks, female, Charles River Laboratories) were anaesthetized by inhalation of 2% isoflurane and placed in a supine position. The mice were administered intranasally with Alexa Fluor 750-labelled IgM-14 to both nostrils of the mice using a fine pipet tip (40 µl total) to achieve the final antibody dose of 1.2 mg kg<sup>-1</sup>. The mice were imaged at predetermined time points after administration (fluorescence ex = 740 nm, em = 790 nm, auto-exposure setting, n = 4 mice in each group) using an IVIS Lumina XRMS Imager (Perkin Elmer). At the time of euthanasia, 20 µl of blood, the heart, lung, liver, spleen, kidney, brain and nasal cavity samples were excised and imaged. Regions of interest (ROIs) were drawn and average radiant efficiency ((p s<sup>-1</sup> cm<sup>-2</sup> sr<sup>-1</sup>)/(µW cm<sup>-2</sup>)) was measured. This parameter represents the sum of the radiance from each pixel inside the ROI divided by the number of pixels. All images were processed using Living Image software (Perkin Elmer) and the same fluorescence threshold was applied for group comparison.

## Mouse infection and antibody protection

The animal study was carried out in accordance with the recommendations for care and use of animals by the Office of Laboratory Animal Welfare, National Institutes of Health. The Institutional Animal Care and Use Committee (IACUC) of the University of Texas Medical Branch approved the animal studies under protocol 1802011. Ten-to-twelve-week-old female BALB/c mice were purchased from Charles River Laboratories and maintained in Sealsafe HEPA-filtered air in/out units. A previously described mouse infection model was used to evaluate antibody protections<sup>1</sup>. Animals were anaesthetized with isoflurane and infected intranasally with 10<sup>4</sup> PFU of mouse-adapted SARS-CoV-2 (CMA4 strain; N501Y)<sup>26</sup>, the P.1 variant or the B.1.351 variant in 50 µl of PBS. Antibodies were intranasally delivered at 6 h before or 6 h after viral infection. Two days after infection, lung samples of infected mice were collected and homogenized in 1 ml PBS using the MagNA Lyser (Roche Diagnostics). The

homogenates were clarified by centrifugation at 15,000 rpm for 5 min. The supernatants were collected for measuring infectious virus titres by plaque assay as mentioned in the neutralization assay. Quantitative PCR with reverse transcription (qRT–PCR) assay was also used for measuring viral RNA (nucleocapsid gene) titres in the lung. In brief, the clarified tissue homogenates were mixed with a fivefold excess of TRIzol LS Reagent (Thermo Fisher Scientific, 10296010). Total RNA was extracted according to the manufacturer's instructions. The extracted RNA was finally dissolved in 40 µl nuclease-free water. Two microlitres of RNA samples were used for qRT–PCR assays using the iTaq SYBR Green one-step kit (Bio-Rad) on the QuantStudio Real-Time PCR systems with fast 96-well module (Thermo Fisher Scientific). The quantification of viral RNA was determined by a standard curve method using an RNA standard (in vitro transcribed 3,839 bp RNA at the nucleotide positions from 26,044 to 29,883 of SARS-CoV-2 genome) and the primers 2019-nCoV\_N2-F (5'-TTACAAACATTGGCCGAAA-3') and 2019-nCoV\_N2-R (5'-GCGCGACAT TCCGAAGAA-3').

## Pharmacokinetic and safety studies

For the pharmacokinetic studies, female BALB/c mice between the ages of 6 and 8 weeks were obtained from Charles River Laboratories. Five groups of mice (three per group) were used. Dosing of IgM-14 was done via the intranasal route according to an approved IACUC protocol. In brief, mice were lightly anaesthetized with isoflurane and 50 µl of IgM-14 at 2 mg ml<sup>-1</sup> (total dose 5 mg kg<sup>-1</sup>) was slowly instilled with a pipette, with the volume split between the two nares. Survival and terminal bleeds were collected by either retro-orbital or cardiac puncture for sample collection at 15 min, 30 min, 1 h, 2 h, 4 h, 8 h, 24 h, 48 h, 72 h and 96 h. Approximately 200 µl of blood was processed for plasma. An ELISA assay was used to determine IgM-14 concentrations in mouse plasma. In brief, Pierce 96-well ELISA plates (Thermo Fisher, 15042) were coated with 100 µl of SARS-CoV-2 spike protein RBD (319-591, C-term His tag, LakePharma, 46438) at 1 µg ml<sup>-1</sup> in 1× PBS overnight at room temperature. After blocking with 3% BSA in PBS at room temperature for 2 h and washing in PBS + Tween-20, samples prepared in PBS containing 3% BSA and 1% diluted mouse plasma were added and incubated for 2h at room temperature. After washing, mouse anti-human lambda-HRP (SouthernBiotech, 9180-05, 1:2,000 diluted) was added and the plates were incubated for 1 h at room temperature. Following washing and addition of SuperSignal ELISA Pico Chemiluminescent Substrate (Thermo Fisher Scientific, 37069), plates were read on a Perkin Elmer EnVision 2104 Multilabel MicroPlate Reader and data were analysed by applying a four-parameter logistic fit to standards using GraphPad Prism. Sample concentrations were interpolated from the standard curve. The lower limit of quantification (LLOQ) for the assay was 0.02 µg ml<sup>-1</sup>.

For the safety study, groups of rats (four per group) were dosed intranasally with IgM-14 or the vehicle control twice daily for five consecutive days. The top dose was 4 mg per kg per day (20 mg total over the 5-day period). All rats survived to the end of the study. Clinical observations, body weight changes, macroscopic observations, clinical pathology observations or organ weight effects were monitored. After the completion of dosing, the nasal cavities of each rat were also removed, sectioned and examined histologically.

## Statistical analysis

All statistical analysis was performed using GraphPad Prism 8 and the statistic tests are described in the indicated figure legends. Nonlinear regression curve fitting was performed to calculate the EC<sub>50</sub>, IC<sub>50</sub> and NT<sub>50</sub> values.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

Data associated with figures are available from the corresponding authors upon reasonable request. The COVID-19 Viral Genome Analysis Pipeline Tracking Mutations tool is available at [https://cov.lanl.gov/content/sequence/TRACK\\_MUT/trackmut.html](https://cov.lanl.gov/content/sequence/TRACK_MUT/trackmut.html). The published structure of the RBD bound to ACE2 is available at the PDB with accession code [6M0J](#). [Source data](#) are provided with this paper.

## References

1. 1.

Ku, Z. et al. Molecular determinants and mechanism for antibody cocktail preventing SARS-CoV-2 escape. *Nat. Commun.* **12**, 469 (2021).

[CAS Article](#) [ADS](#) [Google Scholar](#)

2. 2.

Wibmer, C. K. et al. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat. Med.* **27**, 622–625 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Wang, P. et al. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130–135 (2021).

[CAS](#) [Article](#) [ADS](#) [Google Scholar](#)

4. 4.

Starr, T. N. et al. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850–854 (2021).

[CAS](#) [Article](#) [ADS](#) [Google Scholar](#)

5. 5.

Wölfel, R. et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469 (2020).

[Article](#) [ADS](#) [Google Scholar](#)

6. 6.

Pujadas, E. et al. SARS-CoV-2 viral load predicts COVID-19 mortality. *Lancet Respir. Med.* **8**, e70 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Iwasaki, A. Exploiting mucosal immunity for antiviral vaccines. *Annu. Rev. Immunol.* **34**, 575–608 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

8. 8.

DeFrancesco, L. COVID-19 antibodies on trial. *Nat. Biotechnol.* **38**, 1242–1252 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Weinreich, D. M. et al. REGN-COV2, a neutralizing antibody cocktail, in outpatients with covid-19. *N. Engl. J. Med.* **384**, 238–251 (2021).

[CAS Article](#) [Google Scholar](#)

10. 10.

Kumar, N., Arthur, C. P., Ciferri, C. & Matsumoto, M. L. Structure of the secretory immunoglobulin A core. *Science* **367**, 1008–1014 (2020).

[CAS Article](#) [ADS](#) [Google Scholar](#)

11. 11.

Li, Y. et al. Structural insights into immunoglobulin M. *Science* **367**, 1014–1017 (2020).

[CAS Article](#) [ADS](#) [Google Scholar](#)

12. 12.

Vonarburg, C. et al. Topical application of nebulized human IgG, IgA and IgAM in the lungs of rats and non-human primates. *Respir. Res.* **20**, 99 (2019).

[Article](#) [Google Scholar](#)

13. 13.

Wang, Z. et al. Enhanced SARS-CoV-2 neutralization by dimeric IgA. *Sci. Transl. Med.* **13**, eabf1555 (2021).

[CAS Article](#) [Google Scholar](#)

14. 14.

Galimidi, R. P. et al. Intra-spike crosslinking overcomes antibody evasion by HIV-1. *Cell* **160**, 433–446 (2015).

[CAS Article](#) [Google Scholar](#)

15. 15.

ter Meulen, J. et al. Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS Med.* **3**, e237 (2006).

[Article](#) [Google Scholar](#)

16. 16.

Weitzner, B. D. et al. Modeling and docking of antibody structures with Rosetta. *Nat. Protocols* **12**, 401–416 (2017).

[CAS Article](#) [Google Scholar](#)

17. 17.

Thomson, E. C. et al. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**, 1171–1187 (2021).

[CAS Article](#) [Google Scholar](#)

18. 18.

Liu, Z. et al. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe* **29**, 477–488 (2021).

[CAS Article](#) [Google Scholar](#)

19. 19.

FDA. Fact sheet for health care providers. Emergency use authorization (EUA) of casirivimab and imdevimab, <https://www.fda.gov/media/143892/download> (2020).

20. 20.

FDA. Fact sheet for health care providers. Emergency use authorization (EUA) of bamlanivimab, <https://www.fda.gov/media/143603/download> (2020).

21. 21.

Hou, Y. J. et al. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **182**, 429–446 (2020).

[CAS Article](#) [Google Scholar](#)

22. 22.

Wang, W. et al. Detection of SARS-CoV-2 in different types of clinical specimens. *J. Am. Med. Assoc.* **323**, 1843–1844 (2020).

[CAS](#) [Article](#) [ADS](#) [Google Scholar](#)

23. 23.

Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).

[CAS](#) [Article](#) [ADS](#) [Google Scholar](#)

24. 24.

Gong, S. et al. Anti-HIV IgM protects against mucosal SHIV transmission. *AIDS* **32**, F5–F13 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Keyt, B. A., Baliga, R., Sinclair, A. M., Carroll, S. F. & Peterson, M. S. Structure, function, and therapeutic use of IgM antibodies. *Antibodies* **9**, 53 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Muruato, A. et al. Mouse-adapted SARS-CoV-2 protects animals from lethal SARS-CoV-2 challenge. Preprint at <https://doi.org/10.1101/2021.05.03.442357> (2021).

27. 27.

Xie, X. et al. An infectious cDNA clone of SARS-CoV-2. *Cell Host Microbe* **27**, 841–848 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Xie, X. et al. Engineering SARS-CoV-2 using a reverse genetic system. *Nat. Protocols* **16**, 1761–1784 (2021).

[CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Vorauer-Uhl, K., Wallner, J., Lhota, G., Katinger, H. & Kunert, R. IgM characterization directly performed in crude culture supernatants by a new simple electrophoretic method. *J. Immunol. Methods* **359**, 21–27 (2010).

[CAS Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank W. Strohl for his scientific advice; J. Tang, N. Vora and Y. Ng for their technical support; and J. Jarecki and G. Salazar for their editorial support. This work was supported in part by a Welch Foundation grant AU-0042-20030616 and Cancer Prevention and Research Institute of Texas (CPRIT) grants RP150551 and RP190561 (Z.A.); NIH grants AI134907 and UL1TR001439, awards from the Sealy Smith Foundation, Kleberg Foundation, John S. Dunn Foundation, Amon G. Carter Foundation, Gillson Longenbaugh Foundation and Summerfield Robert Foundation (P.-Y.S.); and NIH grant R15CA182769 and CPRIT RP150656 (X.L.).

## Author information

### Author notes

1. These authors contributed equally: Zhiqiang Ku, Xuping Xie, Paul R. Hinton, Xinli Liu

### Affiliations

1. Texas Therapeutics Institute, Brown Foundation Institute of Molecular Medicine, The University of Texas Health Science Center at Houston, Houston, TX, USA

Zhiqiang Ku, Xiaohua Ye, Hui Deng, Wei Xiong, Junquan Liu, Hang Su, Ningyan Zhang & Zhiqiang An

2. Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX, USA

Xuping Xie, Jing Zou, Yang Liu & Pei-Yong Shi

3. IGM Biosciences, Mountain View, CA, USA

Paul R. Hinton, Dean C. Ng, Deepal Pandya, Sachi Rahman, Yu-An Cao, Kevin B. Carlin, Elizabeth J. Haanes, Bruce A. Keyt & Stephen F. Carroll

4. Department of Pharmacological and Pharmaceutical Sciences, College of Pharmacy, University of Houston, Houston, TX, USA

Xinli Liu & Sujit Biswas

5. Institute for Human Infection and Immunity, University of Texas Medical Branch, Galveston, TX, USA

Antonio E. Muruato & Pei-Yong Shi

6. Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, TX, USA

Antonio E. Muruato & Vineet D. Menachery

7. Sealy Institute for Vaccine Sciences, University of Texas Medical Branch, Galveston, TX, USA

Pei-Yong Shi

8. Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, TX, USA

Pei-Yong Shi

9. Department of Pharmacology and Toxicology, University of Texas Medical Branch, Galveston, TX, USA

Pei-Yong Shi

## Authors

1. Zhiqiang Ku

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Xuping Xie

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Paul R. Hinton  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Xinli Liu  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Xiaohua Ye  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Antonio E. Muruato  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Dean C. Ng  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Sujit Biswas  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Jing Zou  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Yang Liu  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Deepal Pandya  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Vineet D. Menachery  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Sachi Rahman  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Yu-An Cao  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Hui Deng  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Wei Xiong  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Kevin B. Carlin  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Junquan Liu  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Hang Su  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Elizabeth J. Haanes  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Bruce A. Keyt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Ningyan Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Stephen F. Carroll

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Pei-Yong Shi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Zhiqiang An

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

Z.K. isolated the IgG1 monoclonal antibodies, characterized the binding avidities and blocking activities of monoclonal antibodies and performed structural analysis. Z.K. and X.Y. cloned the RBD constructs and purified the proteins, and performed the bioinformatics analysis of virus sequences. H.D. and W.X. provided support with cell culture and transfection. J.L. and H.S. provided support with protein purifications. X.X., A.E.M. and V.D.M. generated mouse-adapted viruses and performed mouse protection studies. X.X. and J.Z. performed neutralization assays. X.X. and Y.L. generated SARS-CoV-2 variants. P.R.H., D.C.N., D.P., S.R. and K.B.C. engineered and produced the IgA1 and IgM monoclonal antibodies, and performed size-exclusion chromatography and RBD ELISA binding. Y.-A.C. performed pharmacokinetics and safety studies. X.L. and S.B. performed antibody bio-distribution. E.J.H., N.Z., B.A.K., S.F.C., P.-Y.S. and Z.A. supervised the study. Z.K. wrote the original draft with input from the team. X.X., P.R.H., X.L., X.Y., E.J.H., N.Z., B.A.K., S.F.C., P.-Y.S. and Z.A. reviewed and edited the manuscript.

## Corresponding authors

Correspondence to [Bruce A. Keyt](#) or [Ningyan Zhang](#) or [Stephen F. Carroll](#) or [Pei-Yong Shi](#) or [Zhiqiang An](#).

## Ethics declarations

### Competing interests

The University of Texas System has filed a patent on the SARS-CoV-2 antibodies and the reverse genetic system and reporter SARS-CoV-2. The University of Texas System and IGM Biosciences have filed a joint patent on the SARS-CoV-2 IgM antibodies. The antibodies are being developed by IGM Biosciences for prophylactic and therapeutic treatment of COVID-19. P.R.H., D.C.N., D.P., S.R., Y-A.C., K.B.C., E.J.H., B.A.K. and S.F.C. are employees of IGM Biosciences. The other authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Galit Alter, Andrianus Boon 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 The epitopes of IgG1 monoclonal antibodies used for engineering and additional characterizations of IgM and IgA1.](#)

**a**, The RBD is shown as a cartoon with the core region coloured in grey and the receptor-binding motif (RBM) coloured in red. **b**, Summary of the binding region on the RBD, key epitope residues and cross-reactivity to SARS-CoV of the six monoclonal antibodies. These epitope residues were mapped using an alanine-scanning RBD mutant library in a previous study<sup>1</sup>. **c**, ELISA binding to RBD by IgG1, IgA1 and IgM isotypes of the indicated monoclonal antibodies. Data are mean of duplicate wells. **d**, Neutralization of live SARS-CoV-2 by IgG1, IgA1 and IgM monoclonal antibodies at 1 µg ml<sup>-1</sup>. Data are mean ± s.d. of triplicates.

### [Source data](#)

## Extended Data Fig. 2 Binding and neutralization characterizations of IgM-06, IgG-06, IgA1-06 and IgA1-14.

**a**, ELISA binding to the spike protein (S) by IgM-06 and IgG-06. Data are mean of duplicate wells. **b**, Binding kinetics of IgM-06 and IgG-06 to the spike protein. **c**, Neutralization of SARS-CoV-2 by IgM-06 and IgG-06. Data are mean of duplicate wells. **d**, Summary of binding EC<sub>50</sub>, association ( $K_{on}$ ), dissociation ( $K_{dis}$ ), avidity ( $K_D$ ) and neutralization NT<sub>50</sub> values of IgM-06 and IgG-06. **e**, Neutralization of SARS-CoV-2 by IgA1-06 and IgA1-14. Data are mean of duplicate wells.

[Source data](#)

## Extended Data Fig. 3 Structural docking of the Fv–RBD complex.

**a, b**, Docking of IgFv-14–RBD (**a**) and IgFv-06–RBD (**b**) complex structures. The RBD is shown as a cartoon and coloured in grey. The Fv is shown as a surface with Vh coloured in cyan and VI coloured in magenta. Antibody epitope residues are shown as sticks and coloured in blue.

## Extended Data Fig. 4 Antibody blocking of RBD and ACE2 interaction as measured by a BLI assay.

**a**, Schematic diagram showing a BLI assay for IgM and IgG blocking of RBD and ACE2 interaction. **b**, A representative binding response curve of the BLI assay. The vertical dashed lines indicate the separation of each binding phase. **c, d**, The normalized ACE2 response curves after blocking by IgM-14 and IgG-14 (**c**) and by IgM-06 and IgG-06 (**d**). **e**, Superposition of IgFv-06–RBD and ACE2–RBD complexes. The IgFv-06 is shown as a surface with Vh coloured in cyan and VI coloured in magenta. The RBD–ACE2 complex is shown as a cartoon with RBD coloured in grey and ACE2 coloured in green. The dashed box indicates steric clash. **f**, IgM-06 and IgG-06 blocking of RBD and ACE2 interaction. The horizontal dashed line indicates 100% blocking.

[Source data](#)

## Extended Data Fig. 5 Construction of SARS-CoV-2 escaping variants and additional neutralization characterizations.

**a**, Schematic diagram showing the construction of the indicated mNeonGreen SARS-CoV-2 viruses using an infectious clone method. **b**, Plaque morphologies of mNeonGreen SARS-CoV-2 viruses with wild-type RBD or the indicated RBD

mutations. **c**, Neutralization of SARS-CoV-2 viruses with K444R (IgG-06-resistant), E484A (IgG-14-resistant) and K444R + E484A (IgG-06 + IgG-14-resistant) mutations by IgM-14 and IgG-14 on A549-ACE2 cells. Data are mean of duplicate wells. **d**, Summary of the  $NT_{50}$  values against indicated mutant viruses and the fold changes of  $NT_{50}$  values between IgM-14 and IgG-14. **e**, Neutralization of indicated SARS-CoV-2 mutant viruses by IgM-06 and IgG-06 on Vero cells. Data are mean of duplicate wells. **f**, Summary of the  $NT_{50}$  values against indicated mutant viruses and the fold changes of  $NT_{50}$  values between IgM-06 and IgG-06. NA, not available.

[Source data](#)

### **Extended Data Fig. 6 Binding kinetics and ACE2-blocking activities of IgM-14, IgG-14, IgM-06 and IgG-06 against selected RBD mutants.**

**a–c**, The binding kinetics of IgM-14 and IgG-14 to wild-type RBD (**a**), E484A RBD (**b**) and K444R + E484A RBD (**c**). **d–f**, The binding kinetics of IgM-06 and IgG-06 to wild-type RBD (**d**), K444R RBD (**e**) and K444R+E484A RBD (**f**). **g, h**, IgM-14 and IgG-14 blocking of E484A RBD (**g**) and K444R + E484A RBD (**h**) interaction with ACE2. **i, j**, IgM-06 and IgG-06 blocking of K444R RBD (**i**) and K444R + E484A RBD (**j**) interaction with ACE2. **k**, Summary of the binding avidities ( $K_D$ ) and ACE2-blocking activities ( $IC_{50}$ ) to indicated RBD proteins by IgM-06 and IgG-06. ND, not determined. \*Half-maximal blocking was not achieved at the highest monoclonal antibody concentration (30 nM) and the  $IC_{50}$  values are defined as  $\geq 90$  nM. **l, m**, The binding kinetics of IgM-14 and IgG-14 to F486S RBD (**l**) and IgM-06 and IgG-06 to K444S RBD (**m**). **n, o**, IgM-14 and IgG-14 blocking of F486S RBD interaction with ACE2 (**n**) and IgM-06 and IgG-06 blocking of K444S RBD interaction with ACE2 (**o**).

[Source data](#)

### **Extended Data Fig. 7 Binding kinetics of IgM-14 and IgG-14 to the 19 RBD mutants.**

**a**, SDS-PAGE images of six RBD proteins with mutations that represent natural escape variants in circulation. **b**, SDS-PAGE image of thirteen RBD proteins with mutations associated with neutralization-resistance to LY-CoV555, REGN-10933 and REGN-10987. The SDS-PAGE gel images were from one experiment. **c–u**, Binding kinetics of IgM-14 (left) and IgG-14 (right) to the indicated RBD mutants. **v**, Fold changes of  $K_{dis}$  and  $K_{on}$  between binding of IgM-14 and IgG-14 to the 19 RBD mutants. The ratios were calculated as indicated in the  $y$  axis.

[Source data](#)

**Extended Data Fig. 8 Antibody blocking of interactions between ACE2 and the 19 RBD mutants, the frequency of RBD mutations and mutational effects on RBD functionality.**

**a–s**, The dose-dependent blocking of the interactions between ACE2 and the indicated RBD mutants. **t**, The correlation between ACE2 blocking IC<sub>50</sub> and neutralization NT<sub>50</sub> values. Two-tailed Spearman correlation was used in the statistical analysis. **u**, The frequency of SARS-CoV-2 circulating variants with indicated RBD mutations. The Tracking Mutations tool was used for analysis as described in the methods. **v**, The effects of indicated RBD mutations on RBD binding affinity to ACE2 and RBD protein expression. The online web source on the sequence-to-phenotype maps of the RBD of SARS-CoV-2 ([https://jbloomlab.github.io/SARS-CoV-2-RBD\\_DMS/](https://jbloomlab.github.io/SARS-CoV-2-RBD_DMS/)) was referred to for analysis. The y axis indicates log<sub>10</sub> scale changes to wild-type RBD. Positive values indicate an improving effect and negative values indicate a decreasing effect by the mutations.

[Source data](#)

**Extended Data Fig. 9 Bio-distribution of AF750-labelled IgM-14 after intranasal delivery in mice.**

**a**, Whole-body imaging of four mice at different time points after a single intranasal dose of IgM-14. **b**, Ex vivo imaging of blood (20 µl) and different organs. **c**, Quantification of average radiant efficiency of blood and indicated organs. Data are mean ± s.d. of four independent mice.

[Source data](#)

**Extended Data Fig. 10 Additional in vivo characterizations of antibody efficacy, pharmacokinetics and safety.**

**a**, Experimental design of prophylactic evaluation of the indicated monoclonal antibodies.  $n = 4$  independent mice for all groups. **b**, Virus PFU titres in the lung samples of mice prophylactically treated with the indicated monoclonal antibodies. **c**, Virus RNA (N gene) titres in the lung samples of mice prophylactically treated with IgM-14. The cut-off for the qRT–PCR method, shown as dotted line, is defined as mean + 2 standard deviations of corresponding RNA copies in the qRT–PCR using lung samples from five uninfected mice. **d**, Viral loads (PFU titres) in the lung samples of mice therapeutically treated with IgM-14 or IgG-14 at the indicated doses. The lines

of median lung viral loads are shown for each group.  $n = 10$  biologically independent mice for all groups except that  $n = 5$  for IgM-14 group. A two-sided Mann–Whitney test was used in the statistical analysis for **b**, **d**. An ordinary one-way ANOVA with Sidak's multiple comparisons was used in the statistical analysis for **c**, **e**. Sequencing analysis of viruses recovered from lung samples of the ten most outlier mice. A representative chromatogram representing the amino acids 483–489 of the RBD is shown to indicate that no mutations of the critical residues E484 and F486 were observed. **f**, The plasma concentrations of IgM-14 after a single intranasal dose of 5 mg kg<sup>-1</sup> in BALB/c mice. Data are mean  $\pm$  s.d. of three independent mice. The values lower than LLOQ (0.02 µg ml<sup>-1</sup>) were defined as 0.01. **g**, Body weight changes of rats after intranasal administrations of 2 mg kg<sup>-1</sup> per dose of IgM-14 or the vehicle control. Data are mean  $\pm$  s.d. of four independent rats. The arrows indicate dosing twice daily for five consecutive days. Statistical differences between IgM-14 and vehicle groups were analysed by a two-sided multiple *t*-test. ns,  $P \geq 0.05$ .

## [Source data](#)

## Supplementary information

### [Supplementary Figure 1](#)

The uncropped gel electrophoresis images. The uncropped images for Fig. 1c are shown on the top panel and the uncropped images for Extended Data Fig. 9 are shown on the bottom pane. The black boxes indicate the regions that were cropped for making the corresponding figures.

### [Reporting Summary](#)

### [Supplementary Tables 1-2](#)

Supplementary Table 1 contains a summary of binding kinetics and ACE2 blocking IC50s to the RBD mutants and Supplementary Table 2 contains a list of primers used for the construction of RBD mutants.

## Source data

### [Source Data Fig. 1](#)

### [Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 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

Ku, Z., Xie, X., Hinton, P.R. *et al.* Nasal delivery of an IgM offers broad protection from SARS-CoV-2 variants. *Nature* **595**, 718–723 (2021).  
<https://doi.org/10.1038/s41586-021-03673-2>

[Download citation](#)

- Received: 23 February 2021
- Accepted: 26 May 2021
- Published: 03 June 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03673-2>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

## Associated Content

Collection

[\*\*Coronavirus\*\*](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03673-2>

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

- Article
- [Published: 07 July 2021](#)

# Metabolic control of T<sub>FH</sub> cells and humoral immunity by phosphatidylethanolamine

- [Guotong Fu](#)<sup>1</sup>,
- [Clifford S. Guy](#)<sup>1</sup>,
- [Nicole M. Chapman](#)<sup>1</sup>,
- [Gustavo Palacios](#)<sup>1</sup>,
- [Jun Wei](#)<sup>1</sup>,
- [Peipei Zhou](#)<sup>1</sup>,
- [Lingyun Long](#)<sup>1</sup>,
- [Yong-Dong Wang](#) [ORCID: orcid.org/0000-0001-8751-9216](#)<sup>2</sup>,
- [Chenxi Qian](#)<sup>1,2</sup>,
- [Yogesh Dhungana](#)<sup>1</sup>,
- [Hongling Huang](#)<sup>1</sup>,
- [Anil KC](#)<sup>1</sup>,
- [Hao Shi](#) [ORCID: orcid.org/0000-0002-2796-629X](#)<sup>1</sup>,
- [Sherri Rankin](#)<sup>1</sup>,
- [Scott A. Brown](#)<sup>1</sup>,
- [Amanda Johnson](#)<sup>3</sup>,
- [Randall Wakefield](#) [ORCID: orcid.org/0000-0002-6233-8109](#)<sup>3</sup>,
- [Camenzind G. Robinson](#) [ORCID: orcid.org/0000-0002-7277-692X](#)<sup>3</sup>,
- [Xueyan Liu](#)<sup>4</sup>,
- [Anthony Sheyn](#)<sup>5,6</sup>,
- [Jiyang Yu](#) [ORCID: orcid.org/0000-0003-3629-4330](#)<sup>2</sup>,
- [Suzanne Jackowski](#) [ORCID: orcid.org/0000-0002-6855-1429](#)<sup>7</sup> &

- [Hongbo Chi](#)   [ORCID: orcid.org/0000-0002-9997-2496<sup>1</sup>](#)

[Nature](#) volume **595**, pages 724–729 (2021)[Cite this article](#)

- 6214 Accesses
- 96 Altmetric
- [Metrics details](#)

## Subjects

- [Follicular T-helper cells](#)
- [Lymphocyte differentiation](#)

## Abstract

T follicular helper ( $T_{FH}$ ) cells are crucial for B cell-mediated humoral immunity<sup>1</sup>. Although transcription factors such as BCL6 drive the differentiation of  $T_{FH}$  cells<sup>2,3</sup>, it is unclear whether and how post-transcriptional and metabolic programs enforce  $T_{FH}$  cell programming. Here we show that the cytidine diphosphate (CDP)–ethanolamine pathway co-ordinates the expression and localization of CXCR5 with the responses of  $T_{FH}$  cells and humoral immunity. Using *in vivo* CRISPR–Cas9 screening and functional validation in mice, we identify ETNK1, PCYT2, and SELENOI—enzymes in the CDP–ethanolamine pathway for de novo synthesis of phosphatidylethanolamine (PE)—as selective post-transcriptional regulators of  $T_{FH}$  cell differentiation that act by promoting the surface expression and functional effects of CXCR5.  $T_{FH}$  cells exhibit unique lipid metabolic programs and PE is distributed to the outer layer of the plasma membrane, where it colocalizes with CXCR5. De novo synthesis of PE through the CDP–ethanolamine pathway co-ordinates these events to prevent the internalization and degradation of CXCR5. Genetic deletion of *Pcyt2*, but not of *Pcyt1a* (which mediates the CDP–choline pathway), in activated T cells impairs the differentiation of  $T_{FH}$  cells, and

this is associated with reduced humoral immune responses. Surface levels of PE and CXCR5 expression on B cells also depend on *Pcyt2*. Our results reveal that phospholipid metabolism orchestrates post-transcriptional mechanisms for T<sub>FH</sub> cell differentiation and humoral immunity, highlighting the metabolic control of context-dependent immune signalling and effector programs.

## 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

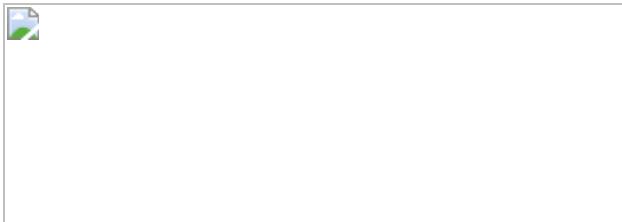
[Rent or Buy](#)

All prices are NET prices.

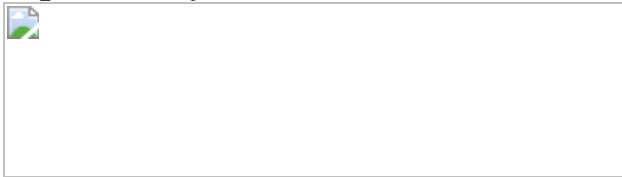
### Additional access options:

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

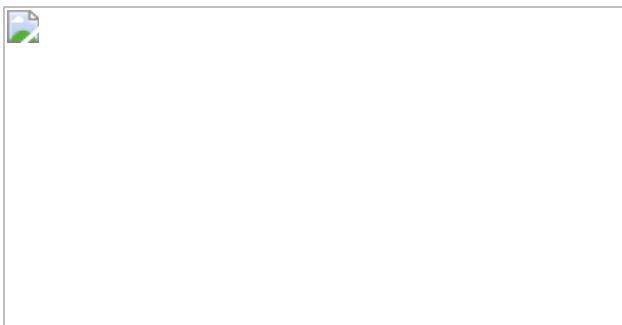
**Fig. 1: In vivo CRISPR–Cas9 screening reveals that the CDP–ethanolamine pathway is critical for the differentiation of T<sub>FH</sub> cells.**



**Fig. 2: The CDP–ethanolamine pathway regulates CXCR5 expression independently of BCL6.**



**Fig. 3: PE is selectively distributed on the outer layer of the T<sub>FH</sub> membrane and stabilizes CXCR5.**



**Fig. 4: Genetic ablation of the CDP–ethanolamine pathway disrupts humoral immunity.**



## Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its Supplementary Information files. Microarray data of wild-type and *Pcyt2*-deficient cells at day 3 after LCMV infection have been deposited in the NCBI Gene Expression Omnibus (GEO) database and are accessible through the GEO SuperSeries accession number: [GSE147190](#). For GSEA using manually curated gene signatures from public datasets, the datasets [GSE21380](#) (ref. <sup>9</sup>), [GSE72188](#) (ref. <sup>17</sup>) and GSE74854 (ref. <sup>37</sup>) from the NCBI GEO database were used to generate the relevant gene signatures (<5% FDR). [Source data](#) are provided with this paper.

## Code availability

The code used for the colocalization index analysis of TIRF–STORM or other co-ordinate-based data has been applied successfully<sup>42</sup> and the accessible link of the code is <https://rdrr.io/cran/colocalization/>.

## References

1. 1.

Crotty, S. T follicular helper cell biology: a decade of discovery and diseases. *Immunity* **50**, 1132–1148 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

2. 2.

Johnston, R. J. et al. Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation. *Science* **325**, 1006–1010 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

3. 3.

Nurieva, R. I. et al. Bcl6 mediates the development of T follicular helper cells. *Science* **325**, 1001–1005 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

4. 4.

Zeng, H. et al. mTORC1 and mTORC2 kinase signaling and glucose metabolism drive follicular helper T cell differentiation. *Immunity* **45**, 540–554 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

5. 5.

Chapman, N. M., Boothby, M. R. & Chi, H. Metabolic coordination of T cell quiescence and activation. *Nat. Rev. Immunol.* **20**, 55–70 (2020).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

6. 6.

Wei, J. et al. Targeting REGNASE-1 programs long-lived effector T cells for cancer therapy. *Nature* **576**, 471–476 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

7. 7.

Huang, H. et al. In vivo CRISPR screening reveals nutrient signaling processes underpinning CD8<sup>+</sup> T cell fate decisions. *Cell* **184**, 1245–1261 (2021).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

8. 8.

Xu, L. et al. The transcription factor TCF-1 initiates the differentiation of T<sub>FH</sub> cells during acute viral infection. *Nat. Immunol.* **16**, 991–999

(2015).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

9. 9.

Choi, Y. S. et al. LEF-1 and TCF-1 orchestrate T<sub>FH</sub> differentiation by regulating differentiation circuits upstream of the transcriptional repressor Bcl6. *Nat. Immunol.* **16**, 980–990 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

10. 10.

Boullaran, C. et al. B lymphocyte-specific loss of Ric-8A results in a G $\alpha$  protein deficit and severe humoral immunodeficiency. *J. Immunol.* **195**, 2090–2102 (2015).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

11. 11.

