

nature

FLUID DYNAMICS

Artificial cilia offer electronic control over surface liquid flows

Collision course
How the upgraded LHC aims to hunt down particle anomalies

The human factor
Why patent law is being undermined by artificial intelligence

Proliferation path
Variable expression of a metabolic enzyme can aid metastasis

WIRELESS
ENERGY



Nature.2022.05.28

[Sat, 28 May 2022]

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

This Week

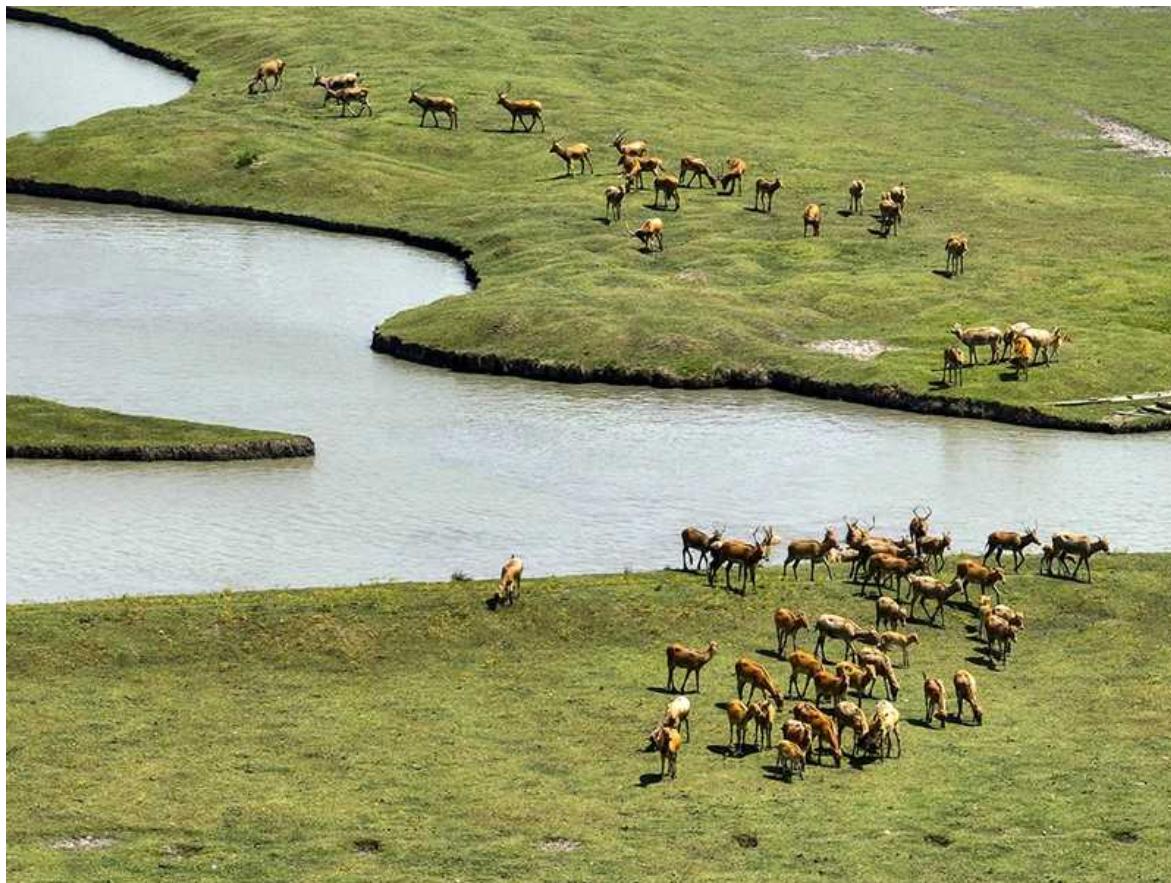
- **Cash and action are needed to avert a biodiversity crisis** [25 May 2022]
Editorial • To stop catastrophic losses of animal and plant species, countries need to move ahead with talks and give generously — despite the COVID-19 pandemic.
- **Schools need research to guide the recovery from COVID disruption** [25 May 2022]
Editorial • Evidence from around the world can help children to catch up, and could improve education as a whole.
- **COVID lessons from Japan: the right messaging empowers citizens** [23 May 2022]
World View • There's no perfect solution to suppress the pandemic, but careful study and communication are key.
- **Here in Ukraine, science continues under air raids** [24 May 2022]
World View • Three months after the Russian invasion, our Kyiv office is open again.