Preite, S. et al. Hyperactivated PI3K $\delta$  promotes self and commensal reactivity at the expense of optimal humoral immunity. *Nat. Immunol.* **19**, 986–1000 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

12. 12.

Lemonnier, F. et al. Recurrent *TET2* mutations in peripheral T-cell lymphomas correlate with T<sub>FH</sub>-like features and adverse clinical parameters. *Blood* **120**, 1466–1469 (2012).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

13. 13.

Johnston, R. J., Choi, Y. S., Diamond, J. A., Yang, J. A. & Crotty, S. STAT5 is a potent negative regulator of T<sub>FH</sub> cell differentiation. *J. Exp. Med.* **209**, 243–250 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

14. 14.

Calzada, E., Onguka, O. & Claypool, S. M. Phosphatidylethanolamine metabolism in health and disease. *Int. Rev. Cell Mol. Biol.* **321**, 29–88 (2016).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

15. 15.

Haynes, N. M. et al. Role of CXCR5 and CCR7 in follicular Th cell positioning and appearance of a programmed cell death gene-1<sup>high</sup> germinal center-associated subpopulation. *J. Immunol.* **179**, 5099–5108 (2007).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

16. 16.

Wang, L., Magdaleno, S., Tabas, I. & Jackowski, S. Early embryonic lethality in mice with targeted deletion of the CTP:phosphocholine cytidylyltransferase alpha gene (*Pcyt1a*). *Mol. Cell. Biol.* **25**, 3357–3363 (2005).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

17. 17.

Liu, X. et al. Genome-wide analysis identifies Bcl6-controlled regulatory networks during T follicular helper cell differentiation. *Cell Rep.* **14**, 1735–1747 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

18. 18.

Shi, J. et al. PD-1 controls follicular T helper cell positioning and function. *Immunity* **49**, 264–274 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

19. 19.

Zhao, M. Lantibiotics as probes for phosphatidylethanolamine. *Amino Acids* **41**, 1071–1079 (2011).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

20. 20.

Irie, A., Yamamoto, K., Miki, Y. & Murakami, M. Phosphatidylethanolamine dynamics are required for osteoclast fusion. *Sci. Rep.* **7**, 46715 (2017).

[PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

21. 21.

Burr, M. L. et al. CMTM6 maintains the expression of PD-L1 and regulates anti-tumour immunity. *Nature* **549**, 101–105 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

22. 22.

Zeng, H. et al. Discrete roles and bifurcation of PTEN signaling and mTORC1-mediated anabolic metabolism underlie IL-7-driven B lymphopoiesis. *Sci. Adv.* **4**, eaar5701 (2018).

[PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [CAS](#) [Google Scholar](#)

23. 23.

Platt, R. J. et al. CRISPR–Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

24. 24.

Oxenius, A., Bachmann, M. F., Zinkernagel, R. M. H. & Hengartner, H. Virus-specific MHC-class II-restricted TCR-transgenic mice: effects on humoral and cellular immune responses after viral infection. *Eur. J. Immunol.* **28**, 390–400 (1998).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

25. 25.

Leonardi, R., Frank, M. W., Jackson, P. D., Rock, C. O. & Jackowski, S. Elimination of the CDP-ethanolamine pathway disrupts hepatic lipid homeostasis. *J. Biol. Chem.* **284**, 27077–27089 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

26. 26.

Zhang, D. et al. Macrophages deficient in CTP:phosphocholine cytidylyltransferase- $\alpha$  are viable under normal culture conditions but are highly susceptible to free cholesterol-induced death. *J. Biol. Chem.* **275**, 35368–35376 (2000).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

27. 27.

Bettini, M. L., Bettini, M., Nakayama, M., Guy, C. S. & Vignali, D. A. Generation of T cell receptor-retrogenic mice: improved retroviral-mediated stem cell gene transfer. *Nat. Protocols* **8**, 1837–1840 (2013).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

28. 28.

Locci, M. et al. Human circulating PD-1<sup>+</sup>CXCR3<sup>-</sup>CXCR5<sup>+</sup> memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses. *Immunity* **39**, 758–769 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

29. 29.

Chen, R. et al. In vivo RNA interference screens identify regulators of antiviral CD4<sup>+</sup> and CD8<sup>+</sup> T cell differentiation. *Immunity* **41**, 325–338 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

30. 30.

Birsoy, K. et al. An essential role of the mitochondrial electron transport chain in cell proliferation is to enable aspartate synthesis. *Cell* **162**, 540–551 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

31. 31.

Sanson, K. R. et al. Optimized libraries for CRISPR–Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

32. 32.

Hanayama, R. et al. Autoimmune disease and impaired uptake of apoptotic cells in MFG-E8-deficient mice. *Science* **304**, 1147–1150 (2004).

[CAS](#) [PubMed](#) [Article](#) [ADS](#) [Google Scholar](#)

33. 33.

Todo, T. et al. Anti-CD3 $\varepsilon$  induces splenic B220 $lo$  B-cell expansion following anti-CD20 treatment in a mouse model of allosensitization. *Int. Immunol.* **24**, 529–538 (2012).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

34. 34.

Zeng, H. et al. mTORC1 couples immune signals and metabolic programming to establish T<sub>reg</sub>-cell function. *Nature* **499**, 485–490 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

35. 35.

Shrestha, S. et al. Treg cells require the phosphatase PTEN to restrain T<sub>H</sub>1 and T<sub>FH</sub> cell responses. *Nat. Immunol.* **16**, 178–187 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

36. 36.

Yu., D. et al. The transcriptional repressor Bcl-6 directs T follicular helper cell lineage commitment. *Immunity* **31**, 457–468 (2009).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

37. 37.

Shaw, L. A. et al. Id2 reinforces T<sub>H</sub>1 differentiation and inhibits E2A to repress T<sub>FH</sub> differentiation. *Nat. Immunol.* **17**, 834–843 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

38. 38.

Li, W. et al. Chimeric antigen receptor designed to prevent ubiquitination and downregulation showed durable antitumor efficacy. *Immunity* **53**, 456–470 (2020).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

39. 39.

Bates, M., Huang, B., Dempsey, G. T. & Zhuang, X. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science* **317**, 1749–1753 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

40. 40.

Guy, C. S. et al. Distinct TCR signaling pathways drive proliferation and cytokine production in T cells. *Nat. Immunol.* **14**, 262–270 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

41. 41.

Liedmann, S. et al. Viral suppressors of the RIG-I-mediated interferon response are pre-packaged in influenza virions. *Nat. Commun.* **5**, 5645 (2014).

[CAS](#) [PubMed](#) [Article](#) [ADS](#) [Google Scholar](#)

42. 42.

Samir, P. et al. DDX3X acts as a live-or-die checkpoint in stressed cells by regulating NLRP3 inflammasome. *Nature* **573**, 590–594 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

43. 43.

Malkusch, S. et al. Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochem. Cell Biol.* **137**, 1–10 (2012).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

44. 44.

Azuma, T. & Kei, T. Super-resolution spinning-disk confocal microscopy using optical photon reassignment. *Opt. Express* **23**, 15003–15011 (2015).

[CAS](#) [PubMed](#) [Article](#) [ADS](#) [Google Scholar](#)

45. 45.

Guagliardi, L. E. et al. Co-localization of molecules involved in antigen processing and presentation in an early endocytic compartment. *Nature* **343**, 133–139 (1990).

[CAS](#) [PubMed](#) [Article](#) [ADS](#) [Google Scholar](#)

46. 46.

Folch, J., Lees, M. & Sloane Stanley, G. H. A simple method for the isolation and purification of total lipids from animal tissues. *J. Biol. Chem.* **226**, 497–509 (1957).

[CAS](#) [Article](#) [Google Scholar](#)

47. 47.

Pasternak, C. A. & Bergeron, J. J. M. Turnover of mammalian phospholipids. *Biochem. J.* **119**, 473–480 (1970).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

48. 48.

Keckesova, Z. et al. LACTB is a tumour suppressor that modulates lipid metabolism and cell state. *Nature* **543**, 681–686 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [ADS](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

The authors acknowledge J. Saravia for human samples; Y. Wang and J. Raynor for editing of the manuscript; M. Hendren for animal colony management; the St. Jude Immunology FACS core facility for cell sorting; the Center for Advanced Genome Engineering (CAGE) for testing sgRNA deletion efficiency; and the Tetramer Facility of the NIH for gp66. Electron microscopy support was provided by the Cell and Tissue Imaging Center at St. Jude Children’s Research Hospital and by NCI P30 CA021765. This work was supported by NIH AI105887, AI131703, AI140761, AI150241, AI150514 and CA221290, by the Lupus Research Alliance, and by ALSAC. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

## Author information

### Affiliations

1. Department of Immunology, St. Jude Children’s Research Hospital, Memphis, TN, USA

Guotong Fu, Clifford S. Guy, Nicole M. Chapman, Gustavo Palacios, Jun Wei, Peipei Zhou, Lingyun Long, Chenxi Qian, Yogesh Dhungana, Hongling Huang, Anil KC, Hao Shi, Sherri Rankin, Scott A. Brown & Hongbo Chi

2. Department of Computational Biology, St. Jude Children’s Research Hospital, Memphis, TN, USA

Yong-Dong Wang, Chenxi Qian & Jiyang Yu

3. Cellular Imaging Shared Resource, St. Jude Children's Research Hospital, Memphis, TN, USA

Amanda Johnson, Randall Wakefield & Camenzind G. Robinson

4. Department of Mathematics, University of New Orleans, New Orleans, LA, USA

Xueyan Liu

5. Department of Surgery, Division of Otolaryngology, St. Jude Children's Research Hospital, Memphis, TN, USA

Anthony Sheyn

6. Department of Otolaryngology, College of Medicine, University of Tennessee Health Science Center, Memphis, TN, USA

Anthony Sheyn

7. Department of Infectious Diseases, St. Jude Children's Research Hospital, Memphis, TN, USA

Suzanne Jackowski

## Authors

1. Guotong Fu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Clifford S. Guy

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Nicole M. Chapman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Gustavo Palacios

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Jun Wei

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Peipei Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Lingyun Long

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Yong-Dong Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Chenxi Qian

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Yogesh Dhungana

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Hongling Huang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Anil KC

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Hao Shi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Sherri Rankin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Scott A. Brown

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Amanda Johnson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Randall Wakefield

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Camenzind G. Robinson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Xueyan Liu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Anthony Sheyn

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Jiyang Yu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Suzanne Jackowski

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Hongbo Chi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

G.F. designed and performed in vitro and in vivo experiments, analysed data and wrote the manuscript; C.S.G. performed imaging experiments and analysed the data with the help of S.R.; N.M.C. analysed human samples; G.P. performed lipidomic experiments; J.W. and L.L. developed the lentiviral sgRNA metabolic library; P.Z., H.H. and A.K. helped with cellular and molecular experiments; Y.-D.W., C.Q. and J.Y. analysed CRISPR–Cas9 screening data; Y.D., H.S. and X.L. performed bioinformatic analyses; S.A.B. helped with serum immunoglobulin analysis; A.J., R.W. and C.G.R. helped with electron microscopy experiments; A.S. provided human tonsil samples; S.J. provided the *Pcyt2<sup>fl</sup>* and *Pcyt1a<sup>fl</sup>* mice and contributed scientific insight and guidance on PE metabolism; H.C. designed experiments, co-wrote the manuscript and provided overall direction.

## Corresponding author

Correspondence to [Hongbo Chi](#).

## Ethics declarations

### Competing interests

H. Chi is a consultant for Kumquat Biosciences. The remaining authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Michael Dustin, Alessio Lanna 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 Pooled \*in vivo\* CRISPR–Cas9 screening and an \*in vivo\* dual transfer system to identify and validate potential regulators of T<sub>FH</sub> cells.](#)

**a**, Diagram of the screening system. Naive Cas9-expressing SMARTA cells were transduced with lentiviral-derived sgRNA metabolic library, expanded *in vitro* and transferred into C57BL/6 mice, which were then infected with LCMV 24 h later. At day 7 after infection, splenic T<sub>FH</sub> (CXCR5<sup>+</sup>SLAM<sup>-</sup>) and T<sub>H</sub>1 (CXCR5<sup>-</sup>SLAM<sup>+</sup>) cells were purified and those gRNAs that were upregulated (corresponding to negative regulators of T<sub>FH</sub> responses) or downregulated (corresponding to positive regulators of T<sub>FH</sub> responses) in T<sub>FH</sub> versus T<sub>H</sub>1 cells ( $|\log_2(T_{FH}/T_H)| > 0.5$ ; adjusted  $P < 0.05$ ) were

determined for metabolism-related genes that establish T<sub>FH</sub> over T<sub>H</sub>1 cell differentiation. **b**, Diagram of in vivo dual transfer system. SMARTA cells transduced with sgRNA viral vectors expressing distinct fluorescent proteins were mixed and transferred into the same C57BL/6 mice, followed by LCMV infection and experimental analyses. **c**, SMARTA cells transduced with the indicated sgRNA viral vectors (Ametrine<sup>+</sup>) were mixed at a 2:1 ratio with sgNTC (mCherry<sup>+</sup>)-transduced SMARTA cells and transferred into C57BL/6 mice, followed by LCMV infection. Analyses of the proportion of donor-derived T<sub>FH</sub> (CXCR5<sup>+</sup>SLAM<sup>-</sup> or PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) and T<sub>H</sub>1 (CXCR5<sup>-</sup>SLAM<sup>+</sup> or PSGL-1<sup>+</sup>Ly6C<sup>+</sup>) cells and quantification of relative T<sub>FH</sub> cell percentage and number (lower right) in the spleen at day 7 ( $n = 4$  mice). **d**, Insertion and deletion (indel) mutations after CRISPR–Cas9 targeted disruption in SMARTA cells transduced with sgNTC or sg*Pcyt2*, through deep sequencing analysis of indels generated at the exonic target site of the *Pcyt2* gene, including 74% of indel events in sg*Pcyt2*-transduced cells compared to 0.6% in sgNTC-transduced cells. **e**, Immunoblot analyses of ETNK1 and PCYT2 in splenic SMARTA cells at day 3 after infection. Asterisk (\*), non-specific band; arrow, the target band. Data are representative of two (**d**, **e**), or at least three (**c**) independent experiments. Data are mean  $\pm$  s.e.m. *P* values are determined by one-way ANOVA (**c**). \*\*\**P* < 0.001. Numbers in quadrants or gates indicate percentage of cells.

#### Source data

#### Extended Data Fig. 2 Validation of the effects of CDP–ethanolamine pathway genes on T<sub>FH</sub> generation during both viral infection and protein immunization.

**a, b**, Quantification of relative proportions and numbers of CXCR5<sup>+</sup>PD-1<sup>+</sup> T<sub>FH</sub> cells (**a**) and CXCR5<sup>-</sup>SLAM<sup>+</sup> T<sub>H</sub>1 cells (**b**) in donor-derived cells from the spleen of mice receiving the indicated sgRNA-transduced SMARTA cells at day 7 after LCMV infection ( $n = 4$  mice). **c**, Indel mutations after CRISPR–Cas9 targeted disruption in SMARTA cells transduced with the indicated sgRNAs through deep sequencing analysis of indels generated at

the exonic target sites of the indicated genes. **d**, Immunoblot analyses of the indicated proteins in the indicated sgRNA-transduced SMARTA cells isolated from the spleen at 3–5 days after adoptive transfer and LCMV infection. Asterisk (\*), non-specific band; arrow, the target band. **e**, SMARTA cells transduced with the indicated sgRNA-expressing vectors (Ametrine<sup>+</sup>) were mixed at a 2:1 ratio with sgNTC (mCherry<sup>+</sup>)-transduced SMARTA cells and transferred into C57BL/6 mice, followed by LCMV infection. Quantification of donor-derived T<sub>FH</sub> and T<sub>H</sub>1 cells from the host spleen at day 7 after infection ( $n = 4$  mice). **f**, OT-II cells transduced with the indicated sgRNAs were transferred into C57BL/6 mice, followed by immunization with NP-OVA and LPS in alum. Quantification of T<sub>FH</sub> cell percentage and number in the spleen at day 7 ( $n = 4$  mice). Data are representative of two (**c–e**), or at least three (**a, b, f**) independent experiments. Data are mean  $\pm$  s.e.m.  $P$  values are determined by one-way ANOVA (**a, b, e, f**). NS, not significant; \* $P < 0.05$ , \*\* $P < 0.01$  and \*\*\* $P < 0.001$ .

#### [Source data](#)

#### [\*\*Extended Data Fig. 3 Dynamic regulation of PE metabolic programs in T<sub>FH</sub> cells.\*\*](#)

**a**, Lipidomic analysis of the lipid content (PC and PS) in naive (CD4<sup>+</sup>) T cells, wild-type T<sub>FH</sub> (transduced with sgNTC), and *Pcyt2*-deficient T<sub>FH</sub> cells ( $n = 3$  samples, each pooled from multiple mice). **b, c**, The dependence on *Pcyt2* for PE alterations during T<sub>FH</sub> cell differentiation from naive CD4<sup>+</sup> T cells. **b**, Heat map showing the 39 significantly downregulated PE molecules in *Pcyt2*-deficient T<sub>FH</sub> cells as compared with wild-type T<sub>FH</sub> cells ( $\log_2$  ratio  $\geq 0.5$ ,  $P < 0.05$ ). **c**, Venn diagram showing the overlap of upregulated (left) or downregulated (right) PE molecules in wild-type T<sub>FH</sub> versus naive CD4<sup>+</sup> T cells compared to wild-type versus *Pcyt2*-deficient T<sub>FH</sub> cells ( $n = 3$  samples, each pooled from multiple mice). **d**, Lipidomic analysis of the lipid content (PE and PC) in wild-type and *Selenoi*-deficient T<sub>FH</sub> cells ( $n = 3$  samples, each pooled from multiple mice). **e**, Primary T<sub>FH</sub> and T<sub>H</sub>1 cells were sorted from LCMV-infected mice

and incubated with [<sup>3</sup>H]-Etn (2  $\mu$ Ci ml<sup>-1</sup>) for 3 h. To assess PE synthesis, lipids were extracted and [<sup>3</sup>H]-Etn incorporation into T<sub>FH</sub> cells was assessed with a scintillation counter (see synthesis column, left graph). To assess PE turnover, after 3 h of [<sup>3</sup>H]-Etn incubation, cells were washed and chased with unlabelled ethanolamine for another 3 h (see turnover column, left graph). The presence of [<sup>3</sup>H] radioactivity in cells and culture medium (right graph) was measured with a scintillation counter ( $n = 4$  samples, each pooled from multiple mice). CPM, counts per minute. Data are representative of at least two independent experiments (a–e). Data are mean  $\pm$  s.e.m.  $P$  values are determined one-way ANOVA (a) or by two-tailed unpaired Student's *t*-test (d, e). NS, not significant; \* $P < 0.05$ , \*\* $P < 0.01$  and \*\*\* $P < 0.001$ .

### Source data

### Extended Data Fig. 4 The CDP–ethanolamine pathway but not the CDP–choline pathway is required for T<sub>FH</sub> cell differentiation.

**a**, Diagram summary of the effects of single or combined deficiency of genes in the CDP–ethanolamine and CDP–choline pathways on T<sub>FH</sub> cell differentiation after LCMV infection. **b**, Immunoblot analyses of the indicated proteins in the indicated sgRNA-transduced SMARTA cells isolated from the spleen at 3–5 days after adoptive transfer and LCMV infection. Asterisk (\*), non-specific band; arrow, the target band. **c**, Summary of the relative proportions of T<sub>FH</sub> (CXCR5<sup>+</sup>SLAM<sup>-</sup>) and T<sub>H</sub>1 (CXCR5<sup>-</sup>SLAM<sup>+</sup>) cells in donor-derived cells in spleen of mice receiving the indicated sgRNA-transduced SMARTA cells at day 7 after LCMV infection ( $n = 4$  mice). **d**, Lipidomic analysis of the lipid content (PC and PE) in wild-type T<sub>FH</sub> (transduced with sgNTC), *Pcyt1a*-deficient T<sub>FH</sub> and *Pcyt1a* and *Pcyt1b* doubly deficient T<sub>FH</sub> cells ( $n = 3$  samples, each pooled from multiple mice). **e**, Flow cytometry analysis (left) and summary of the proportion of wild-type and *Pcyt2*-deficient CXCR5<sup>+</sup> SMARTA cells (middle) and their dilution of CellTrace Violet (CTV; right) at day 2 after LCMV infection ( $n = 3$  mice). **f**, Summary of apoptotic wild-type and

*Pcyt2*-deficient SMARTA cells as analysed by Annexin V and 7-AAD staining (or unstained) in freshly isolated splenocytes at day 2 after LCMV infection ( $n = 3$  mice). **g**, Distribution of SMARTA cells in the follicle in the spleen at day 3 after infection. Scale bar, 50  $\mu$ m ( $n = 50$  sections). **h**, Diagram of T<sub>FH</sub> cell effector functional assay. sgNTC (mCherry<sup>+</sup>) and sg*Pcyt2* (Ametrine<sup>+</sup>)-transduced SMARTA cells (CD45.1<sup>+</sup>) were transferred (first transfer) into C57BL/6 mice (CD45.2<sup>+</sup>), followed by LCMV infection. Seven days later, the fully differentiated wild-type or *Pcyt2*-deficient CXCR5<sup>+</sup>SLAM<sup>−</sup> T<sub>FH</sub> cells (CD45.1<sup>+</sup>) were sorted and equal numbers of these cells were transferred (second transfer) into LCMV-infected mice (CD45.2<sup>+</sup>; this transfer occurred at day 1 after LCMV infection). GC B cell and plasma cell formation was analysed at day 5 after the second adoptive transfer. Data are representative of one (**d**) or at least two (**b, c, e–g**) independent experiments. Data are mean  $\pm$  s.e.m. *P* values are determined by one-way ANOVA (**c, d**) or by two-tailed unpaired Student's *t*-test (**e, f**). NS, not significant; \*\**P* < 0.01 and \*\*\**P* < 0.001. Numbers in gates indicate percentage of cells.

#### Source data

#### Extended Data Fig. 5 The CDP–ethanolamine pathway regulates T<sub>FH</sub> cell differentiation independently of the BCL6–T-bet axis.

**a**, sgNTC or sg*Pcyt2*-transduced SMARTA cells (CD45.1<sup>+</sup>) were transferred into C57BL/6 recipients (CD45.2<sup>+</sup>) that were subsequently infected with LCMV. CD45.1<sup>+</sup> cells were isolated at day 3 for transcriptional profiling by microarray. GSEA of the T<sub>FH</sub> signature in *Pcyt2*-deficient compared to wild-type T cells ( $n = 3$  samples, each pooled from multiple mice). **b, c**, sgNTC, sg*Etnk1*, sg*Pcyt2* or sg*Selenoi*-transduced SMARTA cells were transferred into C57BL/6 recipients, followed by infection with LCMV. Quantification of the percentages of BCL6<sup>+</sup>T-bet<sup>−</sup> and BCL6<sup>−</sup>T-bet<sup>+</sup> cells among donor-derived T cells in the spleen at day 3 (**b**) and day 7 (**c**) after infection ( $n = 4$  mice). **d**, GSEA of BCL6 target genes (identified by ChIP-seq<sup>17</sup>) in *Pcyt2*-deficient compared

to wild-type T cells (samples are the same as Extended Data Fig. 5a;  $n = 3$  samples, each pooled from multiple mice). **e**, GSEA of the  $T_{FH}$ -specific genes that are directly downregulated by BCL6 (in total 48 genes, identified by combining ChIP-seq, RNA-seq and microarray datasets<sup>17</sup>) in *Pcyt2*-deficient compared to wild-type T cells (samples are the same as Extended Data Fig. 5a;  $n = 3$  samples, each pooled from multiple mice). **f**, Analysis and quantification of PSGL-1<sup>-</sup>Ly6C<sup>-</sup> cells, CXCR5<sup>+</sup> cells and CXCR5 MFI among PSGL-1<sup>-</sup>Ly6C<sup>-</sup>  $T_{FH}$  cells at day 3 after infection ( $n = 4$  mice). **g**, sgNTC, sg*Pcyt2*, sg*Bcl6*, or dual sg*Pcyt2* and sg*Bcl6*-transduced SMARTA cells were transferred into C57BL/6 recipients that were subsequently infected with LCMV. Quantification of the proportion of  $T_{FH}$  cells (CXCR5<sup>+</sup>PD-1<sup>+</sup> and CXCR5<sup>+</sup>SLAM<sup>-</sup>) in donor-derived cells from the spleen at day 7 after infection ( $n = 4$  mice). **h**, sgNTC, sg*Etnk1*, sg*Pcyt2* or sg*Selenoi*-transduced SMARTA cells (CD45.1<sup>+</sup>) were transferred into C57BL/6 recipients (CD45.2<sup>+</sup>) that were subsequently infected with LCMV. Quantification of the proportion of  $T_{FH}$  cells (CXCR5<sup>+</sup>PD-1<sup>+</sup>) in donor-derived T cells from the spleen at day 3 after infection ( $n = 4$  mice). **i-o**, SMARTA cells transduced with the indicated sgRNAs were transferred into C57BL/6 recipients, followed by infection with LCMV. **i**, Quantification of the frequency and MFI of CXCR5 on  $T_{FH}$  cells (PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) in the spleen at day 2 after infection ( $n = 3$  mice). **j**, Quantification of the frequency and MFI of PD-1 on  $T_{FH}$  cells (PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) in the spleen at day 2 after infection ( $n = 3$  mice). **k**, SMARTA cells were transduced with sgNTC or sg*Pcyt2* alone or in combination with a retrovirus overexpressing empty vector (Empty-RV) or CXCR5 (CXCR5-RV), followed by adoptive transfer into C57BL/6 mice that were then infected with LCMV. Analysis (top) and quantification (bottom) of the proportion of  $T_{FH}$  cells (CXCR5<sup>+</sup>PD-1<sup>+</sup> and PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) in donor-derived cells from the spleen at day 7 after infection. Numbers above the graphs indicate fold change ( $n = 4$  mice in CXCR5-RV group of CXCR5<sup>+</sup>PD-1<sup>+</sup>  $T_{FH}$  data,  $n = 3$  mice in other groups). **l**, MFI of PD-1, ICOS, Ly6C, PSGL-1, CD44 and CD62L on CD45.1<sup>+</sup> SMARTA cells at day 3 after infection ( $n = 4$  mice). **m**, Quantification of the frequency and MFI of CXCR3 on CD45.1<sup>+</sup> SMARTA cells at day 3 after infection ( $n = 3$  mice). **n**, MFI of CCR7 expression on donor-derived T cells from the host spleen at day 3 after infection ( $n = 3$

mice). **o**, Quantification of phosphorylated AKT (pAKT(S473)) levels in purified wild-type or *Pcyt2*-deficient T<sub>FH</sub> cells (CXCR5<sup>+</sup>SLAM<sup>-</sup>) that were stimulated with CXCL13 for 3 h ( $n = 4$  mice). Data are representative of one (**a, d, e**), at least two (**b, c, g, h**), or at least three (**f, i-o**) independent experiments. Data are mean  $\pm$  s.e.m.  $P$  values are determined by one-way ANOVA (**b, c, f-h, k, l**), or by two-tailed unpaired Student's *t*-test (**i, j, m-o**). NS, not significant; \* $P < 0.05$ , \*\* $P < 0.01$  and \*\*\* $P < 0.001$ . Numbers in quadrants or gates indicate percentage of cells.

### Source data

### Extended Data Fig. 6 PE is selectively distributed on the outer layer of the plasma membrane of T<sub>FH</sub> cells but not T<sub>H1</sub> cells.

**a**, Analysis and quantification of surface levels of PE on T<sub>FH</sub> (PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) and T<sub>H1</sub> (PSGL-1<sup>+</sup>Ly6C<sup>+</sup>) membrane at day 5 after infection ( $n = 4$  mice). **b**, Quantification of the colocalization index of PE and CCR7 or CXCR5 on PSGL-1<sup>-</sup>Ly6C<sup>-</sup> T<sub>FH</sub> cells in Fig. 3b using the CBC algorithm (PE and CCR7,  $n = 33$  cells; PE and CXCR5,  $n = 25$  cells). A colocalization index value of 1 indicates complete colocalization, a value of 0 represents spatial randomness and a value of -1 indicates intermolecular exclusion. **c**, Three-dimensional super resolution confocal image of the entire T<sub>FH</sub> cell on poly-l-lysine-coated coverslips. Surface staining was reconstructed using Imaris software. Scale bars, 2  $\mu$ m. Layers are translucent to readily visualize both colours. Right, quantification of colocalization index of PE and CCR7 or PE and CXCR5 on PSGL-1<sup>-</sup>Ly6C<sup>-</sup> T<sub>FH</sub> cells ( $n = 4$  mice and 20 cells were quantified). **d**, Confocal microscopy imaging of PE distribution on T<sub>FH</sub> (PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) and T<sub>H1</sub> (PSGL-1<sup>+</sup>Ly6C<sup>+</sup>) cells at the indicated time points after LCMV infection. Scale bar, 10  $\mu$ m ( $n = 3$  mice). **e, f**, Lipids associated with Flag-CXCR5 or Flag-CCR7 expressed in SMARTA cells isolated from LCMV-infected mice were immunoprecipitated by anti-Flag M2 magnetic beads and quantified using LC-MS/MS. Analysis of lipid content (PE, PC and PS) (**e**) and heat map showing significantly changed ( $P < 0.05$ ) PE molecules (**f**) in the indicated groups ( $n = 2$  samples, each pooled from multiple mice). Asterisks indicate

PE molecules with the exact mass. **g**, SMARTA cells were transduced with the indicated sgRNA and the retrovirus overexpressing empty vector (Empty-RV) or CXCR5 (CXCR5-RV), followed by adoptive transfer into C57BL/6 mice and LCMV infection. Quantification of the proportion of PE<sup>+</sup> cells in PSGL-1<sup>-</sup>Ly6C<sup>-</sup> T<sub>FH</sub> cells in donor-derived cells from the spleen at day 7. Numbers above the graphs indicate fold change ( $n = 4$  mice). **h**, SMARTA cells were transduced with the indicated sgRNAs, followed by adoptive transfer into C57BL/6 mice and LCMV infection. Quantification of the proportion of PE<sup>+</sup> cells in donor-derived PSGL-1<sup>-</sup>Ly6C<sup>-</sup> T<sub>FH</sub> cells from the spleen at day 7 after infection ( $n = 4$  mice). **i**, Analysis and quantification of PE outer-layer membrane distribution on freshly isolated wild-type and *Pcyt2*-deficient T<sub>FH</sub> (PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) and T<sub>H</sub>1 (PSGL-1<sup>+</sup>Ly6C<sup>+</sup>) cells (T<sub>FH</sub>,  $n = 5$  mice; T<sub>H</sub>1,  $n = 4$  mice). **j**, sgNTC or sg*Pcyt2*-transduced SMARTA cells were transferred into C57BL/6 recipients that were subsequently infected with LCMV. PS exposure in intact T<sub>FH</sub> (PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) and T<sub>H</sub>1 (PSGL-1<sup>+</sup>Ly6C<sup>+</sup>) membrane was assessed by Annexin V and 7-AAD (left) or MFG-E8 and 7-AAD (right) staining analysis at day 5 after infection ( $n = 4$  mice). Data are representative of two (**e–h, j**), or at least three (**a–d, i**) independent experiments. Data are mean  $\pm$  s.e.m. *P* values are determined by one-way ANOVA (**h, j**), or two-tailed unpaired Student's *t*-test (**a–c, g, i**). NS, not significant; \*\*\**P* < 0.001. Numbers in quadrants or gates indicate percentage of cells.

#### [Source data](#)

#### [Extended Data Fig. 7 PE regulates CDP–ethanolamine pathway-mediated T<sub>FH</sub> responses.](#)

**a**, Wild-type and *Pcyt2*-deficient T<sub>FH</sub> cells were sorted from LCMV-infected mice and incubated with [<sup>3</sup>H]-Etn (2  $\mu$ Ci ml<sup>-1</sup>) for 3 h. Lipids were extracted and [<sup>3</sup>H]-Etn incorporated into T<sub>FH</sub> cells was assessed with a scintillation counter ( $n = 4$  samples). **b–e**, Lipid add-back–rescue assay. The sgNTC and sg*Pcyt2*-transduced SMARTA cells were supplemented with the indicated lipids, followed by transfer of these cells to C57BL/6 mice and

LCMV infection. **b**, Diagram showing the time points of lipid treatment. PE was supplemented at days -5, -3 and -1 before adoptive transfer, and LCMV-infected mice were analysed at day 3 after infection (to capture the effects induced by in vitro PE supplementation). **c**, Lipidomic analysis showing the PE content in the indicated groups (left), and Venn diagram showing the rescued PE molecules by diacyl-type and ether-type PE supplementation (right) at day 3 after infection. **d**, Analysis and quantification of the proportion of PE<sup>+</sup> cells in donor-derived T<sub>FH</sub> cells (PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) from the spleen ( $n = 5$  mice). **e**, Analysis and quantification of the proportion of donor-derived CXCR5<sup>+</sup>SLAM<sup>-</sup> T<sub>FH</sub> cells from the spleen ( $n = 5$  mice). **f**, GSEA plot for KEGG signature of ABC transporters in T<sub>FH</sub> versus T<sub>H1</sub> cells (from a public dataset GSE74854 (ref. [37](#))). **g**, Heat map showing the top 22 leading-edge genes from the enrichment plot in **f**. Data are representative of two (**a, c**) or at least three (**d, e**) independent experiments. Data are mean  $\pm$  s.e.m. *P* values are determined by one-way ANOVA (**d, e**) or two-tailed unpaired Student's *t*-test (**a**). \**P* < 0.05, \*\**P* < 0.01 and \*\*\**P* < 0.001. Numbers in gates indicate percentage of cells.

#### [Source data](#)

### [Extended Data Fig. 8 PE stabilizes surface CXCR5 and prevents it from being targeted for lysosome-mediated degradation.](#)

**a**, The internalization of surface CXCR5 over time was traced by confocal imaging. **b**, Wild-type or *Pcyt2*-deficient SMARTA cells were isolated from LCMV-infected mice and labelled with unconjugated PD-1-specific antibody before incubation at 37 °C. The amount of surface PD-1 remaining over time was detected by fluorophore-conjugated secondary antibody staining ( $n = 3$  mice). **c**, Flow cytometry analysis of CXCR5 surface expression showed that proteinase K treatment (0.1 mg ml<sup>-1</sup>) efficiently removed the non-internalized CXCR5 on the cell surface. **d**, CXCR5 recycling assay. Cell-surface CXCR5 was labelled with unconjugated anti-CXCR5 antibody and incubated at 30 °C to allow internalization of antibody-labelled CXCR5. After washing, the remaining surface-bound

antibody was stripped by resuspension in proteinase K, followed by washing and incubation at different time points to allow CXCR5 recycling. The amount of surface CXCR5 recycled over time was detected by fluorophore-conjugated secondary antibody staining and analysed by flow cytometry (left). Right, quantification of percentage of recycled CXCR5 (relative ratio to untreated cells) ( $n = 4$  mice). **e**, Validation of anti-CXCR5 antibody (clone number, EPR23463-30) by immunoblot analysis of CXCR5 expression in wild-type and CXCR5-deficient SMARTA cells that were isolated from LCMV-infected mice. **f**, Immunoblot analysis of CXCR5 and PCYT2 expression in wild-type and *Pcyt2*-deficient SMARTA cells that were isolated from LCMV-infected mice and treated with the indicated concentrations of BafA1 for 8 h. **g**, Immunoblot analysis of CXCR5 expression in wild-type and *Pcyt2*-deficient SMARTA cells isolated from LCMV-infected mice after treatment with the proteasome inhibitor MG-132. **h**, Flow cytometry analysis of CXCR5 surface expression on wild-type and *Pcyt2*-deficient SMARTA cells that were isolated from LCMV-infected mice and treated with the indicated concentrations of BafA1 for 8 h. **i**, Quantification of surface PE levels on human CXCR5 $^{+}$  memory T<sub>FH</sub>, CXCR5 $^{+}$  central memory (T<sub>CM</sub>) and CD45RA $^{+}$  naive-like T cells in peripheral blood ( $n = 4$  donors). **j**, Analysis and quantification of surface PE levels on human tonsil CXCR5 $^{+}$ PD-1 $^{\text{hi}}$  T<sub>FH</sub> cells, transitional CXCR5 $^{+}$ PD-1 $^{+}$  T<sub>FH</sub> cells, CXCR5 $^{-}$  non-T<sub>FH</sub> and CD45RA $^{+}$  naive-like T cells ( $n = 8$  donors). Data are representative of two (**b, e–h**) or at least three (**a, c, d, i, j**) independent experiments. *P* values are determined by one-way ANOVA (**i, j**) or two-tailed unpaired Student's *t*-test (**b, d**). NS, not significant; \**P* < 0.05 and \*\*\**P* < 0.001. Numbers in gates indicate percentage of cells.

#### Source data

#### Extended Data Fig. 9 Loss of *Pcyt2* in activated T cells reveals selective impairments of T<sub>FH</sub> accumulation that are associated with reduced GC responses.

**a**, *Pcyt2* mRNA expression in freshly isolated (day 0) naive CD4 $^{+}$  T cells from wild-type and OX40 $^{\text{Cre}}$ *Pcyt2* $^{\text{fl/fl}}$  mice, or after in vitro anti-CD3/CD28

antibody stimulation for the indicated times ( $n = 4$  samples). **b**, Analysis and quantification of T<sub>FH</sub> cells (CXCR5<sup>+</sup>BCL6<sup>+</sup> or CXCR5<sup>+</sup>PD-1<sup>+</sup>) among B220<sup>-</sup>CD4<sup>+</sup>TCR $\beta$ <sup>+</sup> cells in Peyer's patches (PPs) and mesenteric lymph nodes (mLNs) from wild-type and OX40<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice (8 weeks old; CXCR5<sup>+</sup>BCL6<sup>+</sup> T<sub>FH</sub> cells,  $n = 4$  mice; CXCR5<sup>+</sup>PD-1<sup>+</sup> T<sub>FH</sub> cells,  $n = 3$  mice). **c**, Analysis and quantification of GC B cells (Fas<sup>+</sup>GL7<sup>+</sup>) among B220<sup>+</sup>CD19<sup>+</sup> cells in PPs and mLNs from wild-type and OX40<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice ( $n = 3$  mice). **d**, Immunohistochemistry of GCs in the mLNs of wild-type and OX40<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice. Scale bar, 50  $\mu$ m ( $n = 3$  mice). **e**, *Pcyt1a* mRNA expression in freshly isolated (day 0) naive CD4<sup>+</sup> T cells from wild-type and OX40<sup>Cre</sup>*Pcyt1a*<sup>fl/fl</sup> mice, or after in vitro stimulation with anti-CD3/CD28 antibodies for indicated times ( $n = 3$  samples). **f**, Analysis and quantification of T<sub>FH</sub> cells (CXCR5<sup>+</sup>BCL6<sup>+</sup> or CXCR5<sup>+</sup>PD-1<sup>+</sup>) among B220<sup>-</sup>CD4<sup>+</sup>TCR $\beta$ <sup>+</sup> cells in PPs and mLNs from wild-type and OX40<sup>Cre</sup>*Pcyt1a*<sup>fl/fl</sup> mice (8 weeks old;  $n = 3$  mice). **g**, Analysis and quantification of GC B cells (Fas<sup>+</sup>GL7<sup>+</sup>) among B220<sup>+</sup>CD19<sup>+</sup> cells in PPs and mLNs from wild-type and OX40<sup>Cre</sup>*Pcyt1a*<sup>fl/fl</sup> mice ( $n = 3$  mice). **h**, Mixed bone marrow chimaeras were constructed by mixing bone marrow cells from wild-type or OX40<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice and CD45.1<sup>+</sup> 'spike' mice followed by injection into sub-lethally irradiated *Rag1*<sup>-/-</sup> recipient mice. Quantification of T<sub>FH</sub> cells (CXCR5<sup>+</sup>BCL6<sup>+</sup>) among CD45.1<sup>+</sup>B220<sup>-</sup> CD4<sup>+</sup>TCR $\beta$ <sup>+</sup> or CD45.2<sup>+</sup>B220<sup>-</sup>CD4<sup>+</sup>TCR $\beta$ <sup>+</sup> cells in the PPs and mLNs under steady state ( $n = 3$  mice). **i**, Analysis and quantification of gp66-tetramer-positive CXCR5<sup>+</sup>PD-1<sup>+</sup> T<sub>FH</sub> cells in the spleen of wild-type and OX40<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice at day 7 after infection ( $n = 4$  mice). **j**, Retrogenic mouse-derived naive CD4<sup>+</sup> T cells deficient for *Etnk1* or *Selenoi* were transferred into C57BL/6 mice, followed by LCMV infection. Analyses of the proportion of donor-derived T<sub>FH</sub> (CXCR5<sup>+</sup>SLAM<sup>-</sup>, CXCR5<sup>+</sup>PD-1<sup>+</sup> or PSGL-1<sup>-</sup>Ly6C<sup>-</sup>) cells (left) and quantification of T<sub>FH</sub> cell percentage and number (right) in the spleen at day 7 after infection ( $n = 4$  mice). **k**, Quantification of numbers of T<sub>FH</sub> cells (CXCR5<sup>+</sup>PD-1<sup>+</sup>, CXCR5<sup>+</sup>ICOS<sup>+</sup> or CXCR5<sup>+</sup>Ly6C<sup>-</sup>) or GC B cells in the spleen from wild-type and OX40<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice at day 7 after intraperitoneal immunization with NP-

OVA and LPS in alum ( $n = 4$  mice). **I**, Measurements of anti-NP immunoglobulins in the serum from wild-type and OX40<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice at day 7 after immunization with NP-OVA and LPS ( $n = 16$ , collected from 8 mice). Data are representative of two (**a, e–l**) or at least three (**b–d**) independent experiments. Data are mean  $\pm$  s.e.m.  $P$  values are determined by one-way ANOVA (**j**) or by two-tailed unpaired Student's *t*-test (**b, c, f–i, k, l**). NS, not significant; \* $P < 0.05$ , \*\* $P < 0.01$  and \*\*\* $P < 0.001$ . Numbers in gates indicate percentage of cells.

### Source data

### Extended Data Fig. 10 Surface PE distribution and CXCR5 expression on B cells are dependent on the CDP–ethanolamine pathway.

**a**, Analysis and quantification of the proportions of CXCR5<sup>+</sup>PE<sup>+</sup> cells on mature conventional B cells (B220<sup>hi</sup>IgD<sup>hi</sup>IgM<sup>lo</sup>) or total B220<sup>+</sup>IgM<sup>+</sup> B cells from the spleen, immature B cells (B220<sup>+</sup>IgM<sup>+</sup> cells from the bone marrow (BM)), as well as splenic naive CD4<sup>+</sup> T cells ( $n = 3$  mice). **b**, Quantification of CXCR5 and PE MFI among B220<sup>+</sup> cells in spleen, peripheral lymph nodes (pLNs), mesenteric lymph nodes (mLNs) and Peyer's patches (PPs) of wild-type and *Cd19*<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice ( $n = 3$  mice). **c**, Proportion of migrated B cells from the spleen of wild-type and *Cd19*<sup>Cre</sup>*Pcyt2*<sup>fl/fl</sup> mice was assessed by flow cytometry after treatment with CXCL13 for 3 h ( $n = 3$  mice). **d**, Two-step model of T<sub>FH</sub> cell differentiation. Expression of BCL6 and other transcription factors in activated CD4<sup>+</sup> T cells orchestrates the initiation and commitment to the T<sub>FH</sub> program, leading to the induction of chemokine receptor CXCR5 (left cell). Right cell: In BCL6- and CXCR5-expressing T<sub>FH</sub> cells, the CDP–ethanolamine pathway (composed of ETNK1, PCYT2 and SELENOI), which mediates de novo PE synthesis, acts as a critical post-transcriptional program for the functional maintenance of the T<sub>FH</sub> program. Maintenance of the T<sub>FH</sub> program is essential for GC responses, plasma cell formation and antigen-specific immunoglobulin secretion. By contrast, the CDP–choline pathway that promotes PC synthesis and the PISD-dependent decarboxylation of PS are dispensable

for T<sub>FH</sub> accumulation, as indicated by grey shading. **e**, Mechanistically, the CDP–ethanolamine pathway controls T<sub>FH</sub> responses by stabilizing CXCR5 surface expression, CXCL13–CXCR5-mediated signalling events and cellular trafficking to B cell follicles. De novo synthesis of PE downstream of the CDP–ethanolamine pathway enables PE to interact with CXCR5 and prevents its targeting for lysosome-mediated degradation. Depletion of the CDP–ethanolamine pathway reduces the stability, accelerates the internalization rate and decreases the surface recycling of CXCR5. Data are representative of at least two (**a–c**) independent experiments. Data are mean  $\pm$  s.e.m. *P* values are determined by two-tailed unpaired Student's *t*-test (**b, c**). NS, not significant; \**P* < 0.05 and \*\**P* < 0.01. Numbers in gates indicate percentage of cells.

[Source data](#)

## Supplementary information

### [Supplementary Figure 1](#)

This file contains the uncropped Western blot images with size marker indications.

### [Reporting Summary](#)

### [Supplementary Table 1](#)

Analysis of mouse sgRNA metabolic library CRISPR screening data. This file contains the output of the analysis of mouse metabolic library CRISPR screening data at the gene (Tab a) and sgRNA (Tab b) levels, with the first row showing the label identifying each column. For data analysis, FastQ files obtained after sequencing were demultiplexed using the HiSeq Analysis software (Illumina). Single-end reads were trimmed and quality-filtered using the CLC Genomics Workbench v11 (Qiagen) and matched against sgRNA sequences from the sgRNA metabolic library. Read counts for sgRNAs were normalized against total read counts across all samples. For each sgRNA, the fold change ( $\log_2$ -transformed ratio) for enrichment

was calculated between each of the biological replicates and the input experiment. After merging the quantification results from two sub-libraries, candidate genes were ranked based on the average enrichment of their six gene-specific sgRNAs in Tfh (CXCR5<sup>+</sup>SLAM<sup>-</sup>) relative to Th1 (CXCR5<sup>-</sup> SLAM<sup>+</sup>) cells ( $\log_2$  ratio (Tfh/Th1); adjusted  $P < 0.05$ ), Tfh cells relative to input ( $\log_2$  ratio (Tfh/input); adjusted  $P < 0.05$ ), and Th1 cells relative to input ( $\log_2$  ratio (Th1/input); adjusted  $P < 0.05$ ). The gene level false discovery rate-adjusted  $P$  value was calculated among multiple sgRNAs ( $n = 6$ ) for each gene, using a two-tailed paired Student's *t*-test between  $\log_2$ -transformed average normalized read counts of Tfh, Th1 or input cells, and the  $P$  value was further adjusted using Bonferroni correction with gene size.

## **Supplementary Table 2**

Analysis of lipidomic data in naïve T cells and sgNTC or sg*Pcyt2*-transduced Tfh cells. This file contains the output of the analysis of the lipidomic data from wild-type Tfh, Pcyt2-deficient Tfh and naïve CD4<sup>+</sup> T cells, with the first row showing the label for each column. For data analysis, the Thermo Fisher Scientific LipidSearch software (version 4.1) was used for identification and relative quantification of lipids with the following parameters: precursor and product ion mass tolerance of  $\pm 5$  ppm; main adducts search (M+H, M-H, M+NH<sub>4</sub>, M+CH<sub>3</sub>COO, M+2H, M-2H, M+Na) for mass. All lipid sub-classes were searched within major lipid classes (phospholipids, sphingolipids, glycerolipids and neutral lipids). All individual data files were searched for product ion MS/MS spectra of lipid precursor ions, and the data were merged for each annotated lipid.

Compound Discoverer 3.1 (CD31; Thermo Fisher Scientific) software was used for multivariate statistical and differential expression analyses of individual molecules using a pre-defined workflow from CD31: untargeted metabolomics with statistics; detect ID of unknowns using online databases; map compounds to biological pathways using Metabolika. The mass-to-charge (m/z) data from positive or negative ion modes processed from CD31 was combined and then formatted to comma-separated value (CSV) file and imported into MetaboAnalyst 4.0. The peak areas were normalized using the parameters for sample normalization sum, data transformation

$\log_{10}$  and range data scaling. After normalization, the fold change ( $\log_2$ -transformed ratio) of lipids between each group was calculated.

### **Supplementary Table 3**

PE molecular composition in Selenoi-deficient Tfh cells and in Pcyt2-deficient Tfh cells upon ether-type or diacyl-type PE supplementation. This file contains two tabs (a, b). Tab a shows the 79 PE molecules that were significantly downregulated ( $P < 0.05$ ) upon Selenoi deletion, with 47 and 32 PE molecules being ether-type and diacyl-type PE, respectively. Tab b shows that Pcyt2-deficient SMARTA cells supplemented with ether-type (left) or diacyl-type (right) PE resulted in the generation of both species of PE molecules. PE molecules that were restored by either ether-type or diacyl-type PE supplementation are underlined.

### **Supplementary Table 4**

Expression levels of 48 lipid translocases in Tfh and Th1 cells, and CRISPR-based *in vivo* screening of the 48 lipid translocases. This file contains two tabs (a, b). Tab a shows the average gene expression levels in Tfh and Th1 cells ( $\log_2$  values, from a public dataset GSE74854) of the 48 translocase genes that we examined by CRISPR-based *in vivo* functional testing (see Tab b). Tab b shows the data summary of CRISPR-based *in vivo* functional testing of 48 candidate lipid translocases. SMARTA cells transduced with sgNTC (mCherry $^+$ ) and the indicated gene-specific sgRNA (Ametrine $^+$ ) were mixed at a 1:2 ratio and transferred into C57BL/6 hosts followed by LCMV infection. Positive regulators of cell proliferation (middle columns) or Tfh differentiation (right columns) were defined by quantification of the relative proportions of donor-derived cells (normalized to ‘spike’ cells) or Tfh cells (CXCR5 $^+$ SLAM $^-$ , CXCR5 $^+$ PD-1 $^+$ , PSGL-1 $^-$ Ly6C $^-$ ), respectively (as comparisons, 49.4% reduction and 53.3% increase in Tfh differentiation were observed for deletion of Bcl6 and Blimp1, respectively; data not shown).

## **Source data**

[\*\*Source Data Fig. 1\*\*](#)

[\*\*Source Data Fig. 2\*\*](#)

[\*\*Source Data Fig. 3\*\*](#)

[\*\*Source Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 1\*\*](#)

[\*\*Source Data Extended Data Fig. 2\*\*](#)

[\*\*Source Data Extended Data Fig. 3\*\*](#)

[\*\*Source Data Extended Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 5\*\*](#)

[\*\*Source Data Extended Data Fig. 6\*\*](#)

[\*\*Source Data Extended Data Fig. 7\*\*](#)

[\*\*Source Data Extended Data Fig. 8\*\*](#)

[\*\*Source Data Extended Data Fig. 9\*\*](#)

[\*\*Source Data Extended Data Fig. 10\*\*](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Fu, G., Guy, C.S., Chapman, N.M. *et al.* Metabolic control of T<sub>FH</sub> cells and humoral immunity by phosphatidylethanolamine. *Nature* **595**, 724–729 (2021). <https://doi.org/10.1038/s41586-021-03692-z>

[Download citation](#)

- Received: 17 March 2020
- Accepted: 03 June 2021
- Published: 07 July 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03692-z>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Advertisement

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

- Article
- [Published: 21 July 2021](#)

# EGFR activation limits the response of liver cancer to lenvatinib

- [Haojie Jin<sup>1,2</sup> na1](#),
- [Yaoping Shi<sup>3</sup> na1](#),
- [Yuanyuan Lv<sup>1</sup> na1](#),
- [Shengxian Yuan<sup>4</sup> na1](#),
- [Christel F. A. Ramirez<sup>5</sup>](#),
- [Cor Lieftink<sup>2</sup>](#),
- [Liqin Wang<sup>2</sup>](#),
- [Siying Wang<sup>1</sup>](#),
- [Cun Wang](#) ORCID: orcid.org/0000-0002-4977-2189<sup>1</sup>,
- [Matheus Henrique Dias<sup>2</sup>](#),
- [Fleur Jochems](#) ORCID: orcid.org/0000-0001-7305-7321<sup>2</sup>,
- [Yuan Yang<sup>4</sup>](#),
- [Astrid Bosma<sup>2</sup>](#),
- [E. Marielle Hijmans<sup>2</sup>](#),
- [Marnix H. P. de Groot](#) ORCID: orcid.org/0000-0003-1013-868X<sup>5</sup>,
- [Serena Vegna<sup>5</sup>](#),
- [Dan Cui<sup>3</sup>](#),
- [Yangyang Zhou](#) ORCID: orcid.org/0000-0002-2964-2803<sup>1</sup>,
- [Jing Ling<sup>1</sup>](#),
- [Hui Wang<sup>1</sup>](#),
- [Yuchen Guo<sup>1</sup>](#),
- [Xingling Zheng<sup>1</sup>](#),
- [Nikita Isima<sup>2</sup>](#),

- [Haiqiu Wu<sup>6</sup>](#),
- [Chong Sun ORCID: orcid.org/0000-0001-9008-2158<sup>7</sup>](#) nAff8,
- [Roderick L. Beijersbergen ORCID: orcid.org/0000-0003-0116-4130<sup>2</sup>](#),
- [Leila Akkari ORCID: orcid.org/0000-0001-5479-641X<sup>5</sup>](#),
- [Weiping Zhou ORCID: orcid.org/0000-0002-8464-3454<sup>4</sup>](#),
- [Bo Zhai ORCID: orcid.org/0000-0003-4337-7126<sup>1,3</sup>](#),
- [Wenxin Qin ORCID: orcid.org/0000-0002-1989-3027<sup>1</sup>](#) &
- [René Bernards ORCID: orcid.org/0000-0001-8677-3423<sup>2</sup>](#)

*Nature* volume **595**, pages 730–734 (2021) [Cite this article](#)

- 10k Accesses
- 110 Altmetric
- [Metrics details](#)

## Subjects

- [Cancer genetics](#)
- [Cancer therapy](#)

## Abstract

Hepatocellular carcinoma (HCC)—the most common form of liver cancer—is an aggressive malignancy with few effective treatment options<sup>1</sup>. Lenvatinib is a small-molecule inhibitor of multiple receptor tyrosine kinases that is used for the treatment of patients with advanced HCC, but this drug has only limited clinical benefit<sup>2</sup>. Here, using a kinome-centred CRISPR–Cas9 genetic screen, we show that inhibition of epidermal growth factor receptor (EGFR) is synthetic lethal with lenvatinib in liver cancer. The combination of the EGFR inhibitor gefitinib and lenvatinib displays potent anti-proliferative effects *in vitro* in liver cancer cell lines that express EGFR and *in vivo* in xenografted liver cancer cell lines, immunocompetent mouse models and patient-derived HCC tumours in mice. Mechanistically, inhibition of fibroblast growth factor receptor (FGFR) by lenvatinib

treatment leads to feedback activation of the EGFR–PAK2–ERK5 signalling axis, which is blocked by EGFR inhibition. Treatment of 12 patients with advanced HCC who were unresponsive to lenvatinib treatment with the combination of lenvatinib plus gefitinib (trial identifier NCT04642547) resulted in meaningful clinical responses. The combination therapy identified here may represent a promising strategy for the approximately 50% of patients with advanced HCC who have high levels of EGFR.

## 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](#)

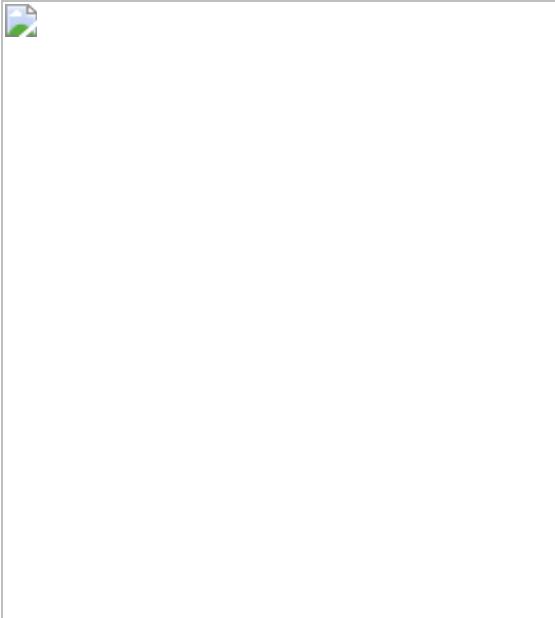
All prices are NET prices.

### Additional access options:

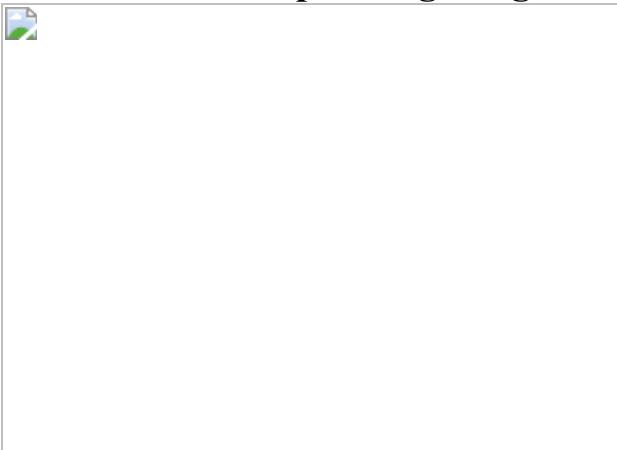
- [Log in](#)

- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

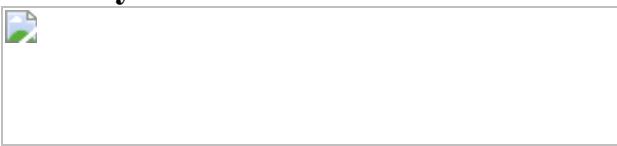
**Fig. 1: A CRISPR–Cas9 screen identifies EGFR as a synthetic lethal target of lenvatinib in liver cancer.**



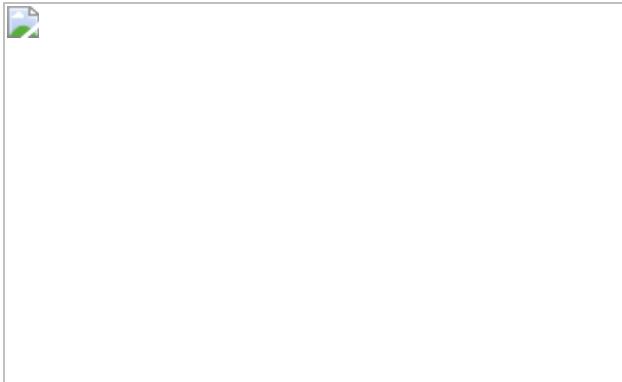
**Fig. 2: Synthetic lethal effects of lenvatinib and EGFR inhibitors on liver cancer cells expressing a high level of EGFR in vitro and in vivo.**



**Fig. 3: Feedback activation of the EGFR–PAK2–ERK5 cascade limits sensitivity of liver cancer cells to lenvatinib.**



**Fig. 4: Responses of patients with HCC to lenvatinib plus gefitinib therapy.**



## Data availability

All data that support the findings of this study are available from the corresponding authors upon reasonable request. The RNA-sequencing data have been deposited in the Gene Expression Omnibus ([GSE157905](#)). [Source data](#) are provided with this paper.

## References

1. 1.

Villanueva, A. Hepatocellular carcinoma. *N. Engl. J. Med.* **380**, 1450–1462 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

2. 2.

Kudo, M. et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet* **391**, 1163–1173 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Sung, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.

*CA Cancer J. Clin.* **71**, 209–249 (2021).

[Article](#) [Google Scholar](#)

4. 4.

Zucman-Rossi, J., Villanueva, A., Nault, J.-C. & Llovet, J. M. Genetic landscape and biomarkers of hepatocellular carcinoma. *Gastroenterology* **149**, 1226–1239 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Llovet, J. M. et al. Sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* **359**, 378–390 (2008).

[CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Finn, R. S. et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N. Engl. J. Med.* **382**, 1894–1905 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Evers, B. et al. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat. Biotechnol.* **34**, 631–633 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Zhu, A. X. et al. SEARCH: a phase III, randomized, double-blind, placebo-controlled trial of sorafenib plus erlotinib in patients with

advanced hepatocellular carcinoma. *J. Clin. Oncol.* **33**, 559–566 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Matsuki, M. et al. Lenvatinib inhibits angiogenesis and tumor fibroblast growth factor signaling pathways in human hepatocellular carcinoma models. *Cancer Med.* **7**, 2641–2653 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Wagle, M.-C. et al. A transcriptional MAPK Pathway Activity Score (MPAS) is a clinically relevant biomarker in multiple cancer types. *NPJ Precis. Oncol.* **2**, 7 (2018).

[Article](#) [Google Scholar](#)

11. 11.

Pirker, R. et al. EGFR expression as a predictor of survival for first-line chemotherapy plus cetuximab in patients with advanced non-small-cell lung cancer: analysis of data from the phase 3 FLEX study. *Lancet Oncol.* **13**, 33–42 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Kimura, T. et al. Immunomodulatory activity of lenvatinib contributes to antitumor activity in the Hepa1-6 hepatocellular carcinoma model. *Cancer Sci.* **109**, 3993–4002 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Li, E., Lin, L., Chen, C. W. & Ou, D. L. Mouse models for immunotherapy in hepatocellular carcinoma. *Cancers (Basel)* **11**, 1800 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

14. 14.

Keng, V. W. et al. A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nat. Biotechnol.* **27**, 264–274 (2009).

[CAS](#) [Article](#) [Google Scholar](#)

15. 15.

He, H. et al. Platelet-derived growth factor requires epidermal growth factor receptor to activate p21-activated kinase family kinases. *J. Biol. Chem.* **276**, 26741–26744 (2001).

[CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Lee, S.-W. et al. EGFR–Pak signaling selectively regulates glutamine deprivation-induced macropinocytosis. *Dev. Cell* **50**, 381–392 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Vaseva, A. V. et al. KRAS suppression-induced degradation of MYC is antagonized by a MEK5–ERK5 compensatory mechanism. *Cancer Cell* **34**, 807–822 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Duncan, J. S. et al. Dynamic reprogramming of the kinaseome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell* **149**, 307–321 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Lito, P. et al. Relief of profound feedback inhibition of mitogenic signaling by RAF inhibitors attenuates their activity in BRAFV600E melanomas. *Cancer Cell* **22**, 668–682 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Kosoff, R. E. et al. Pak2 restrains endomitosis during megakaryopoiesis and alters cytoskeleton organization. *Blood* **125**, 2995–3005 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Kobayashi, S. et al. Transcriptional profiling identifies cyclin D1 as a critical downstream effector of mutant epidermal growth factor receptor signaling. *Cancer Res.* **66**, 11389–11398 (2006).

[CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Philip, P. A. et al. Phase II study of erlotinib (OSI-774) in patients with advanced hepatocellular cancer. *J. Clin. Oncol.* **23**, 6657–6663 (2005).

[CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Chan, S. L. et al. New utility of an old marker: serial  $\alpha$ -fetoprotein measurement in predicting radiologic response and survival of patients with hepatocellular carcinoma undergoing systemic chemotherapy. *J. Clin. Oncol.* **27**, 446–452 (2009).

[CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Llovet, J. M. & Lencioni, R. mRECIST for HCC: performance and novel refinements. *J. Hepatol.* **72**, 288–306 (2020).

[Article](#) [Google Scholar](#)

25. 25.

Ding, X. et al. Genomic and epigenomic features of primary and recurrent hepatocellular carcinomas. *Gastroenterology* **157**, 1630–1645 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Cremolini, C. et al. Early tumor shrinkage and depth of response predict long-term outcome in metastatic colorectal cancer patients treated with first-line chemotherapy plus bevacizumab: results from phase III TRIBE trial by the Gruppo Oncologico del Nord Ovest. *Ann. Oncol.* **26**, 1188–1194 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

[Article](#) [Google Scholar](#)

28. 28.

Jin, H. et al. Regulator of calcineurin 1 gene isoform 4, down-regulated in hepatocellular carcinoma, prevents proliferation, migration, and invasive activity of cancer cells and metastasis of orthotopic tumors by inhibiting nuclear translocation of NFAT1. *Gastroenterology* **153**, 799–811 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Mazières, J. et al. Evaluation of EGFR protein expression by immunohistochemistry using H-score and the magnification rule: re-analysis of the SATURN study. *Lung Cancer* **82**, 231–237 (2013).

[Article](#) [Google Scholar](#)

30. 30.

Sun, C. et al. Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. *Nature* **508**, 118–122 (2014).

[CAS](#) [Article](#) [ADS](#) [Google Scholar](#)

31. 31.

Wang, L. et al. High-throughput functional genetic and compound screens identify targets for senescence induction in cancer. *Cell Rep.* **21**, 773–783 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Schuhmacher, M. et al. The transcriptional program of a human B cell line in response to Myc. *Nucleic Acids Res.* **29**, 397–406 (2001).

[CAS](#) [Article](#) [Google Scholar](#)

33. 33.

Wang, C. et al. Inducing and exploiting vulnerabilities for the treatment of liver cancer. *Nature* **574**, 268–272 (2019).

[CAS Article](#) [ADS](#) [Google Scholar](#)

34. 34.

World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bull. World Health Organ.* **79**, 373–374 (2001).

[PubMed Central](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We thank R. G. H. Beets-Tan for discussion and assistance in MRI image analysis. This work was funded by grants from the European Research Council (ERC 787925 to R.B.), the Dutch Cancer Society through the Oncode Institute, the National Natural Science Foundation of China (82073039 to H.J., 82011530441 to H.J., 81702838 to H.J., 81920108025 to W.Q. and 81801800 to Y.S.), National Science and Technology Key Project of China (2018ZX10732202-002-003 to W.Q. and 2018ZX10302205-002-003 to H.J.), Shanghai Rising-Star Program (19QA1408200 to H.J.), Shanghai Municipal Science and Technology Project (20JC1411100 to W.Q.), 111 project (B21024 to W.Q.), Shanghai Municipal Commission of Health and Family Planning (2018YQ20 to H.J.) and Shanghai Jiao Tong University School of Medicine (YG2019GD01 to H.J. and PY11-17-013 to H.J.). We thank the facilities of The Netherlands Cancer Institute.

## Author information

Author notes

1. Chong Sun

Present address: Immune Regulation in Cancer Group, German Cancer Research Center, Heidelberg, Germany

2. These authors contributed equally: Haojie Jin, Yaoping Shi, Yuanyuan Lv, Shengxian Yuan

## Affiliations

1. State Key Laboratory of Oncogenes and Related Genes, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Haojie Jin, Yuanyuan Lv, Siying Wang, Cun Wang, Yangyang Zhou, Jing Ling, Hui Wang, Yuchen Guo, Xingling Zheng, Bo Zhai & Wenxin Qin

2. Division of Molecular Carcinogenesis, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Haojie Jin, Cor Loeffink, Liqin Wang, Matheus Henrique Dias, Fleur Jochems, Astrid Bosma, E. Marielle Hijmans, Nikita Isima, Roderick L. Beijersbergen & René Bernards

3. Department of Interventional Oncology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Yaoping Shi, Dan Cui & Bo Zhai

4. The Third Department of Hepatic Surgery, Eastern Hepatobiliary Surgery Hospital, Shanghai, China

Shengxian Yuan, Yuan Yang & Weiping Zhou

5. Division of Tumour Biology and Immunology, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Christel F. A. Ramirez, Marnix H. P. de Groot, Serena Vegna & Leila Akkari

6. Department of Cell and Chemical Biology, Leiden University Medical Centre, Leiden, The Netherlands

Haiqiu Wu

7. Division of Immunology, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Chong Sun

## Authors

1. Haojie Jin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Yaoping Shi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Yuanyuan Lv

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Shengxian Yuan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Christel F. A. Ramirez

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Cor Lieftink

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Liqin Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Siying Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Cun Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Matheus Henrique Dias

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Fleur Jochems

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Yuan Yang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Astrid Bosma

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. E. Marielle Hijmans

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Marnix H. P. de Groot

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Serena Vegna

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Dan Cui

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Yangyang Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Jing Ling

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Hui Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Yuchen Guo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Xingling Zheng  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Nikita Isima  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Haiqiu Wu  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Chong Sun  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Roderick L. Beijersbergen  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Leila Akkari  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Weiping Zhou  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Bo Zhai  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Wenxin Qin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

31. René Bernards

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

R.B., W.Q., B.Z. and W.Z. supervised all of the research. R.B., W.Q. and H.J. wrote the manuscript. H.J., Y.L., S.Y., C.F.A.R., S.W., F.J., A.B., Y.Z., J.L., H. Wang, Y.G., X.Z., M.H.P.d.G., S.V. and N.I. designed and performed the preclinical experiments. H.J., C.L. and R.L.B. performed data analysis. B.Z., Y.S. and D.C. designed and performed the clinical study. L.W., C.W., M.H.D., Y.Y., E.M.H., H. Wu, C.S., R.L.B. and L.A. provided advice and technical support for the project. All authors commented on the manuscript.

## Corresponding authors

Correspondence to [Weiping Zhou](#) or [Bo Zhai](#) or [Wenxin Qin](#) or [René Bernards](#).

## Ethics declarations

## Competing interests

H.J. and R.B. are listed as inventors on a patent application using the drug combination described here. The other authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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

## Extended data figures and tables

### Extended Data Fig. 1 A kinome CRISPR screen in the lenvatinib-resistant liver cancer cell line SNU449.

**a**, Liver cancer cells are resistant to lenvatinib treatment in vitro. Short-term viability assay of liver cancer cell lines. Cells were treated with increasing concentrations of lenvatinib for 72 h, and cell viability was determined using CellTiter-Blue. Data are mean  $\pm$  s.e.m. Cell lines with half-maximum inhibitory concentration ( $IC_{50}$ ) values  $< 5 \mu\text{M}$  are classed as relatively sensitive cells, whereas cell lines with  $IC_{50} > 5 \mu\text{M}$  are classed as relatively resistant cells. SNU449, JHH1, SNU182, PLC/PRF5, MHCC97H, HepG2, SNU398, Huh7 and Hep3B,  $n = 8$ ; Huh6 and SK-Hep1,  $n = 7$ . **b**, Scatter plots of  $\log_{10}$ -transformed normalized read counts of independent biological replicates of GeCKO library lentiviral transductions with  $R^2$  values shown ( $n = 3$ ). Replicates from triplicate transductions showed good correlation at the sgRNA level.  $T_0$ , SNU449 cells after puromycin selection;  $T_{\text{untreated}}$ , SNU449 cells cultured for 14 days without drug treatment;  $T_{\text{treated}}$ , SNU449 cells cultured for 14 days with 10  $\mu\text{M}$  lenvatinib. **c**, Fold change distribution of all gRNAs targeting essential genes or nonessential genes in three independent biological replicates. The sgRNAs for essential genes were strongly depleted relative to the sgRNAs targeting nonessential genes, highlighting the quality of the screen. ‘p’, 50 sgRNAs targeting 10 essential genes (red); ‘n’, 50 non-targeting control sgRNAs (blue); ‘x’, 5,860 sgRNAs targeting 504 human kinases (grey). **d**, Average  $\log_2$ -transformed fold change in each individual EGFR gRNA in the pooled CRISPR library ( $n = 3$  independent experiments). Data are mean  $\pm$  s.e.m.  $P$  values were determined by two-sided unpaired Student’s *t*-test

[Source data](#).