Cash and action are needed to avert a biodiversity crisis
[Download PDF](#)

- EDITORIAL
- 25 May 2022

Cash and action are needed to avert a biodiversity crisis

To stop catastrophic losses of animal and plant species, countries need to move ahead with talks and give generously — despite the COVID-19 pandemic.



Ambitious new targets are needed to conserve nature by protecting parks and species.
Credit: Tang Dehong/VCG/Getty

It will take ample time and money to slow the world's catastrophic loss of plant and animal species — and right now, both are running dangerously low. This year, nations are due to agree to an action plan to protect global biodiversity at the 15th Conference of the Parties (COP15) to the United Nations Convention on Biological Diversity. But the meeting is already two years late because of the pandemic, and China, which will host the conference in Kunming, has [yet to set a new date](#).

Now, conflicts over financing are adding to the tension. Conservation groups and advocates suggest that rich nations must donate at least US\$60 billion annually to help less-affluent ones to fund projects such as protecting areas where wildlife can thrive and tackling the illegal wildlife trade that is driving hundreds of species to extinction. This is much more than the \$4 billion to \$10 billion that they are estimated to be spending today, and well below the amount they are giving low- and middle-income countries (LMICs) to fight climate change, which reached around \$50 billion in 2019 according to one estimate. Yet limited overseas development funds are spread ever thinner as donors deal with the pandemic and now the fallout from Russia's invasion of Ukraine. This is where COP15 is meant to deliver: as well as agreeing to the action plan, called the Global Biodiversity Framework, nations will be encouraged to pledge more money.

A mix of [public and private money has started to trickle in](#). Currently, biodiversity funding on the table ahead of COP15 amounts to roughly \$5.2 billion per year, according to estimates by a group of five leading conservation organizations. Most comes from six governments, including France, the United Kingdom and Japan, and the European Union. In April, the Global Environment Facility (GEF) — a multilateral fund to support international environmental agreements — announced that, over the next four years, around \$1.9 billion will go to projects dedicated to biodiversity. However, it's unclear how much of this will come from the coffers that donor countries have already pledged.

Some cash for conservation is coming from private philanthropic donors — such as \$2 billion committed by entrepreneur Jeff Bezos last year. And starting in 2020, a group of financial institutions (now 89 of them) promised to annually report their financing activities and investments that affect biodiversity, and to move away from those that do harm — a form of

ecological accounting that could help to shrink the budget needed to protect biodiversity.

Donors will need to reach much deeper into their pockets to meet the demands of LMICs, the custodians of much of the world's biodiversity. In March, a group of LMICs, led by Gabon, asked for \$100 billion per year in new funding when officials met in Geneva, Switzerland, to discuss progress on the Global Biodiversity Framework. The LMICs want the money placed in a new multilateral fund for biodiversity, separate from, but complementary to, the GEF.

Aside from cash, the fund will need to find a new home and structure — and there are a few options. A proposal from Brazil, circulated at the Geneva meeting, suggests the fund be governed by a board of 24 members, with an equal number from rich and lower-income nations. The board would be responsible for funding decisions and would prioritize projects that help to achieve the biodiversity convention's goals. The pitch generated interest among some countries, but also concerns that it's an attempt by Brazil to divert attention from its failure over the past few years to protect the Amazon rainforest and [prevent other environmental harm](#).

Another option is the Kunming Biodiversity Fund, which China announced in October last year to help LMICs to safeguard their ecosystems. It allocated 1.5 billion yuan (US\$223 million) to seed the fund and invited other countries to contribute, but so far none has. Sources knowledgeable about the fund say that donor countries are reluctant to pitch in because China is holding on too tightly to the reins and is not involving others in its deliberations. Details of how the fund will operate are scarce, but Nature has learnt that China is floating the idea of housing it at the Asian Infrastructure Investment Bank (AIIB), based in Beijing. Set up in 2016, the AIIB has \$100 billion in total capital and 105 members, including Germany, France and the United Kingdom. The AIIB has big green plans. By 2025, it wants half of all infrastructure projects it finances to focus on climate issues. With rigorous oversight and transparency, the AIIB would make a [good home for the Kunming fund](#).

As countries prepare to meet in Nairobi on 20–26 June in a last-ditch attempt to push the biodiversity framework forwards before COP15, China,

as the host, must urgently provide stronger leadership on financing, including more transparency and engagement. Progress will require quick, generous contributions from donor nations — which should [prioritize grants, not loans, for biodiversity projects](#).