**Extended Data Fig. 2 Synthetic lethal effects of lenvatinib and EGFR inhibitors on liver cancer cells expressing a high level of EGFR in vitro.**

**a–f**, EGFR<sup>high</sup> liver cancer cell lines (SNU449, JHH1, Huh6 and SNU182) or EGFR<sup>low</sup> cells (SNU398 and HepG2) were treated with lenvatinib, EGFR inhibitors (gefitinib or erlotinib) or their combination at the indicated concentrations. Growth curves were determined by Incucyte live cell analyses. Data are mean ± s.e.m.,  $n = 5$ . **g–l**, The quantification of three independent assays in Fig. [2b–g](#). Data are mean ± s.e.m.,  $n = 3$  independent experiments. **m–o**, Ectopic EGFR expression confers resistance to lenvatinib in EGFR<sup>low</sup> cells. **m**, Western blot analysis of p-EGFR and total EGFR levels in SNU398 and HepG2 cells infected with control or EGFR overexpression vector. HSP90 served as a control. **n, o**, SNU398 and HepG2 cells expressing control vector or EGFR were cultured with or without lenvatinib (SNU398, 5 µM; HepG2, 10 µM). **n**, The cells were fixed, stained and photographed after 14 days. **o**, Quantification of three independent assays. Data are mean ± s.e.m.,  $n = 3$  independent experiments

[Source data](#).

**Extended Data Fig. 3 The combination of sorafenib and the EGFR inhibitor gefitinib shows no synergy in liver cancer cells in vitro.**

**a–f**, Liver cancer cell lines were treated with sorafenib, the EGFR inhibitor gefitinib or their combination at the indicated concentrations. The cells were fixed and stained after 10–14 days. Representative data from three independent experiments. **g–l**, The quantification of the three independent assays shown in **a–f**. Data are mean ± s.e.m.,  $n = 3$  independent experiments

[Source data](#).

## Extended Data Fig. 4 Bliss independence model evaluating the synergistic effect between lenvatinib or sorafenib with gefitinib in liver cancer cells.

The Bliss independence model was applied to the quantification data of colony-formation experiments in Extended Data Figs. 2g–l, 3g–l. The data in the untreated condition (no lenvatinib or gefitinib) were normalized as 0, indicating no inhibition of cell viability. Then, the additive scores (blissAdd) were generated by multiplying the normalized effect of the two single drugs. We further calculated the reduction in viability by subtracting the mean of the blissAdd from the mean of the measured combination scores (measured). The green background indicates the synergistic effect with  $P \leq 0.05$  and when the reduction in viability is more than 10%. **a–f**, Bliss independence analysis for the combination of gefitinib and lenvatinib shown in Extended Data Fig. 2g–l. **g–l**, Bliss independence analysis for the combination of gefitinib and sorafenib shown in Extended Data Fig. 3g–l. Data are mean  $\pm$  s.e.m.,  $n = 3$  independent experiments.  $P$  values were determined by unpaired two-sided Student's *t*-test

[Source data](#).

## Extended Data Fig. 5 EGFR inhibitors increased the sensitivity of EGFR<sup>high</sup> liver cancer cells to FGFR inhibitors in vitro.

**a, b**, Long-term colony formation assay of 11 liver cancer cell lines treated with the FGFR inhibitors AZD4547 (**a**) and BGJ398 (**b**). Cells were grown in the absence or presence of drugs at the indicated concentrations for 10–14 days, fixed and stained. **c–f**, EGFR<sup>high</sup> liver cancer cell lines (SNU449, JHH1, Huh6 and SNU182) were treated with FGFR inhibitors (AZD4547 or BGJ398), EGFR inhibitor (gefitinib) or their combination at the indicated concentrations. Growth curves were determined by Incucyte live cell analyses. Data are mean  $\pm$  s.e.m.,  $n = 5$ . **g–l**, FGFR inhibitors (AZD4547 or BGJ398) but not the FGFR4-specific inhibitor BLU-554 showed a synergistic effect with EGFR inhibitors in EGFR<sup>high</sup> liver cancer cells in vitro. Liver cancer cell lines with different EGFR expression levels were treated with FGFR inhibitors (AZD4547 or BGJ398), FGFR4-specific

inhibitor (BLU-554), EGFR inhibitors (gefitinib or erlotinib) or their different combinations at the indicated concentrations. The cells were fixed and stained after 10–14 days. **m–o**, SNU449 and JHH1 cells were treated with gefitinib (2.5 µM for SNU449; 0.625 µM for JHH1), BGJ398, sorafenib, or as a combination of two or three drugs. **m**, The cells were fixed, stained and photographed after 14 days. **n, o**, Quantification of three independent assays. Data are mean ± s.e.m.,  $n = 3$  independent experiments.  $P$  values were determined by two-sided unpaired Student's *t*-test

[Source data](#).

**Extended Data Fig. 6 Lenvatinib plus EGFR inhibitors suppresses the ERK1/2 MAPK pathway in EGFR<sup>high</sup> liver cancer cells.**

**a–f**, Immunoblot and quantification of p-EGFR and p-ERK1/2 levels in liver cancer cells after the indicated treatment. EGFR<sup>high</sup> and EGFR<sup>low</sup> liver cancer cell lines were treated with lenvatinib, EGFR inhibitors (gefitinib or erlotinib) or their combination at the indicated concentrations for 24 h. Protein extracts were probed with specific antibodies against EGFR (total and phosphorylated), ERK1/2 (total and phosphorylated) and HSP90 (as a loading control). Data are mean ± s.e.m.,  $n = 3$  independent experiments.  $P$  values were determined by two-sided unpaired Student's *t*-test. **g–j**, Liver cancer cell lines SNU449 (**g**), JHH1 (**h**), Huh6 (**i**) and SNU182 (**j**) were treated with DMSO, lenvatinib, gefitinib or the combination of both drugs for 24 h. The mRNA levels of 10 transcriptional targets of MAPK signalling (*DUSP4*, *DUSP6*, *ETV4*, *ETV5*, *EPHA2*, *EPHA4*, *SPRY2*, *SPRY4*, *PHLDA1* and *CCND1*) were determined by RT–qPCR analysis. *ACTB* (which encodes β-actin) was used as internal control. Data are mean ± s.e.m. ( $n = 3$ )

[Source data](#).

**Extended Data Fig. 7 The combination of lenvatinib and gefitinib suppresses tumour growth without toxicity in cell-line-derived immunodeficient mouse models in vivo.**

**a–j**, Liver cancer cells were grown as tumour xenografts in BALB/c nude mice. After tumour establishment ( $\sim 200 \text{ mm}^3$ ), mice were treated with vehicle, lenvatinib (4 mg kg $^{-1}$ ), gefitinib (80 mg kg $^{-1}$ ) or lenvatinib (4 mg kg $^{-1}$ ) plus gefitinib (80 mg kg $^{-1}$ ), for the indicated number of days. Representative images of H&E, PCNA, Ki-67, cleaved caspase-3 and CD31 staining in SNU449 (**a**) and Huh6 (**b**) xenograft models. Scale bars, 50  $\mu\text{m}$ . **c–j**, Quantification of PCNA $^+$  cells (**c**, **g**), Ki-67 $^+$  cells (**d**, **h**), cleaved caspase-3 $^+$  cells (**e**, **i**) and microvessel density (MVD) (**f**, **j**) per high-power field in representative sections from each group.  $n = 6$  per group. Data are mean  $\pm$  s.e.m.  $P$  values were determined by two-sided unpaired Student's *t*-test. **k**, **l**, The body weight of mice in the SNU449 (**k**) and Huh6 (**l**) xenograft models. Data are mean  $\pm$  s.e.m.  $n = 6$  mice per group

[Source data](#).

**Extended Data Fig. 8 The combination of lenvatinib and gefitinib suppresses tumour growth and extends survival time in HCC PDX mouse models in vivo.**

**a**, **b**, Four HCC PDX models were established and examined for EGFR expression by IHC analyses. **a**, Representative images of H&E and EGFR staining in the four PDX models. Scale bars, 50  $\mu\text{m}$ . **b**, *H*-score of EGFR expression levels in the four PDX models. Data are mean  $\pm$  s.e.m.,  $n = 5$  samples. **c–n**, After tumour establishment ( $\sim 200 \text{ mm}^3$ ), mice were treated with vehicle, lenvatinib (4 mg kg $^{-1}$ ), gefitinib (80 mg kg $^{-1}$ ) or lenvatinib (4 mg kg $^{-1}$ ) plus gefitinib (80 mg kg $^{-1}$ ). **c–f**, Growth curves of four HCC PDX models in mice treated with vehicle, lenvatinib, gefitinib or their combination.  $n = 5$  mice per group. Data are mean  $\pm$  s.e.m.  $P$  values were determined by two-way ANOVA with Tukey multiple comparisons. **g–j**, Kaplan–Meier survival curves of the PDX models after the indicated treatments.  $n = 5$  mice per group.  $P$  values were determined by log-rank Mantel–Cox test. **k–n**, Body weights of mice in the PDX models with the aforementioned treatments were assessed.  $n = 5$  mice per group. Data are mean  $\pm$  s.e.m

[Source data](#).

## Extended Data Fig. 9 The combination of lenvatinib and gefitinib suppresses tumour growth without toxicity in immunocompetent mouse models in vivo.

**a**, Western blot analysis of total EGFR levels in mouse liver tissues and mouse liver cancer cells (Hepa1-6). Representative of three independent experiments. **b**, Hepa1-6 cells were treated with lenvatinib, the EGFR inhibitor gefitinib or their combination at the indicated concentrations. The cells were fixed and stained after 7 days. Representative of three independent experiments. **c**, After intrahepatic inoculation of Hepa1-6 cells in C57BL/6 mice, mice were treated with vehicle, lenvatinib ( $4 \text{ mg kg}^{-1}$ ), gefitinib ( $80 \text{ mg kg}^{-1}$ ) or lenvatinib ( $4 \text{ mg kg}^{-1}$ ) plus gefitinib ( $80 \text{ mg kg}^{-1}$ ) for 2 weeks. Tumour weight was measured. Data are mean  $\pm$  s.e.m.  $n = 6$  mice per group.  $P$  values were determined by two-sided unpaired Student's *t*-test. **d**, Body weights of mice in the Hepa1-6 orthotopic models with the aforementioned treatments were assessed.  $n = 6$  mice per group. Data are mean  $\pm$  s.e.m. **e**, Schematic of gene delivery by hydrodynamic tail vein injection (HDTVi) of the Myc proto-oncogene transposon system and a CRISPR–Cas9 vector targeting the *Trp53* tumour suppressor, which was used to induce HCC 2–3 weeks after hydrodynamic tail vein injections. **f**, Western blot analysis of total EGFR levels in mouse liver tissues and mouse liver cancer cells (*Myc*<sup>OE</sup>*Trp53*<sup>KO</sup>). Representative of three independent experiments. **g**, *Myc*<sup>OE</sup>*Trp53*<sup>KO</sup> mouse liver cancer cells were treated with lenvatinib, EGFR inhibitor gefitinib or their combination at the indicated concentrations, respectively. The cells were fixed and stained after 5 days. Representative of three independent experiments. **h**, Survival curve generated from mice bearing *Myc*<sup>OE</sup>*Trp53*<sup>KO</sup> tumours, treated with vehicle ( $n = 6$ ; median survival of 11 days), lenvatinib ( $4 \text{ mg kg}^{-1}$ ;  $n = 9$ ; median survival of 23 days), gefitinib ( $80 \text{ mg kg}^{-1}$ ;  $n = 6$ ; median survival of 13.5 days) or lenvatinib plus gefitinib ( $n = 9$ ; median survival of 31 days).  $P$  values were determined by two-sided log-rank test. **i**, Body weights of mice in the *Myc*<sup>OE</sup>*Trp53*<sup>KO</sup> mouse liver cancer models with the aforementioned treatments were assessed.  $n = 6\text{--}9$  mice per group. Data are mean  $\pm$  s.e.m. **j, k**, Mice bearing *Myc*<sup>OE</sup>*Trp53*<sup>KO</sup> tumours treated with vehicle, lenvatinib ( $4 \text{ mg kg}^{-1}$ ), gefitinib ( $80 \text{ mg kg}^{-1}$ ) or a combination of both drugs were killed at the end point after treatment. Tumours were

dissociated as single-cell suspensions, and flow cytometry analyses were performed to determine the content of tumour-associated lymphoid cells (NK cells, CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells and regulatory T cells (Tregs)) and myeloid cells (monocytes, neutrophils, tumour-associated macrophages (TAMs) and dendritic cells (DCs)). Data are mean ± s.e.m. *P* values were determined by two-sided unpaired Student's *t*-test. Sample sizes are given in the Methods

[Source data.](#)

**Extended Data Fig. 10 Concurrent inhibition of ERK5 and ERK1/2 shows an enhanced anti-tumour effect in vitro.**

**a**, Quantification of Fig. 3b with three independent experiments. The phosphorylation levels of each protein were normalized based on their total protein levels. *n* = 3 independent experiments. Data are mean ± s.e.m. *P* values were determined by two-sided unpaired Student's *t*-test. **b**, Western blot analysis of EGFR–PAK2–ERK5 cascade in JHH1 cells, treated with lenvatinib, EGFR inhibitors (gefitinib and erlotinib) or their combination at the indicated concentrations for 6 h. HSP90 served as a loading control. **c**, The phosphorylation levels of each protein in **b** were normalized based on their total protein levels. *n* = 3 independent experiments. Data are mean ± s.e.m. *P* values were determined by two-sided unpaired Student's *t*-test. **d**, EGFR expression of SNU449 cells was knocked down by two independent shRNAs, and cells were further treated with lenvatinib (5 μM) for 6 h. Western blot analysis of EGFR–PAK2–ERK5 cascade was performed. HSP90 served as a loading control. The pLKO vector was used in the control experiment. **e**, Quantification of Fig. 3c with three independent experiments. The phosphorylation levels of each protein were normalized based on their total protein levels. *n* = 3 independent experiments. Data are mean ± s.e.m. *P* values were determined by two-sided unpaired Student's *t*-test. **f**, JHH1 cells were treated with lenvatinib (5 μM), PAK inhibitor FRAX1036 (2.5 μM) or their combination for 6 h, and western blot analysis was performed with the indicated antibodies. **g**, The phosphorylation levels of each protein in **f** were normalized based on their total protein levels. *n* = 3 independent experiments. Data are mean ± s.e.m. *P* values were determined by two-sided unpaired Student's *t*-

test. **h, i**, Long-term colony formation assays showing synergistic effects of lenvatinib and the PAK inhibitor FRAX1036 on the proliferation of SNU449 (**h**) and JHH1 (**i**) cells. Representative of three independent experiments. **j, k**, Long-term colony formation assays showing synergistic effects of lenvatinib and ERK5 inhibitor XMD8-92 on the proliferation of SNU449 (**j**) and JHH1 (**k**) cells. Representative of three independent experiments. **l, m**, SNU449 (**l**) and JHH1 (**m**) cells were treated with lenvatinib (5  $\mu$ M), the ERK5 inhibitor XMD8-92 (2.5  $\mu$ M) or their combination for 6 h, and western blot analysis was performed with the indicated antibodies. HSP90 served as a loading control. **n–q**, Knockout of ERK5 using the CRISPR–Cas9 system enhances the response to lenvatinib. The ERK5 knockout efficiency was determined by western blot in SNU449 (**n**) and JHH1 (**p**) cells. HSP90 served as a loading control. The effects of ERK5 knockout on proliferation were indicated by colony formation. The ERK5 knockout SNU449 (**o**) and JHH1 (**q**) cells, or their respective control cells were treated with DMSO or 5  $\mu$ M lenvatinib. After 10 days of culture, cells were fixed, stained and photographed. NT, non-targeting sgRNA. **r**, Quantification of IHC staining in Fig. [3d](#).  $n = 6$  per group. Data are mean  $\pm$  s.e.m.  $P$  values were determined by two-sided unpaired Student's *t*-test

[Source data](#).

## Extended Data Fig. 11 RNA-sequencing analysis of the combination treatment of lenvatinib and gefitinib in EGFR<sup>high</sup> liver cancer cells.

**a–d**, Heat-map representations of the log<sub>2</sub>-transformed gene expression changes (log2FC) in liver cancer cells SNU449 (**a**), JHH1 (**b**), Huh6 (**c**) and SNU182 (**d**) treated with DMSO, 2.5  $\mu$ M gefitinib, 5  $\mu$ M lenvatinib, or the two drugs for 24 h. Differential expression of the whole genome is shown on the basis of each treatment versus DMSO using RNA-sequencing analysis. To show robust changes, only genes that have a mean read count over all samples equal or greater than 100 were included. To prevent too much dilution of the colour-coding, the log<sub>2</sub>-transformed fold-change values were truncated in that lower than -2 were set to -2 and higher than 2

were set to 2. Heat maps depicting all three treatments were then generated by unsupervised hierarchical clustering. One replicate was used for each sample. **e**, For the genes with a mean read count over all samples greater than 100 per line, the top 25% highest and 25% lowest synergy scores were determined and shown as heat maps. These lists were merged and an unsupervised hierarchical clustering was performed. **f**, GSEA analysis of the gene set of Kobayashi ‘EGFR signalling down’ in each comparison. Unsupervised hierarchical clustering of normalized enrichment scores (NES) was used to generate a comprehensive heat-map visualization of the functional transcriptional outputs of the four cell lines. \* $P < 0.001$ . **g**, GSEA analysis of curated gene sets was performed, Kobayashi ‘EGFR signalling down’ and Schuhmacher ‘MYC targets up’ were identified as two of the highest-ranking downregulated gene sets in the combination-treated cells in all four liver cancer cell lines based on additional gene alterations in combination group. **h–k**, A fold change preranked list of each treatment versus DMSO was used to run GSEA against the Hallmark gene sets in SNU449 (**h**), JHH1 (**i**), Huh6 (**j**) and SNU182 (**k**) cells. Unsupervised hierarchical clustering of normalized enrichment scores was used to generate a comprehensive heat-map visualization of the functional transcriptional outputs of each treatment (FDR < 0.1). **l–n**, GSEA indicates that the gene sets of Hallmark ‘MYC targets v1’ (**l**), Hallmark ‘MYC targets v2’ (**m**) and Hallmark ‘KRAS signalling up’ (**n**) were negatively enriched in the combination group based on additional gene alterations.

## Extended Data Fig. 12 Expression and prognostic effect of EGFR in patients with HCC.

**a–d**, IHC staining analyses of EGFR were performed in 298 patients with HCC. The *H*-score method assigned a score of 0–300 to each patient sample, based on the percentage of cells stained at different intensities. The discriminatory threshold was set at 200. All samples were classified as low ( $H < 200$ ; EGFR<sup>low</sup>) or high ( $H \geq 200$ ; EGFR<sup>high</sup>) EGFR expression. **a**, Typical images of EGFR<sup>low</sup> and EGFR<sup>high</sup> cases according to their *H*-score. Scale bars, 200  $\mu\text{m}$ . **b**, Distribution of EGFR expression levels in 298 patients with HCC as represented by various *H*-score ranges. Out of 298 patients with HCC, 157 had *H*-score of  $H \geq 200$  (EGFR<sup>high</sup>). Kaplan–Meier analysis for overall survival (OS, **c**) and time to relapse (TTR, **d**) was

performed according to EGFR levels. Statistical analysis by log-rank Mantel–Cox test. **e**, The CONSORT diagram shows the patient flow through the clinical trial. **f**, H&E and IHC analyses showing high EGFR expression in formalin-fixed paraffin embedded (patient IDs A–L) tissue sections from HCC. All of the biopsies were obtained from liver tumours when patients were diagnosed or treated with primary hepatectomy. Scale bars, 50 µm

[Source data](#).

### **Extended Data Fig. 13 Clinical response of patients with lenvatinib-resistant advanced HCC to the combination of lenvatinib plus gefitinib.**

**a**, Schematic representation of the treatment timeline of patient A with HCC after diagnosis. The dates and treatments administered are annotated along the arrow. MRI scans of patient A were performed during lenvatinib monotherapy (IM-T1), after lenvatinib monotherapy (IM-T2) and after combination therapy (IM-T3). The red label indicates the duration of lenvatinib treatment; the green label indicates the duration of gefitinib treatment. **b**, Serum AFP level of patient A was monitored during lenvatinib monotherapy (AFP-T1), after lenvatinib monotherapy (AFP-T2), and after combination therapy of lenvatinib plus gefitinib (AFP-T3). **c**, Schematic representation of the treatment timeline of patient C with HCC after diagnosis. **d**, Serum AFP level of patient C during the indicated treatment. **e–n**, Serum AFP level of patients B, D–L was monitored before or during lenvatinib monotherapy (AFP-T1), after lenvatinib monotherapy (AFP-T2) and after combination therapy of lenvatinib plus gefitinib (AFP-T3). **o–x**, MRI scans of patients with HCC were performed before or during lenvatinib monotherapy (IM-T1), after lenvatinib monotherapy (IM-T2) and after combination therapy of lenvatinib plus gefitinib (IM-T3). The tumour size of the major target lesions measured by MRI scan in patients with HCC was plotted on the *y* axes against indicated time point and treatment duration on the *x* axes

[Source data](#).

# **Supplementary information**

## **Supplementary Information**

This file contains Supplementary Figures 1-2 (the uncropped blots and FACS gating strategy) and the trial protocol.

## **Reporting Summary**

## **Supplementary Table 1**

Characteristics of 12 recruited HCC patients.

## **Supplementary Table 2**

Kinome-centred CRISPR screen in SNU449 cells.

## **Supplementary Table 3**

Primer sequences used in qRT-PCR analyses.

## **Supplementary Table 4**

Information for antibodies used in the Flow cytometry assays.

## **Peer Review File**

## **Source data**

### **Source Data Fig. 2**

### **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](#)

[Source Data Extended Data Fig. 12](#)

[Source Data Extended Data Fig. 13](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article



Check for  
updates

Cite this article

Jin, H., Shi, Y., Lv, Y. *et al.* EGFR activation limits the response of liver cancer to lenvatinib. *Nature* **595**, 730–734 (2021).  
<https://doi.org/10.1038/s41586-021-03741-7>

## [Download citation](#)

- Received: 08 May 2020
- Accepted: 21 June 2021
- Published: 21 July 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03741-7>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

## Associated Content

Collection

[Cancer at Nature Portfolio](#)

Advertisement

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

- Article
- [Published: 26 May 2021](#)

# Enhancer release and retargeting activates disease-susceptibility genes

- [Soohwan Oh<sup>1</sup>](#),
- [Jiaofang Shao](#) ORCID: [orcid.org/0000-0003-0177-4844<sup>2</sup>](#) na1,
- [Joydeep Mitra](#) ORCID: [orcid.org/0000-0001-5806-7329<sup>3</sup>](#) na1,
- [Feng Xiong<sup>2</sup>](#),
- [Matteo D'Antonio](#) ORCID: [orcid.org/0000-0001-5844-6433<sup>4</sup>](#),
- [Ruoyu Wang](#) ORCID: [orcid.org/0000-0002-3644-1284<sup>2,5</sup>](#),
- [Ivan Garcia-Bassets](#) ORCID: [orcid.org/0000-0002-6941-9585<sup>6</sup>](#),
- [Qi Ma](#) ORCID: [orcid.org/0000-0002-8705-4460<sup>1</sup>](#),
- [Xiaoyu Zhu<sup>2</sup>](#),
- [Joo-Hyung Lee<sup>2</sup>](#),
- [Sreejith J. Nair<sup>1</sup>](#),
- [Feng Yang<sup>1</sup>](#),
- [Kenneth Ohgi<sup>1</sup>](#),
- [Kelly A. Frazer<sup>4,7</sup>](#),
- [Zhengdong D. Zhang<sup>3</sup>](#),
- [Wenbo Li](#) ORCID: [orcid.org/0000-0002-9042-5664<sup>2,5</sup>](#) &
- [Michael G. Rosenfeld](#) ORCID: [orcid.org/0000-0002-1572-156X<sup>1</sup>](#)

[Nature](#) volume 595, pages 735–740 (2021) [Cite this article](#)

- 12k Accesses
- 104 Altmetric

- [Metrics details](#)

## Subjects

- [Gene regulation](#)
- [Genome-wide association studies](#)
- [Transcriptional regulatory elements](#)

## Abstract

The functional engagement between an enhancer and its target promoter ensures precise gene transcription<sup>1</sup>. Understanding the basis of promoter choice by enhancers has important implications for health and disease. Here we report that functional loss of a preferred promoter can release its partner enhancer to loop to and activate an alternative promoter (or alternative promoters) in the neighbourhood. We refer to this target-switching process as ‘enhancer release and retargeting’. Genetic deletion, motif perturbation or mutation, and dCas9-mediated CTCF tethering reveal that promoter choice by an enhancer can be determined by the binding of CTCF at promoters, in a cohesin-dependent manner—consistent with a model of ‘enhancer scanning’ inside the contact domain. Promoter-associated CTCF shows a lower affinity than that at chromatin domain boundaries and often lacks a preferred motif orientation or a partnering CTCF at the cognate enhancer, suggesting properties distinct from boundary CTCF. Analyses of cancer mutations, data from the GTEx project and risk loci from genome-wide association studies, together with a focused CRISPR interference screen, reveal that enhancer release and retargeting represents an overlooked mechanism that underlies the activation of disease-susceptibility genes, as exemplified by a risk locus for Parkinson’s disease (*NUCKS1–RAB7L1*) and three loci associated with cancer (*CLPTM1L–TERT*, *ZCCHC7–PAX5* and *PVT1–MYC*).

## 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

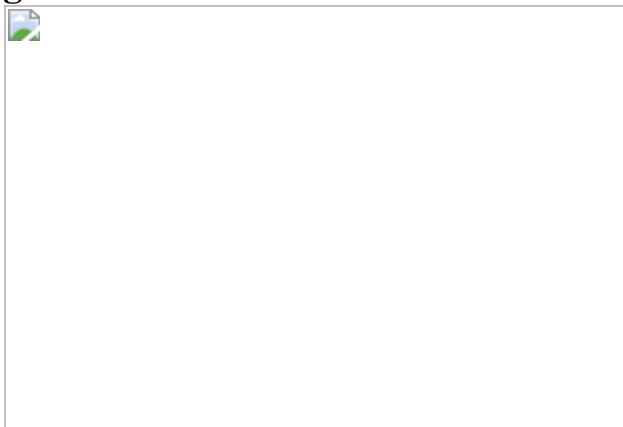
[Rent or Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

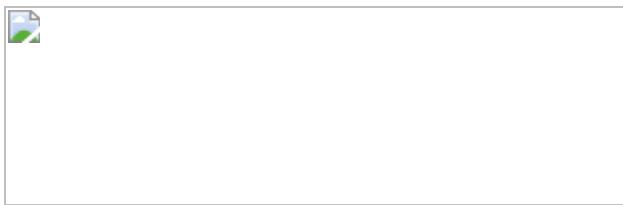
**Fig. 1: Promoter deletion causes ERR.**



**Fig. 2: CTCF binding at promoters is important for their functional engagement with enhancers.**



**Fig. 3: ERR is a potentially common paradigm underlying oncogene activation.**



**Fig. 4: ERR represents an overlooked molecular event that can explain disease susceptibility.**



## Data availability

The datasets generated by the current study are available in the GEO ([GSE115604](#)) and the Sequence Read Archive ([PRJNA412021](#)). For analyses of cancer mutations and GTEx data, codes are deposited in GitHub (<https://github.com/wblilab-uth/ERR-project> and <https://github.com/zdz-lab/ERR>). Other analyses of ChIP-seq peaks, 4C-seq or gene expression in the current study used standard bioinformatics tools and codes, which are available upon request. [Source data](#) are provided with this paper.

# References

1. 1.

Schoenfelder, S. & Fraser, P. Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

2. 2.

Hsieh, T. S. et al. Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *Mol. Cell* **78**, 539–553 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

3. 3.

Gasperini, M. et al. A genome-wide framework for mapping gene regulation via cellular genetic screens. *Cell* **176**, 377–390 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

4. 4.

van Arensbergen, J., van Steensel, B. & Bussemaker, H. J. In search of the determinants of enhancer-promoter interaction specificity. *Trends Cell Biol.* **24**, 695–702 (2014).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

5. 5.

Li, W. et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

6. 6.

Arner, E. et al. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science* **347**, 1010–1014 (2015).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

7. 7.

The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

[ADS](#) [PubMed Central](#) [Article](#) [CAS](#) [PubMed](#) [Google Scholar](#)

8. 8.

Rao, S. S. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

9. 9.

Nora, E. P. et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

10. 10.

Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

11. 11.

Li, Y. et al. The structural basis for cohesin-CTCF-anchored loops. *Nature* **578**, 472–476 (2020).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

12. 12.

Lucas, J. S., Zhang, Y., Dudko, O. K. & Murre, C. 3D trajectories adopted by coding and regulatory DNA elements: first-passage times for genomic interactions. *Cell* **158**, 339–352 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

13. 13.

Guo, Y. et al. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

14. 14.

Jin, E. H. et al. Association between promoter polymorphisms of TFF1, TFF2, and TFF3 and the risk of gastric and diffuse gastric cancers in a Korean population. *J. Korean Med. Sci.* **30**, 1035–1041 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

15. 15.

Weinhold, N., Jacobsen, A., Schultz, N., Sander, C. & Lee, W. Genome-wide analysis of noncoding regulatory mutations in cancer. *Nat. Genet.* **46**, 1160–1165 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

16. 16.

Rheinbay, E. et al. Analyses of non-coding somatic drivers in 2,658 cancer whole genomes. *Nature* **578**, 102–111 (2020).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

17. 17.

Cho, S. W. et al. Promoter of lncRNA gene *PVT1* is a tumor-suppressor DNA boundary element. *Cell* **173**, 1398–1412 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

18. 18.

Zhang, W. et al. A global transcriptional network connecting noncoding mutations to changes in tumor gene expression. *Nat. Genet.* **50**, 613–620 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

19. 19.

Tate, J. G. et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

20. 20.

Satake, W. et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nat. Genet.* **41**, 1303–1307 (2009).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

21. 21.

Liu, Z. et al. LRRK2 phosphorylates membrane-bound Rabs and is activated by GTP-bound Rab7L1 to promote recruitment to the trans-Golgi network. *Hum. Mol. Genet.* **27**, 385–395 (2018).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

22. 22.

Blackwood, E. M. & Kadonaga, J. T. Going the distance: a current view of enhancer action. *Science* **281**, 60–63 (1998).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

23. 23.

Rao, S. S. P. et al. Cohesin loss eliminates all loop domains. *Cell* **171**, 305–320 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

24. 24.

Schwarzer, W. et al. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).

[ADS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

25. 25.

Li, W. et al. Condensin I and II complexes license full estrogen receptor  $\alpha$ -dependent enhancer activation. *Mol. Cell* **59**, 188–202 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

26. 26.

Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

27. 27.

Stadhouders, R. et al. Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions. *Nat. Protocols* **8**, 509–524 (2013).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

28. 28.

van de Werken, H. J. et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat. Methods* **9**, 969–972 (2012).

[PubMed](#) [Article](#) [CAS](#) [Google Scholar](#)

29. 29.

Krijger, P. H. L., Geeven, G., Bianchi, V., Hilvering, C. R. E. & de Laat, W. 4C-seq from beginning to end: a detailed protocol for sample preparation and data analysis. *Methods* **170**, 17–32 (2020).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

30. 30.

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

31. 31.

Hayward, N. K. et al. Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175–180 (2017).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

32. 32.

Schuijers, J. et al. Transcriptional dysregulation of MYC reveals common enhancer-docking mechanism. *Cell Rep.* **23**, 349–360 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

33. 33.

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

34. 34.

MacLeod, D. A. et al. RAB7L1 interacts with LRRK2 to modify intraneuronal protein sorting and Parkinson’s disease risk. *Neuron* **77**, 425–439 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

35. 35.

Latourelle, J. C. et al. Large-scale identification of clinical and genetic predictors of motor progression in patients with newly diagnosed Parkinson’s disease: a longitudinal cohort study and validation. *Lancet Neurol.* **16**, 908–916 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

36. 36.

Pierce, S. E., Tyson, T., Booms, A., Prahl, J. & Coetzee, G. A. Parkinson’s disease genetic risk in a midbrain neuronal cell line. *Neurobiol. Dis.* **114**, 53–64 (2018).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

37. 37.

DeBoever, C. et al. Large-scale profiling reveals the influence of genetic variation on gene expression in human induced pluripotent stem cells. *Cell Stem Cell* **20**, 533–546 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

38. 38.

Panopoulos, A. D. et al. The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. *Cell Res.* **22**, 168–177 (2012).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

39. 39.

Panopoulos, A. D. et al. iPSCORE: a resource of 222 iPSC lines enabling functional characterization of genetic variation across a variety of cell types. *Stem Cell Rep.* **8**, 1086–1100 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

40. 40.

Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

41. 41.

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

42. 42.

van de Geijn, B., McVicker, G., Gilad, Y. & Pritchard, J. K. WASP: allele-specific software for robust molecular quantitative trait locus discovery. *Nat. Methods* **12**, 1061–1063 (2015).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

43. 43.

Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10.33 (2013).

[Article](#) [Google Scholar](#)

44. 44.

Mayba, O. et al. MBASED: allele-specific expression detection in cancer tissues and cell lines. *Genome Biol.* **15**, 405 (2014).

[PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

45. 45.

Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

46. 46.

Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

47. 47.

Battle, A., Brown, C. D., Engelhardt, B. E. & Montgomery, S. B. Genetic effects on gene expression across human tissues. *Nature* **550**,

204–213 (2017).

[ADS](#) [PubMed](#) [Article](#) [Google Scholar](#)

48. 48.

MacArthur, J. et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45**, D896–D901 (2017).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

49. 49.

Li, M. J. et al. GWASdb v2: an update database for human genetic variants identified by genome-wide association studies. *Nucleic Acids Res.* **44**, D869–D876 (2016).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

50. 50.

Nair, S. J. et al. Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. *Nat. Struct. Mol. Biol.* **26**, 193–203 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

51. 51.

Krietenstein, N. et al. Ultrastructural details of mammalian chromosome architecture. *Mol. Cell* **78**, 554–565 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

52. 52.

Fudenberg, G. et al. Formation of chromosomal domains by loop extrusion. *Cell Rep.* **15**, 2038–2049 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

53. 53.