Holding the COP15 meeting must be a priority, too. As China tightens restrictions in the face of a COVID-19 surge, some researchers fear that delays will stretch on, stalling conservation work and leaving less time to meet biodiversity targets. China must either commit to holding the meeting this year or let it proceed elsewhere. One option being quietly discussed is moving the meeting to Canada — home of the United Nations biodiversity convention's secretariat — and this deserves consideration. The world needs an ambitious biodiversity plan now — nature cannot wait.

Nature **605**, 587 (2022)

doi: <https://doi.org/10.1038/d41586-022-01430-7>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01430-7>

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

Schools need research to guide the recovery from COVID disruption
[Download PDF](#)

- EDITORIAL
- 25 May 2022

Schools need research to guide the recovery from COVID disruption

Evidence from around the world can help children to catch up, and could improve education as a whole.



COVID-19 disruptions forced schools to innovate.Credit: David Gray/Getty

One of the most shocking impacts of the COVID-19 pandemic has nothing to do with infections or health. It is that school closures have damaged the education of some 1.6 billion children around the world.

Two years into the pandemic, schools had been completely closed for an average of more than 4.5 months across countries. One in ten countries had closed schools for more than nine months, according to the United Nations cultural organization UNESCO, and millions of children around the world had not gone back at all. Data are still coming in, but they are starting to confirm what everyone feared: that the children facing the biggest setbacks in learning are those who are poor or otherwise disadvantaged. And it is well established that learning losses leave lifelong scars, so it is likely that this will lead to lost opportunities and lower incomes for decades to come.

Many nations want to minimize these losses, but returning schools to business as usual would be a mistake. Instead, they should use this moment to improve teaching and education systems, informed by research. As a [Feature published today](#) makes clear, researchers have built up large bodies of evidence, including randomized controlled trials, that point to cost-effective methods of improving school attendance and learning in both low- and high-income countries. These strategies range from providing information to parents and children about the long-term benefits of education, to helping children comprehend what they read, involving parents in children's education, giving children meaningful feedback on their work and helping students to plan and evaluate their own learning.

Too often, such research is overlooked by educators and ignored by policymakers who mistakenly think they know what works best. But putting evidence-based insights to work in classrooms around the world would help children to recover from the educational damage inflicted by the pandemic. It would also strengthen entire education systems, many of which were failing children well before COVID-19 struck. Many children are denied education by conflict, poverty or politics, including crises such as the war in Ukraine and the Taliban's decision to exclude many girls from school in Afghanistan.

Proponents of evidence-informed education have to be realistic about the limitations of research. One big challenge is the huge variation in classrooms

and schools within countries and around the world. A tutoring programme shown to be effective at one school might not work at another if the children's ages, learning styles or home environments differ, or if it is implemented in a different way. Educational research tends to serve as a guide — but it is not a guarantee that something will work for a particular classroom or child, or when scaled up across a nation. It's therefore invaluable for teachers to be fully involved in research and in applying results.

A bigger problem is that education research is largely decoupled from practice: most educational researchers do not teach; most teachers do not learn about, or take part in, much research. (Compare this to medicine, in which practitioners — doctors — generally learn about research when they train and consult evidence-based guidelines when they practise, and might conduct research themselves.) However, this is not the same everywhere. Researching the effectiveness of lessons is an integral part of teachers' professional development in countries including China and Japan. Other nations should learn from this approach.

Beyond this, there is a growing body of evidence showing how best to get research insights into the classroom. A top-down approach that forces new methods on educators doesn't generally succeed. A better way, argues Rukmini Banerji, who leads Pratham, an educational non-governmental organization in New Delhi, is to encourage teachers and students to try out evidence-backed approaches for themselves.

Learn from disruption

In some cases, the disruption caused by COVID-19 introduced fresh ways of thinking and working in education — as difficult as those changes have been. Schools invented ways of delivering lessons digitally, teachers became more involved in children's social and emotional health, and parents became engaged in what their children were learning at home. Unfortunately, the impact of these innovations hasn't been well studied, because they happened so fast. Researchers and schools should make the most of the data they can collect and, where possible, gather more, so that they can hold on to

innovations that helped — both to assist children now and to strengthen education overall.

It's also important to track cohorts of children to expose the lasting impacts of missed schooling, as well as other consequences of the pandemic. And where innovations and catch-up programmes are in place, their impact should be measured with rigorous research, so that these data are available when learning is next upended.

Some have suggested that children might be able to bounce back quickly from COVID-related school closures by putting on a learning spurt. A more realistic view is that the better-off children will recover quickest, and the pandemic will amplify existing deep inequalities in education. That's why any efforts to help children today — and build the education systems of tomorrow — must focus on the most marginalized and disadvantaged children first.