Hatzis, P. & Talianidis, I. Dynamics of enhancer–promoter communication during differentiation-induced gene activation. *Mol. Cell* **10**, 1467–1477 (2002).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

This work is supported by the Howard Hughes Medical Institute and NIH grants (DK018477, DK039949, HL150521 and NS093066) to M.G.R.; by NIH grants (K22CA204468, R21GM132778, R01GM136922 and ‘4D nucleome’ U01HL156059), a University of Texas Rising STARS Award, the Welch Foundation (AU-2000-20190330), the Gulf Coast Consortium John S. Dunn Foundation and a Cancer Prevention and Research Institute of Texas (CPRIT) Award (RR160083) to W.L.; by NIH grants (HG008118, HL107442, DK105541, DK112155 and NSF-CMMI division award 1728497) to K.A.F.; and by NIH NHGRI grants (R01HG008153 and R01HG008153-S1) to Z.D.Z. W.L. is a CPRIT Scholar in Cancer Research. M.G.R. is an investigator with the Howard Hughes Medical Institute. We thank J. Hightower for assistance with figure preparation, and we acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin (<http://www.tacc.utexas.edu>) for providing high-performance computing resources for part of this work.

## Author information

### Author notes

1. These authors contributed equally: Jiaofang Shao, Joydeep Mitra

## Affiliations

1. Howard Hughes Medical Institute, Department and School of Medicine, University of California San Diego, La Jolla, CA, USA  
Soohwan Oh, Qi Ma, Sreejith J. Nair, Feng Yang, Kenneth Ohgi & Michael G. Rosenfeld
2. Department of Biochemistry and Molecular Biology, McGovern Medical School, University of Texas Health Science Center, Houston, TX, USA  
Jiaofang Shao, Feng Xiong, Ruoyu Wang, Xiaoyu Zhu, Joo-Hyung Lee & Wenbo Li
3. Department of Genetics, Albert Einstein College of Medicine, New York, NY, USA  
Joydeep Mitra & Zhengdong D. Zhang
4. Department of Pediatrics, School of Medicine, University of California San Diego, La Jolla, CA, USA  
Matteo D'Antonio & Kelly A. Frazer
5. Graduate School of Biomedical Sciences, University of Texas MD Anderson Cancer Center and UTHealth, Houston, TX, USA  
Ruoyu Wang & Wenbo Li
6. Department of Medicine, School of Medicine, University of California San Diego, La Jolla, CA, USA  
Ivan Garcia-Bassets
7. Institute for Genomic Medicine, University of California San Diego, La Jolla, CA, USA  
Kelly A. Frazer

Authors

1. Soohwan Oh

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Jiaofang Shao

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Joydeep Mitra

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Feng Xiong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Matteo D'Antonio

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Ruoyu Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Ivan Garcia-Bassets

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Qi Ma

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Xiaoyu Zhu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Joo-Hyung Lee

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Sreejith J. Nair

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Feng Yang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Kenneth Ohgi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Kelly A. Frazer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Zhengdong D. Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Wenbo Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Michael G. Rosenfeld

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

W.L., S.O. and M.G.R. conceived the project. S.O. and W.L. conducted most experiments, with contributions from X.Z., J.-H.L., S.J.N., F.Y. and K.O. J.S., W.L. and S.O. conducted most of the bioinformatic analyses of this paper, except for the GTEx related analyses, which were done by J.M. and Z.D.Z. M.D. and K.A.F. generated and analysed the RNA-seq data for allelic gene expression in human iPS cells. The 4C-seq data were analysed by F.X., R.W. and Q.M. I.G.-B. contributed to the iPS-cell-derived NPC culture and RNA-seq. W.L., M.G.R. and S.O. wrote the manuscript, with input from all authors.

## Corresponding authors

Correspondence to [Wenbo Li](#) or [Michael G. Rosenfeld](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Peter Scacheri 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 Disrupted enhancer–promoter looping alters eRNA and mRNA transcription.

**a**, qRT–PCR results showing the temporal expression kinetics of eRNAs and mRNAs at two loci after treatment with 17-β-oestradiol ( $E_2$ ). mRNAs measured by intronic primers ( $n = 3$ ; biological replicates). Shaded error band represents mean  $\pm$  s.d. **b**, 4C-seq heat maps using *TFF1e*, *P2RY2e* or *KCNK5e* as the viewpoints showing the effects of *RAD21* siRNA depletion on the E-P looping events; yellow and pink highlights depict enhancers and promoters, respectively. The quantified chromosome contact frequency between enhancer and promoter from two replicates of 4C-seq is shown in the bar graph (bottom). **c**, Expression of *TFF1* mRNA (*TFF1m*) and *TFF1* enhancer RNA (*TFF1e*) in wild-type ( $n = 6$ ) and *TFF1p-KO* MCF-7 isogenic cell clones ( $n = 14$ ) was checked by qRT–PCR. The *FOXC1* gene that locates on another chromosome was examined as a control. The WT5 and KO6 lines were used for many subsequent experiments (for example, 4C-seq, GRO-seq and ChIP). Each bar represents data from an independent cell clone and was made from  $n = 3$  data points of technical replicates. The data presented here represent three biological replicates. **d**, qRT–PCR results showing the expression of *TFF1* eRNA and mRNA in wild-type MCF-7 cells versus *TFF1p-KO* cells with two different gRNA pairs (*TFF1p-KO\_gRNA1/2* and *TFF1p-KO\_gRNA3/4*) ( $n = 3$  data points of technical replicates; representative of three independent experiments). **e**, ChIP–qPCR data indicating the binding of RNA Pol II and p300 at the *TFF1* enhancer or promoter in wild-type compared to *TFF1p-KO* cells ( $n = 3$  data points of technical replicates; representative of two biological replicates); their binding at an enhancer region near the *NRIP1* gene (that is, *NRIP1e*) is shown as a control. Data are mean  $\pm$  s.d.; \* $P < 0.05$ ; \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Student's *t*-test.

### Source data

### Extended Data Fig. 2 Promoter deletion affects gene and eRNA transcription in its chromosomal neighbourhood at the *P2RY2*,

## **KCNK5 and PGR loci.**

**a, b**, qRT–PCR showing the mRNA and eRNA expression of *P2RY2* (**a**) and the expression of neighbouring genes (*FCHSD2* and *P2RY6*) (**b**) in wild-type ( $n = 2$ ) versus *P2RY2* promoter KO (*P2RY2p-KO*) ( $n = 2$ ) independent isogenic clones of MCF-7 cells. Each bar represents an independent cell clone ( $n = 3$  data points of technical replicates; representative of two independent experiments). **c**, Top, Hi-C contact matrix and ChIP–seq tracks of CTCF and p300 showing the topology of the chromosomal neighbourhood of *P2RY2* locus. Bottom, UCSC browser screenshots showing the binding of CTCF, cohesin subunit (that is, RAD21) and ER $\alpha$  at *P2RY2* and surrounding gene promoters. **d, e**, qRT–PCR showing the mRNA and eRNA expression of *KCNK5* (**d**) and the expression of neighbouring genes (*KCNK17* and *SAYSD1*) (**e**) in wild-type ( $n = 2$ ) versus *KCNK5* promoter KO (*KCNK5p-KO*) ( $n = 2$ ) independent isogenic cell clones of MCF-7 cells. Each bar represents an independent cell clone ( $n = 3$  data points of technical replicates; representative of two independent experiments). **f**, Top, Hi-C contact matrix and ChIP–seq tracks of CTCF and p300 showing the topology of the chromosomal neighbourhood of the *KCNK5* locus. Bottom, UCSC browser screenshots showing the binding of CTCF, cohesin subunit (that is, RAD21) and ER $\alpha$  at *KCNK5* and surrounding gene promoters. **g, h**, qRT–PCR showing the mRNA and eRNA expression of *PGR* (**g**) and the expression of neighbouring genes (*ARHGAP42* and *TRPC6*) (**h**), in wild-type ( $n = 2$ ) versus *PGR* promoter KO (*PGRp-KO*) ( $n = 2$ ) independent isogenic cell clones of MCF-7 cells. Each bar represents an independent cell clone ( $n = 3$  data points of technical replicates; representative of two independent experiments). **i**, Top, Hi-C contact matrix and ChIP–seq tracks of CTCF and p300 showing the topology of the chromosomal neighbourhood of the *PGR* locus. Bottom, UCSC browser screenshots showing CTCF, cohesin and ER $\alpha$  binding at *PGR* and surrounding gene promoters. Data are mean  $\pm$  s.d.; \* $P < 0.05$ ; \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Student's *t*-test.

[Source data](#)

[Extended Data Fig. 3 Cohesin knockdown affects gene and eRNA transcription through coordinating looping, without](#)

## changing histone methylation.

**a**, qRT–PCR results showing efficient knockdown of *RAD21* by siRNA, which does not affect the mRNA level of *ESR1* (encoding ER $\alpha$ ) ( $n = 3$  data points of technical replicates; representative of two independent experiments). **b**, **c**, qRT–PCR results showing that si*RAD21* decreases the 17- $\beta$ -oestradiol-induced activation of coding genes (**b**), but upregulates the eRNAs next to these genes (**c**) ( $n = 3$  data points of technical replicates; representative of two independent experiments). **d**, MA plots showing the differential expression of RefSeq genes (left) or eRNAs (right) in cells with depletion of *RAD21* versus cells transfected with control siRNA (that is, si*CTL*). Red dots represent significantly changed genes or eRNAs (fold change  $> 1.5$ ; FDR  $< 0.05$ ). The purple bars indicate twofold change. **e**, UCSC genome browser screen shots of GRO-seq results in si*RAD21*-versus si*CTL*-transfected MCF-7 cells at *KCNK5* (left) and *P2RY2* (right) loci; p300 ChIP–seq serves to indicate active enhancers. Yellow highlights denote gene regions; pink highlights denote enhancer regions. **f**, Diagram showing an oppositely regulated transcription of many eRNAs and mRNAs caused by disruption of E-P looping (that is, si*RAD21*). **g**, The tag density plots showed no significant difference of histone marks for promoters (H3K4me3 ChIP–seq) or enhancers (H3K4me1 ChIP–seq). qPCR data represent mean  $\pm$  s.d.; \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Student's *t*-test.

### Source data

## Extended Data Fig. 4 Deletion of the *TFF1* promoter alters the transcription of the cognate *TFF1* enhancer and neighbouring genes.

**a**, Hi-C contact matrix in GM12878 cells (by Juicebox), CTCF-mediated chromatin loops (by ChIA-PET) and ChIP–seq tracks of CTCF and p300 in MCF-7 cells showing the topology (multiple contact domains) of the chromosomal neighbourhood of the *TFF1* locus. Yellow triangles in the Hi-C map denote contact domains; domains are numbered for simplicity. The contact domains containing genes analysed in **b** are highlighted in yellow. **b**, Box plots of gene expression outside of the hosting contact domain of *TFF1* generated by qRT–PCR in wild-type ( $n = 6$ ) and *TFF1p-KO* ( $n = 14$ )

isogenic clones of MCF-7 cells. The box plot centre lines represent medians; box limits indicate the 25th and 75th percentiles as determined by R software; and whiskers extend  $1.5 \times$  IQR from the 25th and 75th percentiles. **c**, qRT–PCR results showing the expression of *TFF3* and *RSPH1* mRNA in wild-type MCF-7 cells as compared to *TFF1p-KO* cells with two different gRNA pairs (*TFF1p-KO\_gRNA1/2* and *TFF1p-KO\_gRNA3/4*) ( $n = 3$  data points of technical replicates; representative of three independent experiments). **d**, ChIP–qPCR data indicating the binding of RNA Pol II and p300 at *TFF3p* in wild-type cells as compared to *TFF1p-KO* cells ( $n = 3$  data points of technical replicates; representative of two independent experiments). **e**, qRT–PCR results showing the expression of *TFF1e* and *TFF3m* after knockdown of *RAD21* by siRNA in wild-type MCF-7 cells versus *TFF1p-KO* cells ( $n = 3$  data points of technical replicates; representative of two independent experiments). **f**, qRT–PCR results showing the expression *TFF1e*, *TFF1m* and *TFF3m* in wild-type MCF-7 cells versus cells with deletion of the *TFF1* enhancer (*TFF1e-KO*), *TFF1* promoter (*TFF1p-KO*) or both (*TFF1e/p-DKO*) ( $n = 3$  data points of technical replicates; representative of three independent experiments). The diagram shows the *TFF1e* deletion in *TFF1p-KO* cells. Data are mean  $\pm$  s.d.; \* $P < 0.05$ ; \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Student's *t*-test.

#### Source data

#### Extended Data Fig. 5 Epigenomic features of the enhancers and promoters that are putative looping pairs, as revealed by their opposite regulation after siRAD21 treatment.

**a**, ChIP–seq profile plots were generated using published data in MCF-7 cells (Supplementary Table 6), showing the differential enrichment of chromatin-looping-related structural factors, including CTCF, cohesin subunit (that is, RAD21), cohesin loading factor (that is, NIPBL) and Mediator subunit (that is, MED1) on the promoters that are downregulated by siRAD21 treatment, as well as on enhancers for which eRNAs are increased by siRAD21. Also refer to Extended Data Fig. 3. **b**, ChIP–seq profile plots showing the differential enrichment of factors that are relevant to enhancer function in MCF-7 cells, including those of p300, ER $\alpha$ , FOXA1

and GATA3. **c**, Diagram showing the non-stoichiometric distribution of chromatin structural factors and other transcription-related factors on functional E-P pairs. These E-P pairs denote RAD21-regulated enhancers (super-activated by siRAD21) and RAD21-dependent promoters (downregulated by siRAD21).

### Extended Data Fig. 6 Promoters exhibit a higher frequency and affinity of CTCF binding than enhancers.

**a**, Percentages of promoters (P), enhancers (E) or random genomic regions (R) that contain a CTCF motif (left) or undergo CTCF binding (right) in MCF-7 cells (based on ENCODE data). The cartoons at the bottom indicate that promoters were further analysed for their features in **c**, on the basis of their distance to an adjacent enhancer. **b**, Similar to **a** but using GM12878 data. **c**, Percentage of MCF-7 promoters that contain a CTCF motif (left) or undergo CTCF binding (right) for three categories of promoters: promoters that have a super-enhancer, a typical enhancer or no enhancer in their 50-kb genomic proximity (cartoon at bottom of **a**). **d**, Similar to **c** but using GM12878 data. All *P* values in **a–e** were generated by two-sided Fisher's exact test. **e**, Top, diagram showing the design of constructs. The binding between GCN4 and ScFV will bring CTCF to the sites at which dCas9/gRNA binds. Bottom, western blots showing the expression of scFv-CTCF-V5 (wild type or Y226A/F228A mutant) and dCas9-10xGCN4 after lentiviral transduction (uncropped images in Supplementary Fig. 1; representative of two independent experiments). **f**, ChIP-qPCR using a V5 antibody indicating comparable binding of V5-tagged wild-type CTCF or CTCF(Y226A/F228A) on *TFF3p* (*n* = 3 data points of technical replicates; representative of two independent experiments). Control gRNAs did not bring the V5-tagged CTCF to *TFF3p*. ChIP was performed with MCF-7 cells expressing dCas9-10xGCN4 and scFV-CTCF (wild type or Y226A/F228A mutant) as indicated. qPCR data in **f** represent mean ± s.d.; \**P* < 0.05; \*\**P* < 0.01, two-tailed Student's *t*-test. The *P* values were based on comparisons between the same-coloured bar (either wild-type or mutant CTCF) in TFF3p-gRNA1/gRNA2 versus CTL-gRNA1/gRNA2 conditions.

[Source data](#)

## Extended Data Fig. 7 Enhancer scanning and alternative models to interpret ERR and enhancer–promoter functional engagement.

**a**, Model describing the process through which an activated enhancer scans its chromosomal neighbourhood inside a contact domain to functionally engage with its cognate promoter target. This enhancer scanning process involves promoter-bound CTCF (green ovals) and is compatible with cohesin-mediated extrusion (dashed lines) in wild-type cells, but it continues to operate in the absence of the cognate promoter (that is, P1) owing to deletion or disease mutation. A hypothetical ‘enhancer engagement’ time is depicted that reflects the relative amount of time in which the active enhancer engages with the neighbouring promoters, correlating with their expression levels. **b**, There are alternative, non-exclusive models that could be largely consistent with our results. For example, one of the other possibilities is a ‘positional effect’ model. In this model, an active enhancer may engage with its preferred promoter, and the two are retained in a transcription-associated E-P loop, quite probably in a nuclear environment such as an interchromatin granule<sup>50</sup>, transcription factory or other perhaps phase-separated structure (pink coloured area). Upon functional loss (deletion, mutation or CTCF loss) of the original promoter, this ‘positional’ effect will follow the enhancer but it will engage with the secondary choices (P3 in the model; for example, *TFF3* as compared to *TFF1*, or *PAX5* as compared to *ZCCHC7*, or *TERT* as compared to *CLPTM1L*). But at a lower frequency or in a small percentage of single cells among a population, the enhancer does engage with other target promoters (such as P2, dotted black line). When the secondary choice is also lost, the entire enhancer will now select the third choice—for example, P2 in the model—to activate the promoter as its newest hierarchical choice. Overall, a few specific results from our study are in better support of the enhancer scanning model. First, it is consistent with the observation that for all single-cell colonies investigated (that is, 14 independent cell clones of *TFF1p-KO* in Extended Data Fig. 1), they consistently exhibited the highest expression of *TFF3* as the new hierarchical promoter choice, precluding models suggesting stochastic promoter choice. A positional effect model may predict that the enhancer

and the target promoter are engaged so that the inhibition of one will reduce the other. This is true in that deletion of *TFF1e* inhibited *TFF1p* (Extended Data Fig. 4f). However, CRISPRi of *TFF1p* resulted in higher activity of *TFF1e* and higher expression of *TFF3* (Fig. 3b). This data better support an enhancer scanning model—a dynamic enhancer–promoter interaction process inside the contact domain, in which *TFF1p* inhibition makes the ‘*TFF1e*-in-action’ preferentially interact with the next target in the hierarchy. Second, *TFF1e* exhibited quite broad chromatin interaction throughout the contact domain (for example, Fig. 1e, by 4C-seq), which is in accord with its ‘scanning’ behaviour. Furthermore, high-resolution Micro-C<sup>2,51</sup> data showed that many ‘stripe loops’ formed in between enhancer and promoter at a finer scale (<50 kb or sometimes <10 kb), consistent with the suggestion that at least a subset of enhancers and promoters are actively scanning or extruding. The positional effect model is fundamentally compatible with the enhancer scanning model. The enhancer scanning model reflects a dynamic process in which the enhancer initiates target searching and it finds a major target as well as many additional minor targets. Were we to snapshot this dynamic process, then at every time point that an enhancer engages with one of its potential targets (for example, *TFF3p* or *TFF1p*), the target promoter would be repositioned closer to the enhancer. Our data suggest that the scanning process of the enhancer requires the cohesin complex, as its depletion by RNA interference reproduced many of the phenomena that are seen in promoter-knockout cells (for example, Fig. 1b, Extended Data Figs. 3, 4e). This is consistent with a proposed loop-extrusion model, rather analogous to the mechanism by which cohesin facilitates the formation of the larger TAD structures<sup>23,24,52</sup>, requiring opposing CTCF motif orientation and high-affinity CTCF binding. However, it cannot be excluded that other chromatin remodelers, or Pol II itself, are the critical driver of enhancer scanning and E-P engagement. In support of this, inhibition of Pol II elongation partially reduced the promoter-centred stripe loops in mouse embryonic stem cells in a Micro-C study<sup>2</sup>. Finally, it is noteworthy that the scanning concept that enhancers (either the DNA–protein complex as an entirety or specific transcription apparatus such as Pol II) travel along chromatin to reach target genes has been extensively discussed as one of the classic models to interpret enhancer activity<sup>22,53</sup>.

## Extended Data Fig. 8 CTCF binding on promoters, but not on enhancers, dictates enhancer–promoter choice.

**a**, Bar graph showing the numbers of RAD21-regulated enhancers and promoters (same set used in Fig. 1b, Extended Data Fig. 5) that engage in chromatin loops in MCF-7 cells as discovered by analysing MCF-7 ChIA-PET data from ENCODE, as compared to their low incidence of looping in K562 ChIA-PET data. **b**, Coloured map showing that, among the functionally looped E-P pairs discovered in **a**, only certain percentages are bound by ERα, FOXA1 or CTCF by ChIP-seq. **c**, The numbers of looped E-P pairs that display various motif directionality of CTCF and FOXA1. The four categories below the plot describe the existence of CTCF or FOXA1 motifs on enhancers (E) or promoters (P); the colours indicate the combinations of motif directionality. For example, ‘CTCF (E) + CTCF (P)’ denotes the coincidence of CTCF motifs on both promoters and enhancers. There was no obvious orientation preference of binding motifs for CTCF. FOXA1 binds to functionally looped enhancers more frequently than CTCF (Extended Data Figs. 5b, 8b), but it also exhibits no preferred motif orientation. **d**, Diagrams showing the position and directionality of CTCF-binding peaks and motifs in the four E-P pairs that we have extensively studied in this work. The colour motif sequence indicates directionality (black for left-pointing motifs, red for right-pointing motifs); the letters in each motif sequence are coloured to match the core CTCF motif nucleotides shown in the canonical CTCF motif by Jaspar (upper left logo); the motif scores after each motif sequence were calculated by the FIMO motif toolset. **e**, qRT-PCR results showing the expression of *TFF1* eRNA and mRNA in wild-type MCF-7 cells versus cells with deletion of the *TFF1* enhancer (*TFF1e-KO*), and MCF-7 cells with deletion of a CTCF peak inside the *TFF1* enhancer (*TFF1e/CTCF-KO*; Supplementary Table 1, Supplementary Fig. 3) ( $n = 3$  data points of technical replicates; representative of two independent experiments). Data are mean  $\pm$  s.d.; \* $P < 0.05$ ; \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Student’s *t*-test.

[Source data](#)

## Extended Data Fig. 9 Cancer mutations in gene promoter regions, and their link to oncogene activation through ERR.

**a**, Plots showing the distribution of ICGC cancer whole-genome sequencing (WGS) mutations in the  $\pm 5$  kb regions near gene TSSs. The two numbers in each plot indicate the mutations in the  $\pm 5$  kb of TSSs, as well as the total numbers of mutations in that cancer type, respectively. Cancer type abbreviations can be found in the ICGC web portal (<http://icgc.org/>), and in the Methods. **b**, The percentage of genes containing ICGC annotated promoter mutations or deletions among all RefSeq genes (all gene promoters; total 27,502), oncogene promoters (OPs, from cosmic database; total 315) and neighbouring gene promoters within  $\pm 200$  kb of cosmic oncogenes (oncogene neighbouring promoters (ONPs); total 1,693). Hypergeometric tests were performed to compare the statistical significance of the enrichment of bearing a mutation or deletion for OPs or ONPs as compared to all RefSeq gene promoters (upon different frequency or size of mutations as indicated in the  $x$  axis). For each test, for example, comparing OPs to all RefGene promoters,  $p[X > k]$  was calculated, which denotes a one-sided probability that a higher percentage of oncogene promoters bears mutations or deletions than the percentage of RefSeq promoters. The random variable  $X$  follows a hypergeometric distribution with parameters  $N$  (number of total RefSeq genes),  $K$  (number of total RefSeq genes bearing promoter mutations or deletions),  $n$  (number of total oncogenes or oncogene neighbours),  $k$  (number of total oncogenes or oncogene neighbours that bear promoter mutations or deletions). The five significant  $P$  values shown in the plot (from left to right) are  $9.2 \times 10^{-4}$ ,  $3.3 \times 10^{-3}$ ,  $5.2 \times 10^{-5}$ ,  $5.3 \times 10^{-3}$  and  $5.6 \times 10^{-4}$ , respectively. mut $>= 1$ : any mutations being identified in ICGC; mut $> 1$ : recurrent promoter mutations identified in more than 1 donors in ICGC release 28; del $< 100$  kb: genetic deletions with length smaller than 100 kb. **c**, Occurrence of somatic mutations and deletions at a selected list of ONPs that were included in our CRISPRi screening (see Fig. 3c). In this plot, the cancer cohorts ( $x$  axis) were ranked by the names of original cancer sites, and the  $y$  axis shows the ONP–OP gene pairs. Gene names before ‘ $\rightarrow$ ’ are those of ONPs, whereas those after ‘ $\rightarrow$ ’ are oncogenes. For example, ‘mutation: CLPTM1L  $\rightarrow$  TERT’ (the fourth row) indicates that TERT is an oncogene listed in COSMIC, and the gene promoter of CLPTM1L was identified as an ONP, which contains somatic mutations in several cancer types. The dot size was scaled by the percentage of affected donors in each of that cancer type or cohort (SamplePerc), and dots of red colour show those samples with mutations, whereas blue dots show those

samples with deletion (<100 kb) covering an ONP. The cancer type abbreviations can be found in the ICGC web portal (<http://icgc.org/>) and in the Methods. **d**, qRT–PCR results showing the expression of gene pairs, *PAX5*–*ZCCHC7*, *MYC*–*PVT1* and *SS18L1*–*MTG2*. Each pair consisted of an oncogene (*PAX5*, *MYC* and *SS18L1*) and its adjacent ONP genes (*ZCCHC7*, *PVT1* and *MTG2*). The ONP promoters were inhibited with specific sgRNAs in MCF-7 cells together with dCas9-KRAB (see the diagrams at the top) ( $n = 3$  data points of technical replicates; representative of two independent experiments). **e, f**, Similar to Fig. 3d, these are results for two other ONP–OP loci after ONP deletion. The Hi-C contact matrix at the top shows the relative location of *PAX5*–*ZCCHC7* and *MYC*–*PVT1* gene pairs in a shared contact domain. qRT–PCR results below show the expression of ONP and OP gene mRNAs (two independent clones each) ( $n = 3$  biological replicates). **g–i**, Oncoplot showing the landscapes of ICGC mutations located in three pairs of OPs as well as their ONPs at the *TERT*–*CLPTM1L* locus (**g**), the *PAX5*–*ZCCHC7* locus (**h**) and the *MYC*–*PVT1* locus (**i**). The percentages of mutations between 0.5–1% are denoted as 1%, and those between 0–0.5% are denoted as 0%. Each column in the oncoplot represents a cancer sample. The bar graphs with numbers (for example, 187, 101, 98) in the right side indicate the numbers of donor samples that contain mutations for that specific gene promoter. The labels on top of each panel indicate the total numbers of donors that contain mutations in any of the promoters in that specific locus, for example, 386 (6.27%) is the sum number of donors that contain mutations for any of the promoters shown in **g**. It is noteworthy that it remains a challenge to directly compare oncogene expression changes in tumour samples that carry rare noncoding mutations, because the rarity of matched RNA-seq and genotype (WGS) precludes a robust statistical analysis. qPCR data in **d–f** represent mean  $\pm$  s.d.; \* $P < 0.05$ ; \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Student's *t*-test.

[Source data](#)

[\*\*Extended Data Fig. 10 Chromosomal interaction in three pairs of oncogene and oncogene-neighbourhood genes exhibiting an ERR-like phenotype.\*\*](#)

**a–c**, UCSC and WashU genome browser screenshots of ChIA-PET and GRO-seq results at the *PAX5*–*ZCCHC7* locus (**a**), *MYC*–*PVT1* locus (**b**) and *TERT*–*CLPTM1L* locus (**c**). The putative enhancers and promoters are highlighted in pink and in yellow, respectively. These cells show strong interactions between the eRNA (enhancer) and the original promoters (that is, *ZCCHC7p* in **a**, *PVT1p* in **b** and *CLPTM1Lp* in **c**). **d**, 4C-seq contact matrix and heat map plots showing the chromatin contacts with the *PAX5* promoter as viewpoint, and its change in 293T cells with *ZCCHC7* promoter knockout. The viewpoint *PAX5* promoter is highlighted in yellow, and the neighbouring enhancer in pink. **e**, 4C-seq contact matrix and heat map plots showing the chromatin contacts with the *TERT* promoter as viewpoint, and its change in 293T cells with *CLPTM1L* promoter knockout. The viewpoint *TERT* promoter is highlighted in yellow, and the neighbouring enhancer in pink. **f**, The numbers of ICGC cancer mutations that reside in the motifs of CTCF, ERα, FOXA1 and GATA3, and random genomic regions near gene TSSs.

### [Extended Data Fig. 11 GTEx data analysis identifies potential ERR gene pairs acting in human populations.](#)

**a**, Schematic overview of GTEx data-processing workflow to identify potential ERR events. Also see Supplementary Figs. [8–10](#) and associated notes. **b**, Histogram showing the number of potential ERR gene pairs that appear in one tissue or more than one. **c**, Histogram showing the number of potential ERR events distributed in each tissue. **d**, qRT–PCR results of CRISPRi showing the expression of *NUCKS1* and *RAB7L1* mRNAs with sgRNA-control (iCTL) or sgRNA specific for the putative enhancer (iEnhancer) in two clones of NPCs expressing dCas9-KRAB (see enhancer location in Fig. [4c,d](#)) ( $n = 3$  data points of technical replicates; representative of two independent experiments). **e**, Published CTCF ChIP–seq screen shots in human-iPS-cell-derived neurons or neuronal cell lines (SK-N-SH) in the *NUCKS1* and *RAB7L1* promoters. The ChIP–seq data tracks were generated by the Cistrome data browser (<http://cistrome.org/>). The location of the three SNPs is shown. **f**, Sanger sequencing results of the *NUCKS1* promoter DNA, showing the allele bias of CTCF binding (comparison between the input DNA and ChIP DNA). Data are mean  $\pm$  s.d.; \* $P < 0.05$ ; \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , two-tailed Student's *t*-test.

## Source data

# Supplementary information

## Supplementary Information

This file contains Supplementary Figures 1-11, Supplementary Notes and Supplementary References.

## Reporting Summary

## Supplementary Tables

This file contains Supplementary Tables 1-7 – see the Table Guide worksheet within the file for details.

## Source data

### Source Data Fig. 1

### Source Data Fig. 2

### Source Data Fig. 3

### Source Data Fig. 4

### Source Data Extended Data Fig. 1

### Source Data Extended Data Fig. 2

### Source Data Extended Data Fig. 3

### Source Data Extended Data Fig. 4

[\*\*Source Data Extended Data Fig. 6\*\*](#)

[\*\*Source Data Extended Data Fig. 8\*\*](#)

[\*\*Source Data Extended Data Fig. 9\*\*](#)

[\*\*Source Data Extended Data Fig. 11\*\*](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Oh, S., Shao, J., Mitra, J. *et al.* Enhancer release and retargeting activates disease-susceptibility genes. *Nature* **595**, 735–740 (2021).  
<https://doi.org/10.1038/s41586-021-03577-1>

[Download citation](#)

- Received: 14 June 2018
- Accepted: 23 April 2021
- Published: 26 May 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03577-1>

# Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03577-1>

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

Structural basis of early translocation events on the ribosome

[Download PDF](#)

- Article
- Open Access
- [Published: 07 July 2021](#)

# Structural basis of early translocation events on the ribosome

- [Emily J. Rundlet<sup>1,2</sup>,](#)
- [Mikael Holm<sup>1</sup>,](#)
- [Magdalena Schacherl ORCID: orcid.org/0000-0002-5478-2509<sup>3</sup>](#),
- [S. Kundhavai Natchiar<sup>1</sup>,](#)
- [Roger B. Altman<sup>1</sup>,](#)
- [Christian M. T. Spahn<sup>3</sup>,](#)
- [Alexander G. Myasnikov ORCID: orcid.org/0000-0003-2607-7121<sup>1</sup>](#) &
- [Scott C. Blanchard ORCID: orcid.org/0000-0003-2717-9365<sup>1</sup>](#)

*Nature* volume 595, pages 741–745 (2021) [Cite this article](#)

- 7575 Accesses
- 179 Altmetric
- [Metrics details](#)

## Subjects

- [Cryoelectron microscopy](#)
- [Enzyme mechanisms](#)
- [Ribosome](#)
- [Single-molecule biophysics](#)

## Abstract

Peptide-chain elongation during protein synthesis entails sequential aminoacyl-tRNA selection and translocation reactions that proceed rapidly (2–20 per second) and with a

low error rate (around  $10^{-3}$  to  $10^{-5}$  at each step) over thousands of cycles<sup>1</sup>. The cadence and fidelity of ribosome transit through mRNA templates in discrete codon increments is a paradigm for movement in biological systems that must hold for diverse mRNA and tRNA substrates across domains of life. Here we use single-molecule fluorescence methods to guide the capture of structures of early translocation events on the bacterial ribosome. Our findings reveal that the bacterial GTPase elongation factor G specifically engages spontaneously achieved ribosome conformations while in an active, GTP-bound conformation to unlock and initiate peptidyl-tRNA translocation. These findings suggest that processes intrinsic to the pre-translocation ribosome complex can regulate the rate of protein synthesis, and that energy expenditure is used later in the translocation mechanism than previously proposed.

[Download PDF](#)

## Main

Faithful translocation requires the ribosome to maintain hold of diverse mRNA and tRNA cargo (the tRNA<sub>2</sub>-mRNA module) while simultaneously allowing their rapid movement between the large and small ribosomal subunits (LSU and SSU; 50S and 30S in bacteria, respectively). In bacteria, translocation is mediated by a highly conserved five-domain (DI–DV) GTPase, elongation factor G (EF-G), the mechanism of which has been examined using biochemical<sup>2,3,4</sup>, structural<sup>5,6,7,8,9,10,11,12</sup> and single-molecule fluorescence energy transfer (smFRET) methods<sup>13,14,15,16,17</sup>. EF-G engages the leading edge of pre-translocation (PRE) ribosome complexes bearing peptidyl-tRNA cargo within the aminoacyl (A) site and deacyl-tRNA in the adjacent peptidyl (P) site to facilitate large-scale conformational changes within and between the ribosomal subunits and tRNA substrates (Fig. 1a).

**Fig. 1: Early kinetic and structural intermediate of tRNA<sub>2</sub>-mRNA translocation.**

---

 **figure1**

---

**a**, Schematic of the translocation reaction coordinate in bacteria depicting SSU body-rotation (blue) and head-swivel (purple). tRNAs are coloured on a gradient from the A (green) to P (yellow) to E (orange) sites. The states enclosed in dashed boxes were characterized in this study. Green (donor, uS13, LD550) and red (acceptor, uL1, LD650) circles denote fluorophore positions (see **b**). FA, fusidic acid. **b**, Population FRET histograms showing FRET evolution over time upon EF-G injection with buffer, SPC (3 mM) or fusidic acid (FA, 400  $\mu$ M).  $n$  represents the number of observed molecules. **c**, Overview of the INT1 ribosome structure captured by SPC, coloured as in **a**.

[Full size image](#)

Within the PRE complex, deacyl- and peptidyl-tRNAs can rapidly and spontaneously unlock from their ‘classical’ positions (PRE-C) after peptide-bond formation to achieve multiple ‘hybrid’ states (PRE-H)<sup>18,19</sup>. Hybrid tRNA conformations, which are achieved by independent or concerted migration of the tRNA 3'-CCA termini to adjacent LSU-binding sites<sup>15</sup> coupled to a global SSU rotation<sup>18,20</sup>, markedly lower the energetic barrier to translocation<sup>21</sup>. By contrast, spontaneous unlocking of the tRNA<sub>2</sub>-mRNA module from the SSU is exceedingly rare<sup>3</sup>. Rapid translocation thus requires the action of EF-G, but how EF-G engages the dynamic PRE complex is actively debated.

Once bound to EF-G, the SSU undergoes a scissor-like conformational change in which its body and head domains rotate in opposing directions (SSU body-rotation reversal and forward head-swivel)<sup>6,7,9,10</sup>. SSU head-swivel carries the tRNA anticodons forward to ‘chimeric hybrid’ positions<sup>9,10</sup>. This process is intimately coupled to the sequential disengagement (unlocking) and engagement (relocking) of tRNA 3'-CCA termini and anticodon elements from LSU and SSU contacts, respectively, en route to their final post-translocation (POST) positions in the P and exit (E) sites. The molecular basis of precise, directional tRNA<sub>2</sub>-mRNA movement, and the role of EF-G-catalysed GTP hydrolysis in this process, remain incompletely understood.

To gain insight into how tRNA<sub>2</sub>-mRNA movement is initiated by EF-G, and the role of GTP hydrolysis in translocation, we used smFRET to guide the capture of six cryo-electron microscopy (cryo-EM) structures of the ribosome in both early and late stages of translocation. A new early-intermediate structure stalled by the antibiotic spectinomycin (SPC) revealed that EF-G engages PRE-H ribosome complexes in an active, GTP-bound conformation to initiate unlocking of the peptidyl-tRNA cargo. The energy liberated by GTP hydrolysis thus facilitates downstream unlocking and relocking events in both subunits that ensure precise directional movement of the tRNA<sub>2</sub>-mRNA module.

## smFRET-guided cryo-EM of translocation

We used smFRET to define reaction conditions that slow translocation sufficiently such that intermediate structures could be captured by cryo-EM. As previously described<sup>13</sup>, the antibiotics SPC and fusidic acid specifically stall transitions after EF-G binding (intermediate states 1 (INT1) and 2 (INT2)), without otherwise altering the translocation reaction coordinate (Fig. 1a, b, Extended Data Fig. 1a–g). The FRET efficiency values of states sampled in the presence of SPC and fusidic acid were indistinguishable from those observed in the absence of the drugs<sup>13</sup>. We initiated pre-steady-state reactions using the same conditions used for smFRET before rapid (within 20 s) transfer to cryo-EM grids. This approach yielded six high-resolution (2.3–2.8 Å) ribosome structures programmed with deacyl-tRNA<sup>Phe</sup> and fMet-Phe-Lys-tRNA<sup>Lys</sup> at sequential stages of translocation (Extended Data Fig. 2, Supplementary Table 1), including the first—to our knowledge—structure of EF-G bound to a ribosome in an active conformation before inorganic phosphate (P<sub>i</sub>) release (designated INT1; Fig. 1c). All structures showed density corresponding to codon–anticodon interactions, post-transcriptional tRNA modifications and a tripeptide-linked peptidyl-tRNA, indicating successful complex capture (Extended Data Fig. 3).

The POST complex, containing classical E- and P-site tRNAs (E/E, P/P) was defined as having 0° of inter-subunit rotation or SSU head-swivel (Supplementary Table 2) and the +1-mRNA position was defined as the nucleotide paired with deacyl-tRNA<sup>Phe</sup> position 37. The observed inter-subunit rotation, SSU head-swivel and tRNA positions (Extended Data Fig. 4, Supplementary Video 1)—together with the temporal order of conformational changes evidenced by smFRET<sup>13</sup> (Extended Data Fig. 1h–o, Supplementary Table 3)—were used to elucidate the molecular underpinnings of tRNA<sub>2</sub>-mRNA translocation.

## SSU unlocking initiates spontaneously

Before EF-G engagement, the PRE-C complex (P/P, A/A) exhibited complete SSU shoulder-domain closure around the peptidyl-tRNA cargo<sup>22,23</sup> (Extended Data Fig. 4b). As anticipated<sup>18,24,25</sup>, spontaneous SSU rotation during PRE-H (P/E, A/P) formation remodelled intersubunit bridges B1 and B2 and shifted the nearly universally conserved G19–C56 base pair in the deacyl-tRNA elbow domain to its fully translocated position in the E site<sup>8,11</sup> (Extended Data Fig. 5a–c, Supplementary Videos 2, 3). We observed two PRE-H conformations in which the peptidyl-tRNA 3'-CCA terminus paired with the LSU P site (Extended Data Fig. 5d). These states represent PRE-H2\* and PRE-H1 conformations<sup>15,16</sup>, wherein the G19–C56 pair in the peptidyl-tRNA elbow remains either fixed against the LSU A-site finger (ASF) or swings by approximately 27 Å towards the E site to engage LSU Helix 84 (H84), respectively (Supplementary Table 4). Both PRE-H conformations exhibited increased SSU body-rotation and head-swivel together with tRNA-bend angle changes (Extended Data Figs. 4, 6, Supplementary Table 2, Supplementary Video 4). Indicative of incomplete translocation on the LSU, the universally conserved, potentially catalytic LSU base A2602 was sequestered away from the peptidyl transferase centre (Extended Data Fig. 5d), which is likely to contribute to the reduced reactivities of PRE-H conformations towards the antibiotic puromycin<sup>16,26</sup>.

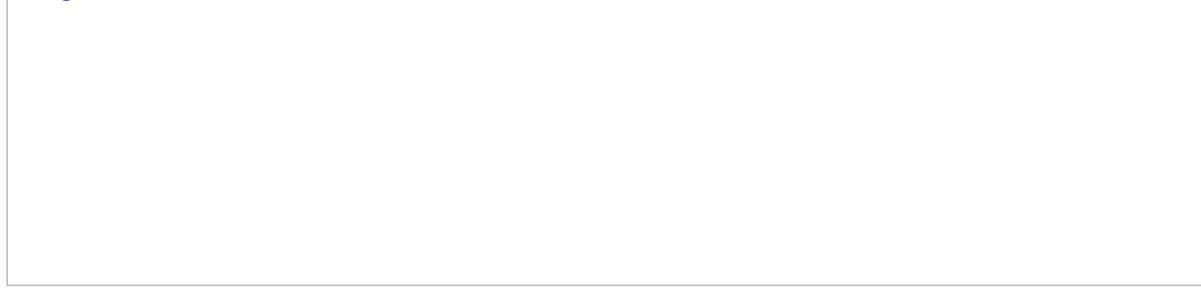
Bending of the tRNA bodies enabled the tRNA anticodons and mRNA to remain in their locked SSU positions during PRE-H formation (Extended Data Fig. 7). The PRE-C–PRE-H1 transition broke SSU shoulder contact with the head domain to partially unlock the grip of the ribosome on the peptidyl-tRNA cargo (Extended Data Fig. 7a). Simultaneously, the universally conserved monitoring base G530 of the SSU shoulder disengaged from the A-site wobble pair to open the mRNA entrance channel (Extended Data Fig. 7b–d, Supplementary Video 5). This spontaneous, partial reversal of SSU domain closure was most pronounced in PRE-H1, potentially contributing to peptidyl-tRNA drop-off from PRE-H states<sup>21,27</sup>.

## EF-G initiates peptidyl-tRNA movement

In the early translocation intermediate (INT1), the tRNA-like DIV of EF-G engaged the minor groove of the peptidyl-tRNA anticodon–mRNA codon minihelix (Fig. 2a, Extended Data Fig. 8). DIV loop II wedged between the monitoring bases of SSU helix 44 (h44; A1492 and A1493) and the codon–anticodon pair to unlock peptidyl-tRNA from the SSU A site, lifting the peptidyl-tRNA–mRNA pair out of the decoding centre (Fig. 2b,c). In contrast to later translocation stages<sup>6,10,11</sup>, loop I in DIV interacted electrostatically with the peptidyl-tRNA phosphate backbone (Extended Data Fig. 8a,b), potentially aiding early EF-G association and positioning.

**Fig. 2: Unlocking of the peptidyl-tRNA decoding centre.**

 figure2



**a**, Locally filtered electron density illustrating shape-specific recognition of the A-site codon–anticodon pair by EF-G (red) in its active, GTP-bound conformation (INT1). **b**, **c**, Unlocking of the tRNA<sub>2</sub>–mRNA decoding centre in the PRE-H1 (**b**) to INT1 (**c**) transition. Peptidyl-tRNA, green; mRNA, pink; H69, grey; h44, blue; h18, cyan; uS12, light blue. Threshold  $\sigma = 5$ .

[Full size image](#)

Consistent with complete SSU unlocking at the leading edge, forward peptidyl-tRNA progression tilted the SSU head away from the body (Extended Data Fig. 4b, [Supplementary Video 5](#)), extracting SSU body base C1397 from mRNA intercalation and shifting the mRNA register relative to G530 (Extended Data Fig. 7b–d). These changes flattened the kink between the A- and P-site codons<sup>28</sup>, modestly relaxed the peptidyl-tRNA bend angle and enabled C1054 of the SSU head to pair with the +7 mRNA (Extended Data Figs. 6, 7c, [Supplementary Table 4](#)). No longer within reach of the tRNA<sub>2</sub>–mRNA module, the A1492 monitoring base and A1913 at the tip of LSU H69 inserted into h44 to relock into their POST positions (Fig. 2c). These findings rationalize how peptidyl-tRNA fixation within the decoding centre efficiently inhibits SSU unlocking and translocation<sup>5,29,30</sup>.

EF-G engagement had a limited effect outside of the decoding centre, maintaining the inter-subunit rotation angle and LSU positions of both tRNAs from PRE-H1. We did, however, observe the formation of interactions between the SSU body and the deacyl-

tRNA anticodon–mRNA codon pair (Extended Data Fig. 9), which is consistent with an allosteric securing of the reading frame in the E site.

## EF-G engages in an active conformation

Coincident with DIV-mediated peptidyl-tRNA unlocking from the decoding centre of the SSU, the G domain (DI) of EF-G packed intimately against the catalytic sarcin–ricin loop (SRL) to rigidify the GTPase-activating centre and shift it away from the SSU ([Supplementary Video 2](#)). To bridge the gap between the decoding centre and the GTPase-activating centre, EF-G adopted an elongated conformation (Fig. 3a, Extended Data Fig. 8c).

**Fig. 3: Overview of the active, GTP-bound conformation of EF-G.**

 [figure3](#)

**a**, Domain architecture of EF-G in its active, GTP-bound conformation (INT1, coloured) and in a post-hydrolysis conformation (INT2, grey, G-domain alignment). **b**, Locally filtered electron density (mesh) in the nucleotide-binding pocket for INT1. **c**, Elongated switch-I (residues 38–68, yellow) contacts with the SRL (grey), the SSU (light blue), DII (orange) and DIII (pink). The conformational change of DII is indicated with an arrow. Threshold  $\sigma = 6$ .

[Full size image](#)

Unexpectedly, the G domain contained strong, continuous electron density for  $\alpha$ ,  $\beta$  and  $\gamma$  phosphates at all thresholds, supporting the presence of a GTP molecule in the nucleotide-binding pocket (Fig. 3b, Extended Data Fig. 10). Congruent with a pre-

hydrolysis state, smFRET experiments revealed that INT1 transit was markedly slowed by the non-hydrolysable GTP analogue GTP $\gamma$ S (Extended Data Fig. 1p–t). Although we cannot unambiguously determine whether EF-G is bound to GTP, GDP-P<sub>i</sub> or a mixture of the two in dynamic exchange<sup>31</sup>, we can conclude that EF-G is capable of unlocking the peptidyl-tRNA cargo from the decoding centre of the SSU before P<sub>i</sub> release. Hence, although pre-hydrolysis EF-G conformations have been trapped on substrates that lack peptidyl-tRNA cargo or on POST complexes using non-hydrolysable GTP analogues<sup>8,9,32</sup>, or using a catalytically dead EF-G mutant (H92A)<sup>12</sup>, the INT1 structure captured here represents the best approximation to date of EF-G bound to its physiological substrate in its active, GTP-bound conformation.

Consistent with an active GTP conformation<sup>33</sup>, the switch-I and catalytic switch-II elements were fully structured to encircle the guanosine nucleotide (Fig. 3c). As observed for G domains of other GTP-bound TRAFAC-family GTPases<sup>33</sup>, the switch-I, switch-II and P-loop regions engaged the  $\beta$  and  $\gamma$  phosphates via Mg<sup>2+</sup> coordination. The catalytic switch-II residue His92 was also positioned 4 Å from the  $\gamma$  phosphate, primed to facilitate GTP hydrolysis (Extended Data Fig. 10a).

In agreement with mutation sites conferring SPC resistance<sup>34</sup>, we observed density for all three SPC rings immediately beneath the SSU P site, approximately 100 Å from the GTP-binding site<sup>35,36</sup> (Extended Data Fig. 11, Supplementary Video 6). Within its physiological INT1 substrate, the methyl substituent of SPC ring C stabilized the interaction of Lys26 of uS5 with h28—an interaction that is likely to prevent further SSU head-swivel at this specific stage of translocation<sup>13,35,36</sup>.

## EF-G engages the rotated ribosome

In its active conformation, the switch-I element of EF-G exhibited a continuous, extended architecture that bridged the G domain with DII and DIII (Fig. 3c). This region is disordered in nearly all EF-G structures both on and off the ribosome, with the exception of an isolated crystal structure of a thermophilic EF-G homologue (EF-G-2) bound to GTP<sup>37</sup> (Extended Data Fig. 8c) and structures of EF-G(H92A) bound to POST ribosomes<sup>12</sup>. The switch-I N terminus interacted with both the rotated SSU body and the LSU, anchoring His38 on the intersubunit bridge B8 fidelity determinant<sup>38</sup> and extending by approximately 19 Å to contact the SRL (Fig. 3c). Because switch-I ordering is contingent on the precise distance between these ribosomal elements, we posit that these stabilizing contacts provide the energy needed for EF-G–GTP binding to unlock the peptidyl-tRNA cargo from the SSU to initiate translocation.

Similar to the structures of GTP-bound EF-G-2<sup>37</sup> and EF-G(H92A)<sup>12</sup>, the extended switch-I structure nucleated a modified β-barrel fold in DII (Fig. 3c, Extended Data Fig. 8d), suggesting that DII has an intramolecular effector role<sup>39</sup>. This non-canonical DII architecture mediated EF-G contact with the conserved U368–A55 tertiary pair where the SSU shoulder and body domains diverge, a region that has been implicated in activating GTP hydrolysis on elongation factor Tu (EF-Tu) during tRNA selection<sup>40</sup>. The modified β-barrel fold also buttressed the switch-I C terminus against the highly conserved DIII helix B3<sup>9,12,32</sup> (Fig. 3c), providing a conduit for information transfer from the SSU shoulder–body interface to the G domain of EF-G. Because this network of contacts is specifically underpinned by interactions with the rotated SSU, we propose that the activation of GTP hydrolysis in EF-G is triggered by formation of the extended switch-I fold or by changes in the SSU rotation angle during later steps of translocation.

## P<sub>i</sub> release remodels the conformation of EF-G

By comparing INT1 with the structure of INT2 stalled by fusidic acid, we obtained additional insights into the role and timing of GTP hydrolysis by EF-G. As anticipated<sup>6,10</sup>, we observed loss of density for the nucleotide γ phosphate and switch I, and a restoration of the canonical DII β-barrel fold in the INT2 complex (Extended Data Figs. 8d, 10d). These post-hydrolysis changes correlated with an upward shift and an approximately 15° rotation of the G domain of EF-G relative to the SRL (Extended Data Fig. 8e), together with inward displacement of the entire GTPase-activating centre towards the LSU central protuberance ([Supplementary Video 2](#)). Despite such extensive remodelling, EF-G DIV loops I and III remained in direct contact with the peptidyl-tRNA anticodon–mRNA codon pair, while losing contact with the tRNA body<sup>6,9,10</sup> (Extended Data Fig. 8b). Consequently, all five EF-G domains reached further into the inter-subunit space, coupled with an approximately 17° hinge-like motion between DIV and DV roughly perpendicular to the SSU interface (Extended Data Fig. 8f, [Supplementary Video 7](#)).

As expected<sup>6,10</sup>, the altered position and conformation of EF-G in INT2 was associated with a scissor-like reverse rotation of the SSU body towards its POST position and forward SSU head-swivel in the direction of translocation (Extended Data Fig. 4b). Such changes collapsed the SPC-binding pocket (Extended Data Fig. 11c), while establishing direct contact between DIV and the SSU head domain<sup>6,9,10</sup> (Extended Data Fig. 8b,d) and a new intersubunit bridge involving the ASF, uS19 and the LSU central protuberance<sup>7,10</sup> (Extended Data Fig. 5c), potentially stabilizing the head-swivel angle. The observed scissor-like conformational changes were reduced in amplitude compared with those found in previous investigations<sup>6,10</sup>, which probably reflects the diffusive nature of SSU head- and body-domain motions and their

sensitivity to ribosome composition and/or experimental condition<sup>13,41</sup>. We infer from these observations that entrance into the INT2 basin liberates a range of intersubunit rotation and SSU head-swivel angles<sup>13</sup>—and related conformational processes in EF-G and the ribosome—that can facilitate GTP hydrolysis and/or P<sub>i</sub> release.

## Head-swivel initiates deacyl-tRNA movement

The INT1–INT2 transition moved the entire tRNA<sub>2</sub>–mRNA module by approximately 8.5 Å towards its POST position (Extended Data Fig. 4c, Supplementary Table 2), enabled in part by the maintenance of anchored stacking interactions between deacyl- and peptidyl-tRNA and the SSU head (Extended Data Fig. 12). Movement of the deacyl-tRNA anticodon triggered release of the C1400 base from the deacyl-tRNA anticodon–mRNA codon pair and disrupted E-site mRNA codon stacking with the SSU 690 loop (Extended Data Fig. 12d, e). Notably, only two of the three E-site codon nucleotides shifted relative to the G926 fiducial marker and the mRNA exit channel (Extended Data Fig. 12), establishing that the tRNA<sub>2</sub>–mRNA module is only partially translocated with respect to the SSU body.

The INT1–INT2 transition also unlocked the interface between uS7 and uS11 (Extended Data Fig. 9) and widened the gap between the L1 stalk and the SSU head at the lagging edge. Simultaneously, the peptidyl-tRNA G19–C56 elbow pair and the A2602 base of the LSU relocked into their fully translocated positions (Extended Data Fig. 5c, d). Movement of the tRNA<sub>2</sub>–mRNA module also relocked SSU bases C1397 and A1493 on the leading edge into their POST positions, intercalated on opposite sides of the downstream mRNA codon<sup>42</sup> (Extended Data Fig. 7b–d). The INT1–INT2 transition therefore completes relocking events on the LSU and in the SSU decoding centre while mediating a distinct SSU unlocking process at the lagging edge of the ribosome. Such changes are likely to contribute to reading frame maintenance while opening pathways through which deacyl-tRNA can shift position and/or dissociate<sup>13,17,42</sup>.

## Discussion

Although snapshots of translocation have been previously reported<sup>6,7,8,9,10,11</sup>, structural information on the initial engagement of GTP-bound EF-G with its physiological substrate has been missing. Our structures reveal persistent engagement of the peptidyl-tRNA cargo during the relay of tRNA<sub>2</sub>–mRNA module unlocking and relocking events on both ribosomal subunits. The sequential unlocking mechanism observed is initiated by PRE-complex dynamics. EF-G engages spontaneously achieved PRE-H conformations in its active, GTP-bound conformation, unlocking the

decoding centre and sending the peptidyl-tRNA on an arc-like trajectory in single-nucleotide increments (Fig. 4), as initially inferred from optical trapping studies of mRNA unwinding<sup>43</sup>. By contrast, deacyl-tRNA movement is not initiated during unlocking at the decoding centre but is instead coordinated by a second SSU unlocking process at the lagging edge, which enables a coupled shift of the entire tRNA<sub>2</sub>-mRNA module in the INT1–INT2 transition (Fig. 4). Notably, translocation also involves the formation of POST-like contacts in both ribosomal subunits (relocking events), which may provide a thermodynamic driving force for forward progression while securing the translation reading frame.

**Fig. 4: Non-uniform tRNA<sub>2</sub>–mRNA movement during translocation.**



**a**, Overlay of the tRNA<sub>2</sub>–mRNA module from the A (green) to P (yellow) to E (orange) sites. **b**, Overlay from **a**, viewed from the codon–anticodon interface. Circles on the tRNAs at position 34 N1 (deacyl-tRNA, left) and N3 (peptidyl-tRNA, right) depict the tRNA trajectories during translocation. **c**, tRNA anticodon–mRNA codon movement during translocation, same perspective as **b**.

[Full size image](#)

Non-competitive elongation-factor binding to the ribosome stipulates distinct recognition features. Our findings support a parsimonious model in which EF-G

preferentially engages rotated PRE-H conformations<sup>16</sup>, whereas EF-Tu recognizes the locked, unrotated ribosome. This model avoids steric clashes between EF-G and constituents of the PRE-C complex and ensures that peptide-bond formation and LSU unlocking have occurred before energy expenditure. Rotated ribosome conformations are also expected to stabilize the extended switch-I structure and the modified DII fold of EF-G, which are likely to be prerequisites for activation of GTP hydrolysis. Such a model helps to explain how PRE-H states lower the energetic barrier to translocation<sup>21</sup>; the unexpectedly high Michaelis constant of EF-G-catalysed translocation; and the dependency of translocation rate on PRE complex composition<sup>3,4,13</sup>. Structural and mechanistic conservation posit that this division of elongation-factor recognition may extend across domains of life.

Although our findings provide information on the molecular basis of early-translocation events, analogous strategies will need to be applied to late-translocation processes, during which deacyl- and peptidyl-tRNA unlock from the SSU head to progress to the POST state (Extended Data Fig. 12). Such events putatively include exaggerated swivel-like motions of the SSU head and relocking of peptidyl-tRNA in its ultimate P-site position<sup>13,17</sup>. Combined structural, smFRET and molecular-dynamics studies will also be vital in defining the precise timing of GTP hydrolysis,  $P_i$  release and the dissociation of GDP-bound EF-G. Delineation of the complete translocation mechanism will provide a deeper understanding of the regulation of translation, including the programmed errors that govern normal physiology and disease<sup>44</sup>.

## 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.

### Buffers and reagents

All experiments were carried out in either polymix buffer A (50 mM Tris-OAc (pH 7.5), 100 mM KCl, 5 mM NH<sub>4</sub>OAc, 0.5 mM Ca(OAc)<sub>2</sub>, 5 mM Mg(OAc)<sub>2</sub>, 6 mM 2-mercaptoethanol, 0.1 mM EDTA, 5 mM putrescine and 1 mM spermidine)<sup>45</sup> or polymix buffer B (30 mM HEPES pH 7.5, 5 mM MgCl<sub>2</sub>, 50 mM NH<sub>4</sub>Cl, 5 mM 2-mercaptoethanol, 2 mM spermidine and 5 mM putrescine)<sup>46</sup>. A cocktail of triplet-state quenchers (1 mM Trolox, 1 mM nitrobenzyl alcohol and 1 mM cyclooctatetraene) and

an enzymatic oxygen scavenging system (protocatechuic acid (PCA)/protocatechuate-3,4-dioxygenase (PCD)) were used for smFRET experiments. Spectinomycin sulfate was purchased from MP Biomedicals. Fusidic acid sodium salt, GTP and GTP $\gamma$ S were from Sigma-Aldrich. GTP was further purified using a Mono Q 5/50 GL anion exchange column (GE Healthcare Life Sciences). Pyruvate kinase, myokinase and phosphoenolpyruvate (PEP) were purchased from Sigma-Aldrich. All other standard reagents were purchased from Sigma-Aldrich or VWR.

## Cryo-EM and smFRET sample preparation

### Purification of ribosomes and elongation factors

Wild-type, uS13- and uL1-labelled ribosomal subunits were purified from *Escherichia coli* BL21 and MRE600 for smFRET and cryo-EM experiments, respectively, as previously described<sup>13,45,47</sup>. EF-Tu<sup>48</sup> and EF-G<sup>15</sup> were purified as previously described. *E. coli* tRNA<sup>fMet</sup>, tRNA<sup>Phe</sup> and tRNA<sup>Lys</sup> were purified<sup>13,16,24</sup> and tRNA<sup>Phe</sup> was labelled with LD655 at the acp<sup>3</sup> modification on nucleotide U47, as described previously<sup>45</sup>. Wild-type, uS13- and uL1-labelled initiation complexes were prepared as previously described<sup>45,47,49</sup>.

## **Preparation of ternary complex for smFRET experiments**

Phenylalanine (2.5 mM), PheRS (0.15  $\mu$ M), pyruvate kinase (0.4  $\mu$ M), myokinase (0.5  $\mu$ M), PEP (3.75 mM), GTP (630  $\mu$ M) and LD655-labelled tRNA<sup>Phe</sup> (250 nM) were combined in charging buffer (50 mM Tris pH 8, 10 mM KCl, 100 mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub>, 1 mM DTT, 2.5 mM ATP and 0.5 mM EDTA) before addition of EF-Tu–EF-Ts (EF-Ts, elongation factor thermostable) (1  $\mu$ M). The resulting mixture was incubated for 10 min at 37 °C to aminoacylate the tRNA (aa-tRNA) and form a ternary complex (EF-Tu–aa-tRNA–GTP). Before injection into the microscope flow cell for smFRET imaging, ternary complex was diluted 40× (to a final concentration of 6 nM) in imaging polymix buffer containing 0.5 mM GTP.

## **Preparation of Phe-tRNA<sup>Phe</sup> ternary complex for cryo-EM experiments**

Phenylalanine (1 mM), PheRS (0.2  $\mu$ M), pyruvate kinase (0.6  $\mu$ M), myokinase (0.6  $\mu$ M), PEP (0.4 mM), GTP (1 mM) and tRNA<sup>Phe</sup> (1.6  $\mu$ M) were combined in charging buffer before addition of EF-Tu–EF-Ts (8  $\mu$ M). The resulting mixture was incubated for 10 min at 37 °C to aminoacylate the tRNA and form a ternary complex. Successful aminoacylation was confirmed by fast protein liquid chromatography (FPLC).

## **Preparation of Lys-tRNA<sup>Lys</sup> ternary complex for cryo-EM experiments**

Lysine (1 mM), LysRS (0.6  $\mu$ M), pyruvate kinase (0.6  $\mu$ M), myokinase (0.6  $\mu$ M), PEP (0.4 mM), GTP (1 mM) and tRNA<sup>Lys</sup> (3  $\mu$ M) were combined in charging buffer and incubated for 15 min at 37 °C to aminoacylate the tRNA. EF-Tu–EF-Ts (15  $\mu$ M) was added to the mixture and incubated for 5 min at 37 °C to form a ternary complex. Successful aminoacylation was confirmed by FPLC.

## **Preparation of elongator POST complexes for cryo-EM**

All reactions were performed in the presence of a GTP regeneration system<sup>50</sup>. Initiation complexes at a concentration of approximately 3  $\mu$ M were prepared with MFK mRNA (Biotin-5'-CAA CCU AAA ACU UAC ACA CCC UUA GAG GGA CAA UCG **AUG UUC AAA GUC UUC AAA GUC AUC-3'**) and fMet-tRNA<sup>fMet</sup> in the P site. mRNA nucleotide position 40 corresponds to the +1 position. Initiation complexes were incubated with ternary complex containing Phe-tRNA<sup>Phe</sup> (around 1.6  $\mu$ M) for 5 min at 37 °C to form the PRE translocation complex. PRE complexes were incubated with sub-stoichiometric concentrations of GTP-bound EF-G (300 nM) for 10 min at 37 °C to form the elongator POST translocation complex (fMet-Phe-tRNA<sup>Phe</sup> in the P site). Additional reagents were added to the mixture to aminoacylate free tRNA in solution. Elongator POST complexes were pelleted over a 37% sucrose cushion containing buffer A at 437,000g in a TLA-100.3 rotor (Beckman) for 4 h at 4 °C to remove EF-G and deacetyl-tRNA. Pelleted elongator complexes were resuspended in buffer A for a final concentration of 9  $\mu$ M and were flash-frozen.

### **Preparation of elongator complexes for cryo-EM**

Elongator POST complexes containing fMet-Phe-tRNA<sup>Phe</sup> in the P site were thawed and diluted in buffer B with 1 mM GTP for a final concentration of 2  $\mu$ M ribosomes. For preparation of the translocation intermediate samples, SPC (INT1) or fusidic acid (INT2) were added to the dilution buffer. SPC was used at its half-maximum inhibitory concentration ( $IC_{50}$ ) for translocation inhibition (3 mM)<sup>51</sup>. Fusidic acid was used at near-saturating concentration (400  $\mu$ M)<sup>13,52</sup>. The elongator POST complexes were incubated with Lys-tRNA<sup>Lys</sup> ternary complex (2  $\mu$ M final) for around 30 s at 25 °C to fill the A site. The resulting elongator PRE complex (P-site tRNA<sup>Phe</sup>; A-site fMet-Phe-Lys-tRNA<sup>Lys</sup>) was either added to cryo-EM grids directly (PRE) or incubated with EF-G (5  $\mu$ M final) in the absence (POST) or presence of SPC (INT1) or fusidic acid (INT2) for around 5–10 s before the solution was applied to cryo-EM grids.

### **Cryo-EM grid preparation**

Cryo-EM grids were prepared using a Vitrobot Mark IV plunge-freezing device (Thermo Fisher Scientific). For each experiment, 3 µl of sample was applied to Quantifoil R 1.2/1.3 holey carbon Cu 300 mesh (INT1 and INT2) or Au 300 mesh (PRE and POST) grids that had been glow-discharged ( $\text{Ar}/\text{O}_2$ ) for 20 s using a Solarus II Plasma Cleaning system (Gatan). Grids were incubated in the Vitrobot chamber for 10 s at 10 °C at 95% humidity before blotting (6 s; blot force –5) and plunge freezing into liquid ethane.

## smFRET imaging of translocation

Ribosomes programmed with 5'-biotinylated mRNA substrates containing P-site-bound fMet-tRNA<sup>fMet</sup> and displaying the codon UUC in the A site were immobilized on passivated coverslips as described previously<sup>13,45</sup>. The ribosomes were then incubated for 2 min with ternary complex containing either LD655-labelled Phe-tRNA<sup>Phe</sup> or unlabelled Phe-tRNA<sup>Phe</sup>, leading to stoichiometric formation of either PRE ribosomes containing A-site LD655-labelled fMet-Phe-tRNA<sup>Phe</sup>, P-site tRNA<sup>fMet</sup> and LD550-labelled uS13 or PRE ribosomes containing A-site fMet-Phe-tRNA<sup>Phe</sup>, P-site tRNA<sup>fMet</sup> and LD550-labelled uS13 and LD650-labelled uL1. To initiate translocation, EF-G with either 1 mM GTP or 1 mM GTPγS, with or without 3 mM SPC or 400 µM fusidic acid, was delivered to the flow cell by stopped-flow injection. All smFRET experiments were carried out at 25 °C. The time-evolution of the FRET signal was then recorded using a home-built total-internal-reflection-based fluorescence microscope<sup>53</sup> with laser (532 nm) illumination at 0.1 kW cm<sup>-2</sup> at a time resolution of 40 or 400 ms. Donor and acceptor fluorescence intensities were extracted from the recorded movies and FRET efficiency traces were calculated using custom software implemented in MATLAB R2015b. FRET traces were selected for further analysis according to the following criteria: a single catastrophic photobleaching event; at least 8:1 signal-to-background-noise ratio and 6:1 signal-to-signal/noise ratio; less than four donor–fluorophore blinking events; a correlation coefficient between donor and acceptor <0.5. The resulting smFRET traces were analysed using hidden Markov model idealization methods as implemented in the SPARTAN software package (v.3.7.0)<sup>53</sup>. In all idealizations, transitions between all states were allowed. The model used for uS13 to peptidyl-tRNA FRET had four states (FRET

values:  $0.14 \pm 0.04$ ;  $0.30 \pm 0.03$ ;  $0.50 \pm 0.05$ ;  $0.75 \pm 0.06$ ); the model for uS13 to uL1 FRET had three FRET states (FRET values:  $0.73 \pm 0.05$ ;  $0.48 \pm 0.08$ ;  $0.27 \pm 0.04$ ). To compare the translocation kinetics under different conditions from the idealized FRET traces, we constructed normalized cumulative distributions over the arrival time to the POST state, defined as the 0.50 FRET state for the uS13 to peptidyl-tRNA signal and the 0.27 FRET state for the uS13 to uL1 signal.

## Cryo-EM data collection

Cryo-EM data were collected using a Titan Krios G3i (Thermo Fisher Scientific) transmission electron microscope equipped with a K3 direct electron detector and post column GIF (energy filter). K3 gain references were acquired just before data collection. Data collection was performed using SerialEM software (v.3.7.1)<sup>54</sup> with image shift protocol (nine images were collected with one defocus measurement per nine holes). Movies were recorded at defocus values from  $-0.5 \mu\text{m}$  to  $-1.5 \mu\text{m}$  at a magnification of 105,000 $\times$ , which corresponds to the pixel size of  $0.826 \text{ \AA}$  per pixel at the specimen level (super-resolution  $0.413 \text{ \AA}$  per pixel) for the apo PRE, POST and INT1 structures. During the 2.4-s exposure, 60 frames (0.04 s per frame,  $1.4596 e^-$  per frame per  $\text{\AA}^2$ ) were collected with a total dose of around  $87 e^-$  per  $\text{\AA}^2$ . The first frame was discarded. Motion correction was performed on raw super-resolution movie stacks and binned twofold using MotionCor2 software<sup>55</sup>. Cryo-EM data for the INT2 complex was collected at a magnification of 82 kx ( $1.06 \text{ \AA}$  per pixel; super-resolution  $0.53 \text{ \AA}$  per pixel), with a total dose of around  $70 e^-$  per  $\text{\AA}^2$ . CTF parameters were determined using CTFFind4<sup>56</sup> and refined later in Relion<sup>57</sup> (v.3.1) and cryoSPARC<sup>58</sup> (v.3). Before particle picking, good micrographs were qualified by power spectrum. Particles were picked using cisTEM<sup>59</sup> and the coordinates were transferred to Relion (see below for details of classification and refinement). Sharpened and locally filtered maps were used to aid in model building. Electron density map values were normalized to mean = 0 and standard deviation ( $\sigma$ ) = 1 in UCSF Chimera using the vop scale function. For detailed information on data collection parameters and model-building statistics see Extended Data Fig. 2 and Supplementary Table 1.

## Cryo-EM data processing for the apo PRE structures

Prior to particle picking, good micrographs were qualified by power spectrum (7,183 movie stacks). Particles were picked within cisTEM (659,777 particles). After extraction in Relion (fourfold binned), several rounds of the 2D classification were performed in cryoSPARC. An Ab initio structure was built in cryoSPARC and then used as a reference for 3D classification in Relion. Particles from good classes (534,348 particles) were then re-extracted (twofold binned) and refined in Relion followed by CtfRefine and 3D classification into 10 classes. Class 4 possessed an unrotated SSU (225,423 particles) and class 9 possessed a rotated SSU (160, 291 particles). Unrotated SSU and rotated SSU classes were individually subjected to 3D refinement in Relion and sorted further into 5 classes using 3D classification. From 3D classification of the unrotated particles, two classes contained classical A- and P-site tRNAs, which were combined (109,769 particles) and run through 3D refinement (un-binned) yielding the PRE-C structure. From 3D classification of the rotated particles, one class showed evidence of hybrid P-site tRNA and classical A-site tRNA (PRE-H2\*; 33,330 particles). Two classes from the unrotated 3D classification contained weak density for A-site tRNA, which were combined (126,699 particles) and further classified with an A-site mask to improve ligand density. From this focused classification of the A-site tRNA, one class contained hybrid deacyl-tRNA and peptidyl-tRNA (PRE-H1; 51,685 particles). After 3D classification, particles from PRE-C, PRE-H2\* and PRE-H1 classes were re-extracted with the full pixel size and refined in Relion according to the gold-standard criteria.

## Cryo-EM data processing for the apo POST structure

Before particle picking, good micrographs were qualified by power spectrum (2,834 movie stacks). Particles were picked within cisTEM (439,422 particles). After extraction in Relion (fourfold-binned), particles were refined in Relion followed by CtfRefine and 3D classification into 10 classes. Five classes possessed an unrotated SSU, which were combined (120,513 particles), re-extracted and refined in Relion (twofold-binned). To improve the density for E-site tRNA, we performed a focused 3D classification using an E-site mask. This yielded two classes with solid E-

site tRNA density, which were combined (34,688 total particles), re-extracted with the full pixel size and refined in Relion according to the gold-standard criteria for the apo POST complex structure.

## Cryo-EM data collection and processing for the INT1 structure

Before particle picking, good micrographs were qualified by power spectrum (11,916 movie stacks). Particles were picked within cisTEM (1,001,439 particles). After extraction in Relion (fourfold-binned), several rounds of the 2D classification were performed in cryoSPARC. Particles from good 2D classes (652,128 particles) were then refined in Relion and sorted using 3D classification into 6 classes. One class possessed a rotated SSU (184,857 particles), which was refined and further classified in Relion into 8 classes. One of these classes contained EF-G (33,688 particles). These particles were re-extracted with the full pixel size and refined in Relion according to the gold-standard criteria for the INT1 complex structure.

## Cryo-EM data collection and processing for the INT2 structure

Before particle picking, good micrographs were qualified by power spectrum (6,651 movie stacks). Particles were picked within cisTEM (1,259,307 particles). After extraction in Relion (fourfold-binned), several rounds of 2D classification were performed in cryoSPARC. Particles from good 2D classes (639,984 particles) were then refined in Relion and sorted using 3D classification into 6 classes. One class possessed a rotated SSU with EF-G bound (113,540 particles), which was refined in Relion. To improve occupancy of the tRNAs and EF-G, we performed a focused 3D classification using a ligand mask into three classes in Relion. One class contained two tRNAs and EF-G (33,008 particles). These particles were re-extracted with the full pixel size and refined in Relion according to the gold-standard criteria for the INT2 complex structure.

## Molecular model building

Models of 50S (starting model PDB ID: 4YBB<sup>60</sup>), 30S (starting model PDB ID: 4YBB<sup>60</sup>), tRNA<sup>Lys</sup> (starting model PDB ID: 5E81<sup>61</sup>), tRNA<sup>Phe</sup> (starting

model PDB ID: 4WRO<sup>62</sup>), EF-G (starting model PDB ID: 4V9O<sup>32</sup>) and ribosomal protein L7/L12 (starting model PDB ID: 1CTF<sup>63</sup>) were fitted into EM maps and refined through iterative rounds of manual model building in Coot (v.0.9.4.1)<sup>64</sup>, refinement of RNA with ERRASER<sup>65</sup> and real-space refinement using Phenix (v.1.19-4092)<sup>66</sup>. mRNA nucleotide 40 corresponds to the +1 position. The nascent peptide and mRNA were built de novo in Coot. ATP molecules were modelled between 23S (1) U369 and A404 and (2) U40 and U441. An ATP was also modelled in the LSU E site for PRE-C. Polyamines were modelled into tubular unassigned density displaying the appropriate surrounding electrochemical environment. Notably, putrescine molecules were modelled in the E site of PRE-H2\* and PRE-H1 proximal to 16S A790. The acp<sup>3</sup> modification on U47 of tRNA<sup>Phe</sup> and tRNA<sup>Lys</sup> was also modelled de novo as follows: the 3-amino-3-carboxypropyl moiety was added to position 3 of the pyrimidine ring of uridine monophosphate, saved as a novel modified RNA-nucleotide acp<sup>3</sup>U with ligand code 3au. Restraints for refinement were generated using phenix.elbow<sup>67</sup>. Models were validated using phenix.validation\_cryoem<sup>68</sup> with the built-in MolProbity<sup>69</sup> scoring. See Supplementary Table 1 for more information. In each complex, we also observed fragmented electron density for the Shine–Dalgarno-like/Anti-Shine–Dalgarno minihelix, which was modelled in PRE-C, as well as ribosomal protein uS1.

## Figure preparation

Molecular graphics and analyses were performed with UCSF Chimera<sup>70</sup> or ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. Unsharpened maps from Relion Refine3D were used for figure images, threshold  $\sigma = 6$ , unless otherwise stated. Angle and distance measurements were performed in UCSF Chimera using the Fit in Map and the Distance tools. All figures were prepared using structures and models aligned on the LSU core, unless otherwise noted. The LSU core used was simulated 3 Å density of high-resolution ribosome crystal structure PDB ID: 4YBB<sup>60</sup> in UCSF Chimera (molmap) with the following mobile elements omitted: uL5, uL6, uL9, uL10, uL11, uL120, uL31, H34 (709–723), A-site finger (ASF; H38; 866–906), the L11 stalk (1045–1112),

H69 (1908–1925), the L1 stalk (2093–2198), H83/84 (2297–2318), the SRL (2651–2667) and 5S. Root mean square deviation (r.m.s.d.) heat maps were prepared in UCSF Chimera using the Matchmaker tool for proteins and nucleic acids. Rotation angles and axes illustrations were generated using the Measure Rotation tool in UCSF Chimera. Electron density was coloured using the Colour Zone tool with a 3 Å radius. Figures were compiled in Adobe Illustrator (Adobe). The mRNA kink in Supplementary Table 2 was measured between mRNA positions +3 (C1' of the last nucleotide of the deacyl-tRNA codon) and +4 (C1' of the first nucleotide of the peptidyl-tRNA codon).

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

PDBs and cryo-EM 3D maps for all structures are available through the Protein Data Bank (PDB) and the Electron Microscopy Data Bank (EMDB), respectively, as follows: PRE-C, [7N1P](#), [EMD-24120](#); PRE-H2\*, [7N30](#), [EMD-24135](#); PRE-H1, [7N2U](#), [EMD-24133](#); INT1-SPC, [7N2V](#), [EMD-24134](#); INT2-FA, [7N2C](#), [EMD-24132](#); POST, [7N31](#), [EMD-24136](#).

## References

1. 1.

Kurland, C. G. Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* **26**, 29–50 (1992).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

2. 2.

Rodnina, M. V., Savelsbergh, A., Katunin, V. I. & Wintermeyer, W. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* **385**, 37–41 (1997).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

3. 3.

Savelsbergh, A. et al. An elongation factor G-induced ribosome rearrangement precedes tRNA-mRNA translocation. *Mol. Cell* **11**, 1517–1523 (2003).

[CAS](#) [PubMed](#) [Article](#) [PubMed Central](#) [Google Scholar](#)

4. 4.

Borg, A. & Ehrenberg, M. Determinants of the rate of mRNA translocation in bacterial protein synthesis. *J. Mol. Biol.* **427**, 1835–1847 (2015).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

5. 5.

Brilot, A. F., Korostelev, A. A., Ermolenko, D. N. & Grigorieff, N. Structure of the ribosome with elongation factor G trapped in the pretranslocation state. *Proc. Natl Acad. Sci. USA* **110**, 20994–20999 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

6. 6.

Zhou, J., Lancaster, L., Donohue, J. P. & Noller, H. F. How the ribosome hands the A-site tRNA to the P site during EF-G-catalyzed translocation. *Science* **345**, 1188–1191 (2014).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

7. 7.

Ratje, A. H. et al. Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. *Nature* **468**, 713–716 (2010).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

8. 8.

Tourigny, D. S., Fernández, I. S., Kelley, A. C. & Ramakrishnan, V. Elongation factor G bound to the ribosome in an intermediate state of translocation. *Science* **340**, 1235490 (2013).

[PubMed](#) [Article](#) [CAS](#) [Google Scholar](#)

9. 9.

Zhou, J., Lancaster, L., Donohue, J. P. & Noller, H. F. Crystal structures of EF-G–ribosome complexes trapped in intermediate states of translocation. *Science* **340**, 1236086 (2013).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

10. 10.

Ramrath, D. J. F. et al. Visualization of two transfer RNAs trapped in transit during elongation factor G-mediated translocation. *Proc. Natl Acad. Sci. USA* **110**, 20964–20969 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

11. 11.

Gao, Y.-G. et al. The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* **326**, 694–699 (2009).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

12. 12.

Li, W. et al. Activation of GTP hydrolysis in mRNA–tRNA translocation by elongation factor G. *Sci. Adv.* **1**, e1500169 (2015).

[ADS](#) [PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

13. 13.

Wasserman, M. R., Alejo, J. L., Altman, R. B. & Blanchard, S. C. Multiperspective smFRET reveals rate-determining late intermediates of ribosomal translocation. *Nat. Struct. Mol. Biol.* **23**, 333–341 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

14. 14.

Chen, J., Petrov, A., Tsai, A., O’Leary, S. E. & Puglisi, J. D. Coordinated conformational and compositional dynamics drive ribosome translocation. *Nat. Struct. Mol. Biol.* **20**, 718–727 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

15. 15.

Munro, J. B., Altman, R. B., O’Connor, N. & Blanchard, S. C. Identification of two distinct hybrid state intermediates on the ribosome. *Mol. Cell* **25**, 505–517 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

16. 16.

Wang, L., Altman, R. B. & Blanchard, S. C. Insights into the molecular determinants of EF-G catalyzed translocation. *RNA* **17**, 2189–2200 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

17. 17.

Alejo, J. L. & Blanchard, S. C. Miscoding-induced stalling of substrate translocation on the bacterial ribosome. *Proc. Natl Acad. Sci. USA* **114**, E8603–E8610 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

18. 18.

Valle, M. et al. Locking and unlocking of ribosomal motions. *Cell* **114**, 123–134 (2003).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

19. 19.

Moazed, D. & Noller, H. F. Intermediate states in the movement of transfer RNA in the ribosome. *Nature* **342**, 142–148 (1989).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

20. 20.

Cornish, P. V., Ermolenko, D. N., Noller, H. F. & Ha, T. Spontaneous intersubunit rotation in single ribosomes. *Mol. Cell* **30**, 578–588 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

21. 21.

Semenkov, Y. P., Rodnina, M. V. & Wintermeyer, W. Energetic contribution of tRNA hybrid state formation to translocation catalysis on the ribosome. *Nat. Struct. Biol.* **7**, 1027–1031 (2000).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

22. 22.

Ogle, J. M. et al. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* **292**, 897–902 (2001).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

23. 23.

Ogle, J. M., Murphy, F. V., Tarry, M. J. & Ramakrishnan, V. Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**, 721–732 (2002).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

24. 24.

Dunkle, J. A. et al. Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science* **332**, 981–984 (2011).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

25. 25.

Budkevich, T. et al. Structure and dynamics of the mammalian ribosomal pretranslocation complex. *Mol. Cell* **44**, 214–224 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

26. 26.

Sharma, D., Southworth, D. R. & Green, R. EF-G-independent reactivity of a pre-translocation-state ribosome complex with the aminoacyl tRNA substrate puromycin supports an intermediate (hybrid) state of tRNA binding. *RNA* **10**, 102–113 (2004).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

27. 27.

Verma, M. et al. A short translational ramp determines the efficiency of protein synthesis. *Nat. Commun.* **10**, 5774 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

28. 28.

Yusupov, M. M. et al. Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–896 (2001).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

29. 29.

Holm, M., Borg, A., Ehrenberg, M. & Sanyal, S. Molecular mechanism of viomycin inhibition of peptide elongation in bacteria. *Proc. Natl Acad. Sci. USA* **113**, 978–983 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

30. 30.

Wang, L. et al. Allosteric control of the ribosome by small-molecule antibiotics. *Nat. Struct. Mol. Biol.* **19**, 957–963 (2012).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

31. 31.

Algire, M. A., Maag, D. & Lorsch, J. R. P<sub>i</sub> release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol. Cell* **20**, 251–262 (2005).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

32. 32.

Pulk, A. & Cate, J. H. D. Control of ribosomal subunit rotation by elongation factor G. *Science* **340**, 1235970 (2013).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

33. 33.

Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299–1304 (2001).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

34. 34.

Lee, R. E. et al. Spectinamides: a new class of semisynthetic antituberculosis agents that overcome native drug efflux. *Nat. Med.* **20**, 152–158 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

35. 35.

Borovinskaya, M. A., Shoji, S., Holton, J. M., Fredrick, K. & Cate, J. H. D. A steric block in translation caused by the antibiotic spectinomycin. *ACS Chem. Biol.* **2**, 545–552 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

36. 36.

Carter, A. P. et al. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**, 340–348 (2000).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

37. 37.

Connell, S. R. et al. Structural basis for interaction of the ribosome with the switch regions of GTP-bound elongation factors. *Mol. Cell* **25**, 751–764 (2007).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

38. 38.

McClory, S. P., Leisring, J. M., Qin, D. & Fredrick, K. Missense suppressor mutations in 16S rRNA reveal the importance of helices h8 and h14 in aminoacyl-tRNA selection. *RNA* **16**, 1925–1934 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

39. 39.

Gasper, R. & Wittinghofer, F. The Ras switch in structural and historical perspective. *Biol. Chem.* **401**, 143–163 (2019).

[PubMed](#) [Article](#) [CAS](#) [Google Scholar](#)

40. 40.

Schmeing, T. M. et al. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688–694 (2009).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

41. 41.

Whitford, P. C., Blanchard, S. C., Cate, J. H. D. & Sanbonmatsu, K. Y. Connecting the kinetics and energy landscape of tRNA translocation on the ribosome. *PLoS Comput. Biol.* **9**, e1003003 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

42. 42.

Flis, J. et al. tRNA translocation by the eukaryotic 80S ribosome and the impact of GTP hydrolysis. *Cell Rep.* **25**, 2676–2688.e7 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

43. 43.

Wen, J.-D. et al. Following translation by single ribosomes one codon at a time. *Nature* **452**, 598–603 (2008).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

44. 44.

Dever, T. E., Dinman, J. D. & Green, R. Translation elongation and recoding in eukaryotes. *Cold Spring Harb. Perspect. Biol.* **10**, a032649 (2018).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

45. 45.

Blanchard, S. C., Kim, H. D., Gonzalez, R. L., Jr, Puglisi, J. D. & Chu, S. tRNA dynamics on the ribosome during translation. *Proc. Natl Acad. Sci. USA* **101**, 12893–12898 (2004).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

46. 46.

Ferguson, A. et al. Functional dynamics within the human ribosome regulate the rate of active protein synthesis. *Mol. Cell* **60**, 475–486 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

47. 47.

Blanchard, S. C., Gonzalez, R. L., Kim, H. D., Chu, S. & Puglisi, J. D. tRNA selection and kinetic proofreading in translation. *Nat. Struct. Mol. Biol.* **11**, 1008–1014 (2004).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

48. 48.

Burnett, B. J. et al. Direct evidence of an elongation factor-Tu/Ts-GTP aminoacyl-tRNA quaternary complex. *J. Biol. Chem.* **289**, 23917–23927 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

49. 49.

Wasserman, M. R. et al. Chemically related 4,5-linked aminoglycoside antibiotics drive subunit rotation in opposite directions. *Nat. Commun.* **6**, 7896 (2015).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

50. 50.

Jelenc, P. C. & Kurland, C. G. Nucleoside triphosphate regeneration decreases the frequency of translation errors. *Proc. Natl Acad. Sci. USA* **76**, 3174–3178 (1979).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

51. 51.

Bruhn, D. F. et al. Aminomethyl spectinomycins as therapeutics for drug-resistant respiratory tract and sexually transmitted bacterial infections. *Sci. Transl. Med.* **7**, 288ra75 (2015).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

52. 52.

Borg, A. et al. Fusidic acid targets elongation factor G in several stages of translocation on the bacterial ribosome. *J. Biol. Chem.* **290**, 3440–3454 (2015).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

53. 53.

Juette, M. F. et al. Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat. Methods* **13**, 341–344 (2016).

[PubMed](#) [PubMed Central](#) [Article](#) [CAS](#) [Google Scholar](#)

54. 54.

Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

[Article](#) [Google Scholar](#)

55. 55.

Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

56. 56.

Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

[PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

57. 57.

Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

[PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

58. 58.

Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

59. 59.

Grant, T., Rohou, A. & Grigorieff, N. *cisTEM*, user-friendly software for single-particle image processing. *eLife* **7**, e35383 (2018).

[PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

60. 60.

Noeske, J. et al. High-resolution structure of the *Escherichia coli* ribosome. *Nat. Struct. Mol. Biol.* **22**, 336–341 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

61. 61.

Rozov, A. et al. Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat. Commun.* **7**, 10457 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

62. 62.

Rozov, A., Demeshkina, N., Westhof, E., Yusupov, M. & Yusupova, G. Structural insights into the translational infidelity mechanism. *Nat. Commun.* **6**, 7251 (2015).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

63. 63.

Leijonmarck, M. & Liljas, A. Structure of the C-terminal domain of the ribosomal protein L7L12 from *Escherichia coli* at 1.7 Å. *J. Mol. Biol.* **195**, 555–579 (1987).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

64. 64.

Casañal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of electron cryo-microscopy and crystallographic data. *Protein Sci.* **29**, 1055–1064 (2020).

[Article](#) [CAS](#) [Google Scholar](#)

65. 65.

Chou, F.-C., Sripakdeevong, P., Dibrov, S. M., Hermann, T. & Das, R. Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat. Methods* **10**, 74–76 (2013).

[CAS](#) [PubMed](#) [Article](#) [Google Scholar](#)

66. 66.

Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

67. 67.

Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D* **65**, 1074–1080 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

68. 68.

Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D* **74**, 814–840 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

69. 69.

Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

70. 70.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

71. 71.

Chen, C. et al. Single-molecule fluorescence measurements of ribosomal translocation dynamics. *Mol. Cell* **42**, 367–377 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

72. 72.

Munro, J. B., Wasserman, M. R., Altman, R. B., Wang, L. & Blanchard, S. C. Correlated conformational events in EF-G and the ribosome regulate translocation. *Nat. Struct. Mol. Biol.* **17**, 1470–1477 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

73. 73.

Chen, C. et al. Elongation factor G initiates translocation through a power stroke. *Proc. Natl Acad. Sci. USA* **113**, 7515–7520 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Article](#) [Google Scholar](#)

[Download references](#)

## Acknowledgements

We acknowledge support from the Cryo-Electron Microscopy and Tomography Center, the High-Performance Computing Center and the Single-Molecule Center at St. Jude Children's Research Hospital. We also acknowledge guidance from D. Miller and J. Bollinger from the St. Jude Biomolecular X-ray Crystallography Center for their input and support of our molecular-model building and computational infrastructure. We thank D. Terry, R. Kiselev and other members of the Blanchard laboratory for their expertise and efforts to enable the single-molecule investigations performed and for their review of the manuscript during the preparation and completion of this research. This work was supported by grant funding from the National Institutes of Health (GM079238) to S.C.B., the NIH T32 (GM115327-Tan) to E.J.R., the Swedish Research Council (2017-06313) to M.H. and the German Research Foundation (DFG FOR 1805) to C.M.T.S.

# Author information

## Affiliations

1. Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN, USA

Emily J. Rundlet, Mikael Holm, S. Kundhavai Natchiar, Roger B. Altman, Alexander G. Myasnikov & Scott C. Blanchard

2. Tri-Institutional PhD Program in Chemical Biology, Weill Cornell Medicine, New York, NY, USA

Emily J. Rundlet

3. Institut für Medizinische Physik und Biophysik, Charité – Universitätsmedizin Berlin, Berlin, Germany

Magdalena Schacherl & Christian M. T. Spahn

## Authors

1. Emily J. Rundlet

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Mikael Holm

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Magdalena Schacherl

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. S. Kundhavai Natchiar  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Roger B. Altman  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Christian M. T. Spahn  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Alexander G. Myasnikov  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Scott C. Blanchard  
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

E.J.R. and R.B.A. prepared the biochemical components used for this investigation. A.G.M. and E.J.R. prepared cryo-EM grids, and collected and processed cryo-EM data. E.J.R., M.S. and S.K.N. built and refined molecular models. M.H. performed smFRET experiments. E.J.R., S.C.B., M.S. and C.M.T.S. analysed structural data. All authors wrote the manuscript.

## Corresponding author

Correspondence to [Scott C. Blanchard](#).

# Ethics declarations

## Competing interests

S.C.B. and R.B.A. hold equity interests in Lumidyne Technologies. The remaining authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks Joseph Puglisi 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 smFRET investigations of translocation.

**a–g**, smFRET data on translocation inhibition by SPC. **a**, Schematic of translocation indicating the positions of the donor (LD550, uS13 N terminus) and acceptor (LD655, peptidyl-tRNA U47) fluorophores, and FRET efficiency values for the indicated states. **b–e**, Population FRET histograms showing time evolution of FRET between uS13 and peptidyl-tRNA at a time resolution of 400 ms on delivery of 5  $\mu$ M EF-G in either buffer A (**b**, **d**) or B (**c**, **e**), without (**b**, **c**) or with 3 mM SPC (**d**, **e**). *N* indicates the number of observed molecules. **f**, Example smFRET trace from the data in **b–e**. EF-G, injected approximately 4 s after the start of data acquisition, rapidly binds to the PRE complex (0.14 FRET), converting it into INT1, and subsequently INT2 (both 0.3 FRET), before achieving the POST state (0.5 FRET). **g**, Kinetic analysis of data as in **b–e**, demonstrating an approximately tenfold potentiating effect of buffer B on SPC inhibition. The points represent mean accumulation of translocated ribosomes with time, defined as reaching the 0.5 FRET state. Solid lines represent fits of bi-

exponential functions to the data. Error bars indicate standard deviations derived from three technical experimental replicates. **h–o**, Approximate positions of published donor (green sphere) and acceptor (red sphere) dyes on the INT1 structure. **h**, Donor tRNA<sup>Phe</sup> (Cy3, s<sup>4</sup>U8), acceptor fMet-Phe-Lys-tRNA<sup>Lys</sup> (Cy5, acp<sup>3</sup>U47)<sup>13,15</sup>. **i**, Donor tRNA<sup>Phe</sup> (Cy3, acp<sup>3</sup>U47), acceptor fMet-Phe-Lys-tRNA<sup>Lys</sup> (Cy5, acp<sup>3</sup>U47)<sup>13</sup>. **j**, Donor uS13 (LD550, N-terminal ACP), acceptor fMet-Phe-Lys-tRNA<sup>Lys</sup> (Cy5, acp<sup>3</sup>U47)<sup>13</sup>. **k**, Donor uL11 (Cy3, residue 87), acceptor fMet-Phe-Lys-tRNA<sup>Lys</sup> (Cy5, acp<sup>3</sup>U47)<sup>71</sup>. **l**, Donor uS13 (LD550, N-terminal ACP), acceptor uL5 (LD650, N-terminal ACP)<sup>13</sup>. **m**, Donor bL9 (Cy3, residue N11C), acceptor bS6 (Cy5, residue D41C)<sup>20</sup>. **n**, Donor fMet-Phe-Lys-tRNA<sup>Lys</sup> (Cy3, acp<sup>3</sup>U47), acceptor EF-G (Cy5, C-terminal SFP)<sup>72</sup>. **o**, Donor EF-G (bifunctional rhodamine, residues 467–474), acceptor uL11 (Cy5, residue 87)<sup>73</sup>. **p**, Schematic of translocation indicating the positions of donor (LD550, uS13 N terminus) and acceptor (LD650, uL1) fluorophores and FRET efficiency values for the indicated states in **q–t**. **q–t**, Population FRET histograms showing time evolution of FRET between uS13 and uL1 upon injection of 5 μM EF-G with either 1 mM GTP (**q**, Apo), 1 mM GTP and 3 mM SPC (**r**), 1 mM GTPγS (**s**) or 1 mM GTP and 400 μM fusidic acid (**t**, FA), revealing that the non-hydrolysable GTP analogue GTPγS stalls the ribosome in both the same states as SPC (INT1) and fusidic acid (INT2), whereas SPC stalls only in INT1 and fusidic acid stalls only in INT2. *N* indicates the number of observed molecules.

### Extended Data Fig. 2 Fourier shell correlation and local resolution for cryo-EM structures along the translocation reaction coordinate.

**a–f**, Fourier shell correlation (FSC) resolution curves for PRE-C (**a**), PRE-H2\* (**b**), PRE-H1 (**c**), POST (**d**), INT1 (**e**) and INT2 (**f**) structures obtained by masking the two half maps and calculating the cross-resolution between the masked volumes in Relion. Resolution was estimated using the 0.143 cut-off criterion (dotted line). **g–l**, Local resolution electron density maps for PRE-C (**g**), PRE-H2\* (**h**), PRE-H1 (**i**), POST (**j**), INT1 (**k**) and INT2 (**l**). CP, central protuberance. LSU, grey background; SSU, blue background. Threshold  $\sigma = 5$ . See also Methods and Supplementary Table 1.

## Extended Data Fig. 3 Structural evidence of tRNA identity for cryo-EM structures along the translocation reaction coordinate.

**a**, Locally filtered electron density for cognate tRNA anticodon–mRNA codon (pink) interactions with deacyl-tRNA<sup>Phe</sup> (AAG, left, yellow/orange) and peptidyl-tRNA<sup>Lys</sup> (UUU, right, green) from PRE-C (top) to POST (bottom). SSU head (purple) and body (blue). Threshold  $\sigma = 6$ . **b–d**, Locally filtered electron density for modified tRNA bases in support of tRNA assignment from PRE-C (top) to POST (bottom): deacyl-tRNA<sup>Phe(AAG)</sup> ms<sup>2</sup>i<sup>6</sup>A37 (MIA, **b**); peptidyl-tRNA<sup>Lys(UUU)</sup> t<sup>6</sup>A37 (T6A, **c**); peptidyl-tRNA<sup>Lys(UUU)</sup> mnm<sup>5</sup>s<sup>2</sup>34 (U8U, **d**). Arrows designate defining density for each modification. Coloured as in **a**. **e**, Locally filtered electron density for the nascent peptide (fMet-Phe-Lys, red). Deacyl-tRNA, yellow/orange; peptidyl-tRNA, green; A2602, dark green; P loop, dark blue. Threshold  $\sigma = 5$  for PRE-C, INT1 and INT2. Threshold  $\sigma = 7$  for PRE-H2\*, PRE-H1 and POST.

## Extended Data Fig. 4 Global conformational changes within the ribosome that define the translocation reaction coordinate.

**a**, Schematic of the translocation reaction coordinate in bacteria depicting SSU rotation (blue) with respect to the LSU (grey) and SSU head-swivel (purple) processes. tRNAs are coloured on a gradient from the A (green) to P (yellow) to E (orange) sites. Deacyl-tRNA dissociation (orange) can occur at multiple steps after INT2. tRNA positions are depicted in chimeric-hybrid notation (ssu head SSU BODY/LSU). **b**, SSU conformational changes accompanying each sequential translocation step, viewed from inside the intersubunit space (left) and towards the intersubunit space from the head domain (inset), coloured by r.m.s.d. at each SSU residue for each transition. Degree of shoulder domain closure, SSU body-rotation and SSU head-swivel as compared to POST is indicated as ‘total’. Cylindrical axes and bolded text indicate the degree of SSU body-rotation (transparent black axis, right, LSU core alignment) and SSU head-swivel (solid black axis, right, SSU body alignment) as compared to the previous state on the reaction coordinate. Threshold  $\sigma = 5$ . **c**, Deacyl-tRNA (yellow/orange),

peptidyl-tRNA (green/yellow) and EF-G (red) movements during translocation. Current tRNA and EF-G positioning (solid coloured, outlined), previous position (transparent colour, no outline) and next position (white, solid outline). Alignment on the LSU core. Camera perspective is identical for all images. See also Supplementary Table 2.

**Extended Data Fig. 5 LSU interactions with tRNA and the SSU in each characterized ribosome structure along the translocation reaction coordinate.**

**a**, Location of intersubunit bridge B1 on the ribosome (top, left) and overlay of uL31 model (dark blue) depicting the conformational change during translocation (top, right). Sphere is positioned on residue Ile66 (Ca). Bottom, bridge B1 interactions between the LSU central protuberance and the SSU head domain (purple) in each structural state, including the ASF (H38, dark green) to uS13 (B1a), uL5 to uS13 (B1b) and uL31 to uS13/uS19 (B1c). Threshold  $\sigma = 4$ . **b**, Location of the L1 stalk, uL5 (dark grey), uL16 (light grey) and ASF (dark green) on the INT1 ribosome (top). L1 stalk positioning and interaction with deacyl-tRNA (yellow/orange) elbow in each classified structure (bottom). **c**, Bridge B1 interactions between the central protuberance and deacyl-/peptidyl-tRNA from PRE-C (top) to POST (bottom). P-site tRNA G19–C56 pair contacts include LSU H84 base A2309 and uL5 N-terminal Arg80. A-site tRNA contacts include G19–C56 pair coordination by ASF base A896 and TΨC stem packing against uL16 residues Arg6 and Lys7. A site (green circle), P site (yellow circle), E site (orange circle) regions are shown. **d**, Locally filtered electron density illustrating peptidyl transferase centre interactions along the translocation interaction coordinate. Watson–Crick pairing between the 3'-CCA tRNA ends and the peptidyl transferase centre. P loop (LSU rRNA 2251–2253, 2450–2451), dark blue; A loop (LSU rRNA 2553–2555, 2582–3585, dark grey; peptide, red; LSU base A2602, dark green; deacyl-tRNA, yellow-orange; peptidyl-tRNA, green; Threshold  $\sigma = 7$ . Camera perspective is identical for images within each panel. Alignment on the LSU core. See also [Supplementary Video 2](#).

## Extended Data Fig. 6 Changes in tRNA conformation during translocation.

Alignment of deacyl-tRNA (rows 1 and 2, orange) and peptidyl-tRNA (bottom, green) globally (left column), on the tRNA acceptor arm domain (positions 1–6 and 50–72, middle column, grey oval) or on the tRNA anticodon stem loop (positions 30–40, right column, pink circle). Deacyl-tRNA is coloured on a gradient from yellow to orange based on the position on the reaction coordinate (PRE-C, PRE-H2\*, PRE-H1, INT1, INT2 and POST). Peptidyl-tRNA is coloured on a gradient from green to yellow. Blue circle annotates the position of the tRNA elbow. Alignment is on the outlined circle. See also Supplementary Table 2 and Video 4.

## Extended Data Fig. 7 SSU unlocking on the leading-edge during translocation.

**a**, Locally filtered electron density illustrating the interaction between the SSU shoulder domain (cyan) A532 base, the SSU head domain (purple) G1206 of h34 and uS3 (dark purple) 193 loop near the A-site mRNA and EF-G (red) binding site. **b**, Locally filtered electron density illustrating the points of contact involved in unlocking the tRNA<sub>2</sub>–mRNA module from the SSU body domain monitoring bases (A1492, A1493) in the A site. Peptidyl-tRNA, green; mRNA, pink; H69, grey; h44, blue; h28, dark blue; h18, cyan; uS12, light blue; h34, purple; EF-G, red. **c**, Locally filtered electron density and molecular models from **b**, viewed from beneath the SSU A site. Coloured as in **b**. **d**, Cartoon schematic depicting A-site unlocking. Dotted lines indicate weak electron density. Coloured as in **b**, **c**. Camera perspective is identical for images in each panel. Alignment on LSU core. Threshold  $\sigma = 5$ . See also Fig. 2.

## Extended Data Fig. 8 Overview of the EF-G conformations and interactions in INT1 and INT2.

**a**, EF-G binding site (left) and domain architecture (middle and right) in its active, GTP conformation (INT1). Model for INT2 EF-G is shown as transparent in the right panel. **b**, Interaction between EF-G DIV (red) and

the mRNA (pink) peptidyl-tRNA (green) codon–anticodon minihelix in the A site in the INT1 (left) and INT2 (right) structures (top). Locally filtered electron density illustrating the interaction between EF-G DIV loop I Lys504 and peptidyl-tRNA (row 2), EF-G DIV loop II Gly542 and C1210 of SSU h34 (purple, row 3) and EF-G DIV loop II His584 and peptidyl-tRNA (row 4). Threshold  $\sigma = 3$ . **c**, GTP/GDP-P<sub>i</sub> conformation of EF-G (INT1, coloured) compared to a published structure of an EF-G homologue (grey, EF-G-2, PDB ID: 1WDT<sup>37</sup>). Degree of bend angle between the G domain and DIV compared to INT1 is indicated. Conformational change in DII is indicated with an arrow. Coloured as in **a**. **d**, EF-G DII conformational change (arrow) in the INT1 (left) to INT2 (right) transition, oriented by the G domain (Asn37), switch I (Trp52) and DII (His367, Asn369 and Arg371), rearranging DII contact with SSU h5 via Phe329 and Arg362. Coloured as in **a**. **e**, EF-G rotation of approximately 15° around the SRL (grey) and into the A site in the INT1 (left) to INT2 (right) transition. This movement placed SSU head nucleotide C1210 (h34, purple) in direct contact with EF-G DIV in INT2. Coloured as in **a**. Threshold  $\sigma = 5$ . **f**, Conformational rotation of the superdomain DI–DIII of EF-G in the INT1 (solid) to INT2 (transparent) transition (approximately 17°, left). Models of EF-G were superimposed on the G domain. r.m.s.d. of EF-G in the INT1 to INT2 transition mapped on the INT1 structure (right). Density absent in the INT2 structure is depicted in dark blue. Alignment on the LSU core, unless otherwise stated.

### [Extended Data Fig. 9 E-site mRNA codon interactions with SSU h23, uS7 and uS11.](#)

Locally filtered electron density illustrating the interaction between mRNA and the SSU in the E site. The black arrow annotates the –1 mRNA position, which flips in base orientation during translocation. The orange arrow annotates the gap between SSU proteins uS7 and uS11. mRNA, pink; uS7, purple; uS11, cyan; h23, blue; h24, light blue; h28, dark blue; deacyl-tRNA, yellow; peptidyl-tRNA, green. A putrescine molecule (PUT, grey) is modelled proximal to the –1 ribose in the PRE-H2\* and PRE-H1 states. Alignment on the LSU core. Threshold  $\sigma = 5$ .

## Extended Data Fig. 10 EF-G nucleotide-binding pocket in INT1 and INT2.

**a**, INT1 EF-G domain architecture (left) and zoom in of locally filtered electron density illustrating nucleotide interactions, including catalytic residues His92 of switch II and G2661 and G2662 of the SRL (grey, right). Threshold  $\sigma = 6$ . **b**, Locally filtered electron density at different thresholds ( $\sigma = 4, 6, 8, 10$ ) for the phosphates in the nucleotide-binding pocket of EF-G in INT1. **c**, Alternative modelling of GDP-P<sub>i</sub> (dark blue) in the nucleotide-binding pocket of INT1 compared to GTP (orange). Ligand cross correlation (CC) value reported from the corresponding PDB validation reports. Distances were measured between  $\gamma$  phosphate,  $\beta$  phosphate and His92 amine group. **d**, Locally filtered electron density illustrating the nucleotide-binding pocket of EF-G in the INT2 structure. Threshold  $\sigma = 6$ . G domain, red; DII, dark orange; DIII, strawberry; DIV, hot pink; DV, yellow/orange; P loop, green; switch I, yellow; switch II, lime; Mg<sup>2+</sup>, dark green. See also Fig. 3.

## Extended Data Fig. 11 SPC-binding sites evidenced in INT1.

**a**, SPC-binding site (red) on the SSU between the head (purple) and body (blue) domains beneath the SSU P site. **b**, Zoom-in of the primary SPC-binding site showing that the convex face of SPC sits in the major groove of h34 (purple, between A1193 and G1064) and that the concave face of SPC sits between h28 (dark blue, G1387) and uS5 (cyan). Lys26 of uS5 reaches to interact with the h28 phosphate G1387. **c**, Overlay of the SSU SPC-binding pocket for INT1 (coloured) and INT2 (white), showing collapse of the SPC binding site and reorientation of Lys26 during SSU head-domain swivel. SPC shown in ball-and-stick and transparent-sphere representation. Distance changes for h28 (G1387 C1'), h34 (A1196 C1') and h35 (C1066 C1') are indicated. **d**, Electron density indicative of a second SPC-binding site near the exit tunnel of the LSU between uL22 and nucleotides G488 and A1284. Threshold  $\sigma = 6$ . **e**, Overlay of the SSU from PRE-C (light grey), PRE-H2\* (light blue), PRE-H1 (blue), INT1 (light purple), INT2 (purple) and POST (dark grey) aligned on the SSU body domain to illustrate head-domain swivel. See also [Supplementary Video 6](#).

## Extended Data Fig. 12 Coordinated translocation of the tRNA<sub>2</sub>-mRNA module.

Overview of the helical architecture involved in tRNA<sub>2</sub>-mRNA translocation (left). **a–f**, Interactions between the SSU and the tRNA<sub>2</sub>-mRNA module depicted with molecular model and electron density (left) together with a schematic representation (right) for PRE-C (**a**), PRE-H2\* (**b**), PRE-H1 (**c**), INT1 (**d**), INT2 (**e**) and POST (**f**). A-site contacts with the tRNA<sub>2</sub>-mRNA module depicted here include (1) anchored h34 (plum) SSU head base C1054 stacking on tRNA at position 34, (2) intercalation of h28 (dark blue) C1397 into mRNA (pink) downstream of the A-site codon (PRE +7/+8, POST +10/+11) and (3) hydrogen bonding of the h18 (cyan) SSU shoulder G530 base with the A-site wobble position. P-site interactions include (4) anchored stacking of h31 (purple) SSU head base m<sup>2</sup>G966 at tRNA position 34, (5) h28 C1400 base stacking against the P-site wobble position, (6) A-minor interactions of SSU head bases G1338/A1339 with the tRNA minor groove, (7) electrostatic interactions of uS9 (light purple) C-terminal Arg130 residue with the anticodon U-turn motif and (8) hydrogen bonding of the h28 G926 base with the phosphate backbone between the –1 and +1 mRNA positions, anchored by A1505 (aqua). E-site interactions include (9) the anchored stacking of the SSU body h23 (light blue) G693 base against the –3 mRNA base. Distances in the schematic are not to scale. Dotted residues display weak electron density. Grey circles depict contacts that are unchanged from the previous state, yellow circles depict contacts that are different from the previous state. Peptidyl-tRNA, green; deacyl-tRNA, orange. Camera perspective is identical for all images. Alignment on the LSU core. Threshold  $\sigma = 6$ .

## Supplementary information

### Supplementary Information

This file contains Supplementary Tables 1-4 and Supplementary Figure 1.

### Reporting Summary

## Peer Review File

### Supplementary Video 1

tRNA movement through the ribosomal intersubunit space during tRNA<sub>2</sub>-mRNA translocation. Cross-section of tRNA binding sites for morphed electron density illustrating tRNA movement and ribosomal conformational changes associated with the transitions from PRE-C to PRE-H2\* to PRE-H1 to INT1 to INT2 to POST states. Modeled tRNAs and EF-Gs are shown in ribbon representation. Peptidyl-tRNA<sup>Lys</sup>, green; deacyl-tRNA<sup>Phe</sup>, orange; EF-G, red. Alignment on the LSU core. Threshold  $\sigma = 5.5$ .

### Supplementary Video 2

SSU head dynamics during tRNA<sub>2</sub>-mRNA translocation. Morphed electron density illustrating conformational changes of the SSU (30S) head (purple), the LSU (50S) L1 stalk (left), the central protuberance (CP, centre) and the GTPase activating centre (GAC, right) associated with the transitions from PRE-C to PRE-H2\* to PRE-H1 to INT1 to INT2 to POST states. Peptidyl-tRNA<sup>Lys</sup>, green; deacyl-tRNA<sup>Phe</sup>, orange; 50S, grey; 30S body domain, blue EF-G, red. Alignment on the LSU core. Threshold  $\sigma = 5.5$ .

### Supplementary Video 3

SSU rotation and head domain swivel during tRNA<sub>2</sub>-mRNA translocation. Cross-section of the intersubunit space for morphed electron density illustrating tRNA movement and conformational changes of the SSU (30S) body (blue) and head (purple) domains associated with the transitions from PRE-C to PRE-H2\* to PRE-H1 to INT1 to INT2 to POST states. Clockwise SSU rotation from PRE-C to PRE-H, counterclockwise rotation from INT1 to INT2 and from INT2 to POST. Peptidyl-tRNA<sup>Lys</sup>, green; deacyl-tRNA<sup>Phe</sup>, orange; LSU (50S), grey; EF-G, red; mRNA, pink. Alignment on the LSU core. Threshold  $\sigma = 5.5$ .

### Supplementary Video 4

tRNA conformational changes during translocation. Conformational heterogeneity of deacyl-tRNA (yellow to orange) and peptidyl-tRNA (green to yellow) during translocation from PRE-C to PRE-H2\* to PRE-H1 to INT1 to INT2 to POST states. Red ribbon represents reversal of the translocation reaction coordinate (POST to PRE-C). Alignment on the anticodon. See also Extended Data Fig. 6.

## Supplementary Video 5

Unlocking of the A-site mRNA channel and peptidyl-tRNA movement during translocation. Morphed electron density illustrating dynamics of the SSU (30S) shoulder (blue) and head (purple) domains associated with the transitions from PRE-C to PRE-H2\* to PRE-H1 to INT1 to INT2 to POST states. Key SSU nucleotides A532 of h18, A1492/A1493 of h44 and A1913 of H69 are annotated. Peptidyl-tRNA<sup>Lys</sup>, green; deacyl-tRNA<sup>Phe</sup>, orange; LSU (50S), grey; EF-G, red; mRNA, pink. Alignment on the LSU core. Threshold  $\sigma = 5.5$ .

## Supplementary Video 6

Spectinomycin binding site in the INT1 structure. Illustration of the hydrogen bond network (cyan) with rRNA of the SSU (30S, purple) and uS5 Lys26 (orange) in the spectinomycin (SPC, grey) binding pocket. In part 2, electron density is coloured by element. See also Extended Data Fig. 11.

## Supplementary Video 7

Conformational changes in EF-G during the INT1-INT2 transition. Molecular morph of EF-G in the transition from INT1 to INT2. Peptidyl-tRNA<sup>Lys</sup>, green; deacyl-tRNA<sup>Phe</sup>, orange; LSU (50S), grey; SSU (30S) body, blue; and SSU head, purple; EF-G, coloured by domain. Alignment on the LSU core. Threshold  $\sigma = 5.5$ .

## **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



Check for  
updates

## Cite this article

Rundlet, E.J., Holm, M., Schacherl, M. *et al.* Structural basis of early translocation events on the ribosome. *Nature* **595**, 741–745 (2021).  
<https://doi.org/10.1038/s41586-021-03713-x>

## Download citation

- Received: 11 January 2021
- Accepted: 09 June 2021
- Published: 07 July 2021
- Issue Date: 29 July 2021

- DOI: <https://doi.org/10.1038/s41586-021-03713-x>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03713-x>

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

- Article
- [Published: 05 July 2021](#)

# Structure of a mammalian sperm cation channel complex

- [Shiyi Lin](#) [ORCID: orcid.org/0000-0002-9026-8275](#)<sup>1,2,3 na1</sup>,
- [Meng Ke](#) [ORCID: orcid.org/0000-0002-9213-7439](#)<sup>1,2,3 na1</sup>,
- [Yuqi Zhang](#) [ORCID: orcid.org/0000-0001-6899-111X](#)<sup>1,2,3</sup>,
- [Zhen Yan](#) [ORCID: orcid.org/0000-0002-1999-8451](#)<sup>1,2,3</sup> &
- [Jianping Wu](#) [ORCID: orcid.org/0000-0001-9869-7691](#)<sup>1,2,3</sup>

[Nature](#) volume **595**, pages 746–750 (2021) [Cite this article](#)

- 10k Accesses
- 72 Altmetric
- [Metrics details](#)

## Subjects

- [Cryoelectron microscopy](#)
- [Permeation and transport](#)
- [Reproductive biology](#)

## Abstract

The cation channel of sperm (CatSper) is essential for sperm motility and fertility<sup>1,2</sup>. CatSper comprises the pore-forming proteins CATSPER1–4 and multiple auxiliary subunits, including CATSPER $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\varepsilon$ ,  $\zeta$ , and EFCAB9<sup>1,3,4,5,6,7,8,9</sup>. Here we report the cryo-electron microscopy (cryo-

EM) structure of the CatSper complex isolated from mouse sperm. In the extracellular view, CATSPER1–4 conform to the conventional domain-swapped voltage-gated ion channel fold<sup>10</sup>, following a counterclockwise arrangement. The auxiliary subunits CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\varepsilon$ —each of which contains a single transmembrane segment and a large extracellular domain—constitute a pavilion-like structure that stabilizes the entire complex through interactions with CATSPER4, 1, 3 and 2, respectively. Our EM map reveals several previously uncharacterized components, exemplified by the organic anion transporter SLCO6C1. We name this channel–transporter ultracomplex the CatSpermosome. The assembly and organization details of the CatSpermosome presented here lay the foundation for the development of CatSpermosome-related treatments for male infertility and non-hormonal contraceptives.

## 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.

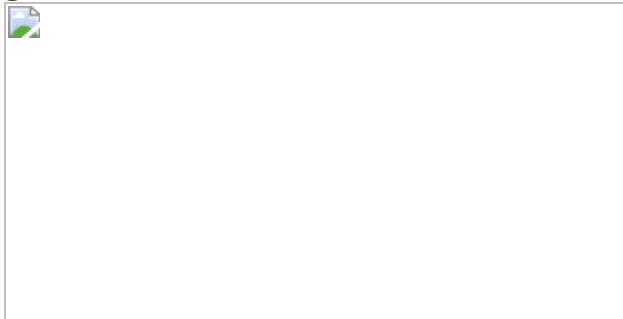
### **Additional access options:**

- [Log in](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

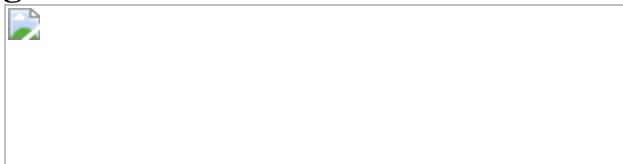
**Fig. 1: Cryo-EM reconstruction of the mouse CatSpermasome.**



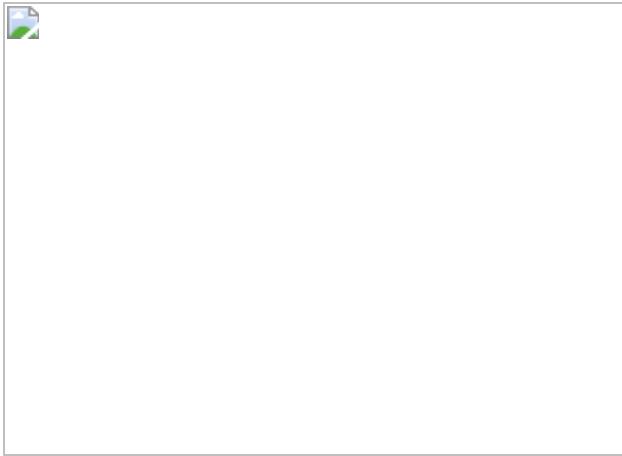
**Fig. 2: The heterotetrameric channel formed by CATSPER1–4.**



**Fig. 3: Interactions between CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$  and CATSPER1–4.**



**Fig. 4: Previously uncharacterized components of the CatSpermasome.**



## Data availability

The cryo-EM maps of the mouse CatSpermasome and the corresponding atomic coordinate have been deposited in the Electron Microscopy Data Bank and the Protein Data Bank under the accession codes [EMD-31076](#) and [7ZEEB](#), respectively. The mass spectrometry data have been deposited in the MassIVE database (<https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp>) under the accession number MSV0000987325. All data analysed during this study are included in this Article and its Supplementary Information. Any other relevant data are available from the corresponding author upon reasonable request.

## References

1. 1.

Ren, D. et al. A sperm ion channel required for sperm motility and male fertility. *Nature* **413**, 603–609 (2001).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

2. 2.

Ren, D. & Xia, J. Calcium signaling through CatSper channels in mammalian fertilization. *Physiology* **25**, 165–175 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Quill, T. A., Ren, D., Clapham, D. E. & Garbers, D. L. A voltage-gated ion channel expressed specifically in spermatozoa. *Proc. Natl Acad. Sci. USA* **98**, 12527–12531 (2001).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

4. 4.

Lobley, A., Pierron, V., Reynolds, L., Allen, L. & Michalovich, D. Identification of human and mouse CATSPER3 and CATSPER4 genes: characterisation of a common interaction domain and evidence for expression in testis. *Reprod. Biol. Endocrinol.* **1**, 53 (2003).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Liu, J., Xia, J., Cho, K. H., Clapham, D. E. & Ren, D. CatSper $\beta$ , a novel transmembrane protein in the CatSper channel complex. *J. Biol. Chem.* **282**, 18945–18952 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

6. 6.

Wang, H., Liu, J., Cho, K. H. & Ren, D. A novel, single, transmembrane protein CATSPERG is associated with CATSPER1 channel protein. *Biol. Reprod.* **81**, 539–544 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

7. 7.

Chung, J. J., Navarro, B., Krapivinsky, G., Krapivinsky, L. & Clapham, D. E. A novel gene required for male fertility and functional

CATSPER channel formation in spermatozoa. *Nat. Commun.* **2**, 153 (2011).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

8. 8.

Chung, J. J. et al. CatSperζ regulates the structural continuity of sperm Ca<sup>2+</sup> signaling domains and is required for normal fertility. *eLife* **6**, e23082 (2017).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

9. 9.

Hwang, J. Y. et al. Dual sensing of physiologic pH and calcium by EFCAB9 regulates sperm motility. *Cell* **177**, 1480–1494 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Long, S. B., Campbell, E. B. & Mackinnon, R. Voltage sensor of Kv1.2: structural basis of electromechanical coupling. *Science* **309**, 903–908 (2005).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

11. 11.

Whitaker, M. Calcium at fertilization and in early development. *Physiol. Rev.* **86**, 25–88 (2006).

[MathSciNet](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Okabe, M. The cell biology of mammalian fertilization. *Development* **140**, 4471–4479 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

13. 13.

Wang, H. F., McGoldrick, L. L. & Chung, J. J. Sperm ion channels and transporters in male fertility and infertility. *Nat. Rev. Urol.* **18**, 46–66 (2021).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Chung, J. J. et al. Structurally distinct  $\text{Ca}^{2+}$  signaling domains of sperm flagella orchestrate tyrosine phosphorylation and motility. *Cell* **157**, 808–822 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Seifert, R. et al. The CatSper channel controls chemosensation in sea urchin sperm. *EMBO J.* **34**, 379–392 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Yoshida, M. & Yoshida, K. Sperm chemotaxis and regulation of flagellar movement by  $\text{Ca}^{2+}$ . *Mol. Hum. Reprod.* **17**, 457–465 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

17. 17.

Strunker, T. et al. The CatSper channel mediates progesterone-induced  $\text{Ca}^{2+}$  influx in human sperm. *Nature* **471**, 382–386 (2011).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

18. 18.

Lishko, P. V., Botchkina, I. L. & Kirichok, Y. Progesterone activates the principal  $\text{Ca}^{2+}$  channel of human sperm. *Nature* **471**, 387–391 (2011).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

19. 19.

Avenarius, M. R. et al. Human male infertility caused by mutations in the CATSPER1 channel protein. *Am. J. Hum. Genet.* **84**, 505–510 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

20. 20.

Hildebrand, M. S. et al. Genetic male infertility and mutation of CATSPER ion channels. *Eur. J. Hum. Genet.* **18**, 1178–1184 (2010).

[CAS](#) [Google Scholar](#)

21. 21.

Jin, Z. R. et al. Roles of CatSper channels in the pathogenesis of asthenozoospermia and the therapeutic effects of acupuncture-like treatment on asthenozoospermia. *Theranostics* **11**, 2822–2844 (2021).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

Avidan, N. et al. CATSPER2, a human autosomal nonsyndromic male infertility gene. *Eur. J. Hum. Genet.* **11**, 497–502 (2003).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

23. 23.

Zhang, Y. et al. Sensorineural deafness and male infertility: a contiguous gene deletion syndrome. *J. Med. Genet.* **44**, 233–240 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

24. 24.

Luo, T. et al. A novel copy number variation in CATSPER2 causes idiopathic male infertility with normal semen parameters. *Hum. Reprod.* **34**, 414–423 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

25. 25.

Wang, J. X. et al. Patient with CATSPER3 mutations-related failure of sperm acrosome reaction with successful pregnancy outcome from intracytoplasmic sperm injection (ICSI). *Mol. Genet. Genom. Med.* **9**, e1579 (2020).

[Google Scholar](#)

26. 26.

Brown, S. G. et al. Homozygous in-frame deletion in CATSPER in a man producing spermatozoa with loss of CatSper function and compromised fertilizing capacity. *Hum. Reprod.* **33**, 1812–1816 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

27. 27.

Quill, T. A. et al. Hyperactivated sperm motility driven by CATSPER2 is required for fertilization. *Proc. Natl Acad. Sci. USA* **100**, 14869–14874 (2003).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

28. 28.

Qi, H. et al. All four CatSper ion channel proteins are required for male fertility and sperm cell hyperactivated motility. *Proc. Natl Acad. Sci. USA* **104**, 1219–1223 (2007).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Jin, J. et al. Catsper3 and Catsper4 are essential for sperm hyperactivated motility and male fertility in the mouse. *Biol. Reprod.* **77**, 37–44 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

30. 30.

Nikpoor, P., Mowla, S. J., Movahedin, M., Ziaeef, S. A. M. & Tiraihi, T. CatSper gene expression in postnatal development of mouse testis and in subfertile men with deficient sperm motility. *Hum. Reprod.* **19**, 124–128 (2004).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

31. 31.

Long, J. E., Lee, M. S. & Blithe, D. L. Male contraceptive development: update on novel hormonal and nonhormonal methods. *Clin. Chem.* **65**, 153–160 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

32. 32.

Cheng, C. Y. & Mruk, D. D. New frontiers in nonhormonal male contraception. *Contraception* **82**, 476–482 (2010).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

33. 33.

Kirichok, Y., Navarro, B. & Clapham, D. E. Whole-cell patch-clamp measurements of spermatozoa reveal an alkaline-activated  $\text{Ca}^{2+}$  channel. *Nature* **439**, 737–740 (2006).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

34. 34.

Tao, X., Lee, A., Limapichat, W., Dougherty, D. A. & MacKinnon, R. A gating charge transfer center in voltage sensors. *Science* **328**, 67–73 (2010).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

35. 35.

Xie, J. F. et al. Structure of the human sodium leak channel NALCN in complex with FAM155A. *Nat. Commun.* **11**, 5831 (2020).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Hagenbuch, B. & Stieger, B. The SLCO (former SLC21) superfamily of transporters. *Mol. Aspects Med.* **34**, 396–412 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Fietz, D. et al. Membrane transporters for sulfated steroids in the human testis—cellular localization, expression pattern and functional analysis. *PLoS One* **8**, e62638 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

38. 38.

Suzuki, T. et al. Identification and characterization of novel rat and human gonad-specific organic anion transporters. *Mol. Endocrinol.* **17**, 1203–1215 (2003).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

39. 39.

Abbott, G. W. Chansporter complexes in cell signaling. *FEBS Lett.* **591**, 2556–2576 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

40. 40.

Li, N. N. et al. Structure of a pancreatic ATP-sensitive potassium channel. *Cell* **168**, 101–110.e10 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

41. 41.

Perez, G. I. et al. Prolongation of ovarian lifespan into advanced chronological age by *Bax*-deficiency. *Nat. Genet.* **21**, 200–203 (1999).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Lei, J. & Frank, J. Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. *J. Struct. Biol.* **150**, 69–80 (2005).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

43. 43.

Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods*

**14**, 331–332 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

44. 44.

Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

[CAS](#) [Google Scholar](#)

45. 45.

Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Scheres, S. H. W. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

50. 50.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

[CAS](#) [PubMed](#) [Google Scholar](#)

51. 51.

Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

52. 52.

Smart, O. S., Neduvvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. P. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph. Model.* **14**, 354–360 (1996).

[CAS](#) [Google Scholar](#)

53. 53.

Yang, J. et al. Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl Acad. Sci. USA* **117**, 1496–1503 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

54. 54.

Nicholls, R. A., Fischer, M., McNicholas, S. & Murshudov, G. N. Conformation-independent structural comparison of macromolecules with ProSMART. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 2487–2499 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

55. 55.

Amunts, A. et al. Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485–1489 (2014).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

56. 56.

DeLano, W. L. *The PyMOL Molecular Graphics System* <http://www.pymol.org> (2002).

[Download references](#)

## Acknowledgements

We thank N. Yan and H. Yu for critical reading of the manuscript; M. Jiang and Y. Ru for help with the immunofluorescence experiments; the Cryo-EM Facility and HPC Center of Westlake University for providing cryo-EM and computation support; S. Feng and the Mass Spectrometry & Metabolomics Core Facility of Westlake University for protein sample MS analysis; and J. Bao, D. Wu and the Laboratory Animal Resources Center of Westlake University for help with animal maintenance and IVF experiments. This work was supported by Westlake Laboratory (Westlake Laboratory of Life Sciences and Biomedicine) (W101486022101) and an Institutional Startup Grant from the Westlake Education Foundation (101486021901) to J.W.

## Author information

## Author notes

1. These authors contributed equally: Shiyi Lin, Meng Ke

## Affiliations

1. Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, China

Shiyi Lin, Meng Ke, Yuqi Zhang, Zhen Yan & Jianping Wu

2. Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, China

Shiyi Lin, Meng Ke, Yuqi Zhang, Zhen Yan & Jianping Wu

3. Institute of Biology, Westlake Institute for Advanced Study, Hangzhou, China

Shiyi Lin, Meng Ke, Yuqi Zhang, Zhen Yan & Jianping Wu

## Authors

1. Shiyi Lin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Meng Ke

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Yuqi Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Zhen Yan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Jianping Wu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

## Contributions

J.W. conceived and supervised the project; S.L. and Y.Z. prepared the protein sample under the guidance of Z.Y. and J.W.; S.L. collected the cryo-EM data and M.K. calculated the cryo-EM map; J.W. built the model; all authors contributed to data analysis; and J.W. wrote the manuscript.

## Corresponding author

Correspondence to [Jianping Wu](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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

## Extended data figures and tables

## Extended Data Fig. 1 Mouse genotyping, confocal imaging of spermatozoa and endogenous purification of the CatSpermasome from mouse sperm.

**a**, The genotype of each mouse was verified by two PCR reactions. Top, schematic of the genotyping procedure. Bottom, a group of representative PCR results. NC, negative control (empty template); WT, wild type; KI/+, heterozygote knock-in; KI/KI, homozygote knock-in. In total, 49 wild-type, 676 KI/+, and 207 KI/KI mice were verified. **b**, An drawing of a mouse sperm. CatSper was mainly distributed at the principal piece of spermatozoa. **c**, EGFP fluorescence was detected in the principal piece (red arrows) of knock-in mouse spermatozoa, but not wild-type mouse spermatozoa. Blue arrows, auto-fluorescence signal observed in the middle piece of spermatozoa. Shown here are one of the two images taken for each sample. Scale bars, 30  $\mu$ m. **d**, Schematic of CatSpermasome purification. **e**, The purified protein sample was subjected to gel filtration analysis. The peak fractions of CatSpermasome (arrow) were collected and concentrated for cryo-EM and MS studies. Inset, the cross-linked protein sample was visualized on SDS-PAGE by silver staining. The corresponding protein band (arrow) in a separate gel without staining was cut out for MS analysis. For gel source data, see Supplementary Fig. 1. **f**, A representative EM micrograph of the CatSpermasome sample stained with uranyl acetate (one micrograph out of five in total for the negative staining sample). Scale bar, 50 nm.

## Extended Data Fig. 2 Mass spectrometric analysis of the purified CatSpermasome.

**a**, MS samples were the same as used for the cryo-EM study. MS detected proteins are shown in order of decreasing confidence. All previously characterized CatSper components are listed (yellow). The top six entries with highest peptide spectrum match (PSM) values are all CATSPER proteins. Most of the contaminating proteins are cytoskeletal proteins. The newly identified components SLCO6C1 and CATSPER $\eta$  are highlighted in red. TMEM249, which is probably another new component, is shaded in

light blue. **b**, Representative MS spectra for specific peptides of CATSPER1, SLCO6C1, CATSPER $\eta$ , and TMEM249.

### Extended Data Fig. 3 Cryo-EM analysis of mouse CatSpermasome.

**a**, A representative motion-corrected micrograph of the CatSpermasome cryo-EM sample out of a dataset of 16,648 images. Scale bar, 50 nm. **b**, Two-dimensional class averages. Box size, 430 Å. **c**, Gold standard FSC curves for the 3D reconstructions. The curves were calculated with masks for the entire protein (overall map), and for masked regions of corresponding maps. See **f** for each map. **d**, Validation of the final structure models. FSC curves of the final refined model versus the summed map that it was refined against (black); of the model versus the first half-map (blue); and of the model versus the second half-map (red). The small difference between the blue and red curves indicates that the refinement of the atomic coordinates did not suffer from overfitting. **e**, Angular distribution of the particles of the final reconstruction generated by cryoSPARC. **f**, Flowchart of EM data processing (see ‘Image processing’ in Methods).

### Extended Data Fig. 4 Structural features of the cytosolic regions.

**a**, The cytosolic region consists of two separate but interacting parts. Cytosolic map 1 is in close contact with the S6 segments of CATSPER2 and CATSPER3. A density surrounded in the bottom of cytosolic map 1 (shown in green) is connected to the density of the S6 segment of CATSPER3 and may belong to the carboxyl end of CATSPER3. Owing to limited resolution, the identity of cytosolic map 1 remains to be determined. Cytosolic map 2, however, is likely to be the subcomplex of EFCAB9 and CATSPER $\zeta$ . The maps were generated in ChimeraX. **b**, Predicted structures of EFCAB9 by tFold and CATSPER $\zeta$  by trRosetta. EFCAB9 has two EF-hand motif-containing lobes, which is very similar to calmodulin. CATSPER $\zeta$  consists mainly of  $\alpha$ -helices. **c**, Docking of the predicted structures of EFCAB9 and CATSPER $\zeta$  into cytosolic map 2. The two lobes of EFCAB9 are in a compact conformation instead of the extended

conformation in the predicted structure. The main body of CATSPER $\zeta$  (light cyan in **b**) can be fitted into the remaining density near EFCAB9 in cytosolic map 2. Several fitted  $\alpha$ -helices are indicated by arrows.

### Extended Data Fig. 5 EM maps of CATSPER1–4.

**a**, Electron density maps of each segment of the transmembrane helices of CATSPER1–4. The boundaries of each displayed segment are labelled. The densities, shown as blue meshes, are contoured at 3–4 $\sigma$  in PyMOL. **b**, Electron density map of the selectivity filter and the pore helices. The densities are contoured at 4 $\sigma$ . Two tentatively assigned Na $^+$  ions are shown as purple spheres. **c**, Electron density maps of detergent-like molecules. These densities may also belong to cholesterol or steroid hormones under physiological conditions. Three GDN molecules are tentatively assigned to these densities. The densities are found in a semi-open cavity formed by the S3, S4 and S4–5 segments of CATSPER1, but not CATSPER2–4, whose corresponding cavities are smaller. **d**, Structural superimposition of CATSPER1–4 indicates that CATSPER1 has a larger cavity for binding of detergent-like molecules.

### Extended Data Fig. 6 Structural details and sequence alignment of the pore domain and VSDs.

**a**, Overall structure of the pore domain of CATSPER1–4. The critical DDDD residues in the selectivity filter are shown as sticks. Each S6 segment contains a  $\pi$ -helix turn (red arrows). **b**, Sequence alignment of the selectivity filter and the pore helices among mouse CATSPER1–4, human CATSPER1–4, and rabbit Ca $_v$ 1.1. The invariant Thr and Trp residues are shaded cyan. The DDDD residues in the selectivity filter of CATSPER1–4 and the corresponding residues EEEE in Ca $_v$ 1.1 are highlighted red. **c**, Sequence alignment of the VSDs among CATSPER1–4. The boundary of each segment is shaded light grey. Positively charged residues on S4 segments are shaded blue and residues corresponding to positions R1–R6 are boxed. An1 and CTC residues on segments S2 and S3 are shaded purple. **d**, Structural comparison of the VSDs among CATSPER1–4. The four VSDs are superimposed relative to CTC and An1 on S2. For visual

clarity, the S1 segments are omitted and only the side chains of aligned residues and R4 residues on the S4 segments are shown.

### **Extended Data Fig. 7 Structure of the auxiliary subunits CATSPER $\beta$ , $\gamma$ , $\delta$ and $\epsilon$ .**

**a**, The overall structures of CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$  share similar domain organizations. The structures are shown in cartoon form and the sugar moieties in the glycosylation sites are shown as sticks. The NTDs,  $\beta$ -propeller domains, Ig-like domains, stem domains and transmembrane domains are coloured green, blue, orange, yellow and salmon, respectively. The head domain in CATSPER $\beta$  and the NTD2 domain in CATSPER $\epsilon$  are coloured slate and cyan, respectively. **b**, Domain organization of CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$ . The boundaries for each domain and the identified glycosylation sites are labelled. The disulfide bonds are indicated by orange lines. See Extended Data Table 2 for details. **c**, Inter-subunit interactions among CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$ , shown in four side views. CATSPER $\beta$  is domain coloured and CATSPER $\gamma$ ,  $\delta$  and  $\epsilon$  are coloured as in Fig. 1. The side openings formed by two adjacent subunits are indicated by dotted lines. For visual clarity, the transmembrane helices are omitted. **d**, Extracellular view of the auxiliary subunits. The top of the channel is sealed by the Ig-like domains of CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$ . Bottom, close-up view of the interactions among the Ig-like domains (box in top image). The residues that mediate the interactions in the interface are shown as sticks. Hydrogen bonds are indicated by red dashed lines. **e**, Interactions between the transmembrane domains of CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$  and the adjacent VSDs of CATSPER1–4. The residues that contribute to the interface interactions are shown as sticks. Potential hydrogen bonds are indicated by red dashed lines. The stem domains of CATSPER $\beta$  and CATSPER $\epsilon$  are further from the adjacent channel subunits than those of CATSPER $\gamma$  and CATSPER $\delta$  (double-headed arrows).

### **Extended Data Fig. 8 Representative EM maps of the auxiliary subunits CATSPER $\beta$ , $\gamma$ , $\delta$ and $\epsilon$ .**

The selected segments cover almost every domain of CATSPER $\beta$ ,  $\gamma$ ,  $\delta$  and  $\varepsilon$ . Representative densities for the glycosylation sites from each subunit are also presented. The densities, shown as blue meshes, are contoured at 4–5 $\sigma$  in PyMOL.

### **Extended Data Fig. 9 Structural analysis and sequence alignment of SLCO6C1.**

**a**, Side view (left) and central slice view (right) of the electron density map of SLCO6C1 and the adjacent components. The map was generated in ChimeraX. **b**, SLCO6C1 is captured in an outward-facing conformation. The structure is shown in cylindrical helices cartoon mode and coloured by domain. Right, the N and C domains of SLCO6C1 each contain six transmembrane helices and are pseudo-symmetric, with an r.m.s.d. of ~6 Å when superimposed. **c**, Sequence alignment among mouse SLCO6C1, rat SLCO6C1 and human SLCO6A1. The glycosylation site (red triangle) and the residues that may be involved in interactions with CATSPER $\varepsilon$  (yellow triangles) are conserved among species. The uniport IDs for the aligned sequences are: mSLCO6C1: Q3V161; rSLCO6C1: G3V7R7; and hSLCO6A1: Q86UG4.

### **Extended Data Fig. 10 Structural characterization of CATSPER $\eta$ and TMEM249.**

**a**, Side view (left) and central slice view (right) of the electron density maps of CATSPER $\eta$  and the adjacent components. The map was generated in ChimeraX. **b**, Electron density maps of CATSPER $\eta$ . The map allows accurate assignment of the side chains of the bulky residues. The densities, shown as blue meshes, are contoured at 4 $\sigma$  in PyMOL. **c**, Predicted structure of TMEM249 by tFold perfectly fits into the density near CATSPER $\eta$  with minor adjustments. The structure has the characteristics of two transmembrane helices and a  $\beta$ -sheet (arrows) in the cytosolic domain. **d**, TMEM249 interacts with CATSPER $\eta$  and VSD4 in the transmembrane and cytosolic regions, respectively. Unlike CATSPER $\eta$ , TMEM249 does not interact with the stem domain of CATSPER $\beta$ .

## [Extended Data Fig. 11 Immunofluorescence detection of representative CatSpermosome components in wild-type sperm.](#)

CATSPER1, CATSPER4, CATSPER $\beta$  and TMEM249 are mainly distributed in the principal piece of sperm (red fluorescence signal). The sperm head, middle piece and principal piece are indicated by black, blue, and red arrows, respectively. NC, negative control (without primary antibody). Shown here are one of three images taken for each sample. Scale bars, 20  $\mu$ m.

### **Extended Data Table 1 Statistics for data collection and structural refinement**

[Full size table](#)

### **Extended Data Table 2 Summary of model building of the mouse CatSpermosome**

[Full size table](#)

## **Supplementary information**

### [Supplementary Figure 1](#)

Uncropped raw data gel from Extended Data Fig. 1 - SDS-PAGE gel for Extended Data Fig. 1e showing the BS3-crosslinked purified mouse CatSpermosome sample by silver staining.

### [Reporting Summary](#)

### [Video 1](#)

: Overall EM map of the mouse CatSpermosome The displayed overall composite map was in the same color scheme as Fig. 1. The movie was generated in ChimeraX

## **Rights and permissions**

## [Reprints and Permissions](#)

## About this article



Check for  
updates

## Cite this article

Lin, S., Ke, M., Zhang, Y. *et al.* Structure of a mammalian sperm cation channel complex. *Nature* **595**, 746–750 (2021).  
<https://doi.org/10.1038/s41586-021-03742-6>

## [Download citation](#)

- Received: 17 March 2021
- Accepted: 21 June 2021
- Published: 05 July 2021
- Issue Date: 29 July 2021
- DOI: <https://doi.org/10.1038/s41586-021-03742-6>

## Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

# Associated Content

## Sperm CatSper ion channel swims into sharper focus

- David E. Clapham
- Raymond E. Hulse

Advertisement

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03742-6>

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