Nature **605**, 588 (2022)

doi: <https://doi.org/10.1038/d41586-022-01429-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01429-0>

COVID lessons from Japan: the right messaging empowers citizens
[Download PDF](#)

- WORLD VIEW
- 23 May 2022

COVID lessons from Japan: the right messaging empowers citizens



There's no perfect solution to suppress the pandemic, but careful study and communication are key.

- [Hitoshi Oshitani](#) 0

Through six waves of COVID-19 in Japan, the number of cases and deaths per capita has been significantly lower than in other G7 countries. This is despite having the world's oldest population, and being densely packed. Yes, Japan has high vaccination rates, especially for older people, and masking is common. But neither of these is a full explanation. Deaths were low even before vaccines were available, and masks are common across Asia.

Japan has sought to understand the spread and risks of the disease and apply that to minimizing deaths and hospitalizations while maintaining social and economic activities. Trade-offs among these factors can be uneasy. Strong social pressure probably helped to boost protective measures, such as mask wearing, and minimized risky behaviours. Overall, the government quickly equipped its people with information to take protective action and avoided rigid prescriptions.

In 2003, I was the officer responsible for emerging diseases at the World Health Organization (WHO) Western Pacific regional office when the outbreak of severe acute respiratory syndrome (SARS) occurred: it was contained within eight months, with fewer than 1,000 deaths. When I first learnt of a similar coronavirus identified in China in people with pneumonia — SARS-CoV-2 — I thought perhaps the outbreak would follow a similar path.

I soon realized otherwise. With SARS, most people became severely ill. With COVID-19, many cases are mild or asymptomatic — and, unlike SARS, people can spread the disease without being sick. In other words, COVID-19 is far less ‘visible’ and so is harder to contain.

Japan’s constitution prohibits strict lockdowns, so another strategy was needed to suppress transmission. Heading into the pandemic, Japan had more than 8,000 public-health nurses across 400 public-health centres conducting ‘retrospective’ contact tracing for diseases, such as tuberculosis, to identify how people became infected —and that system was quickly adapted to COVID-19.

By the end of February 2020, scientists had identified many clusters of transmission and realized that most infected people did not infect anyone else, but a few infected many. From my past work, I knew that respiratory viruses are mainly transmitted through aerosols. My colleagues and I looked for common risk factors among superspreading events to come up with a more effective public-health message for the public. It incorporated early indications that SARS-CoV-2 could spread through aerosols.

This led us to warn against the ‘3Cs’ (*sanmitsu*): closed environments, crowded conditions and close-contact settings. Even as other countries

focused on disinfection, Japan promoted this concept extensively, by asking people to avoid high-risk activities such as karaoke bars, nightclubs and indoor dining. People largely complied. A panel of artists, academics and journalists named *sanmitsu* Japan's buzzword of the year in 2020.

Since the beginning of the pandemic, we've tracked how superspreading events differ. Other parts of the world have continued to flirt with 'going back to normal' by totally lifting restrictions, often in service of the economy, only to see cases soar again, with significant numbers of deaths. Simple solutions that help only the privileged and immunocompetent individuals cannot be accepted as a 'new normal' while vulnerable people bear the brunt of such policies. Current data suggest that Japanese citizens are adapting. In late April and early May, Japan celebrated its Golden Week holidays. This year, there were almost no special restrictions on when restaurants had to close or whether they could serve alcohol. Crowds were up, but smaller than in the years before the pandemic, and precautions, such as finding ventilated spaces, were emphasized. In the earlier waves, people would relax as cases ebbed, prompting a subsequent wave. But behaviour after the surge earlier this year seems different, even with no restrictive measures in place.

The situation is becoming more complicated. People are reluctant to accept strict measures, even with the upsurge of cases, because vaccine coverage is high and Omicron fatality rates are lower. There are more interventions available, especially in a high-income country such as Japan: booster vaccinations, antivirals, better clinical care and public-health measures, such as CO₂ monitors to track ventilation in public buildings.

But there is no one silver bullet that can eliminate the virus. Certainly, Japan's response has not been perfect and has received criticism. It is true that the country's initial testing capacity was limited, but extensive testing is not enough to suppress transmission.

Scientists and government advisers have to grapple with the fact that we do not yet know the right balance in the long term. They must understand that the behaviour of both the virus and people is subject to change — and adjust recommendations as such changes unfold.

Often, phrases such as ‘exit strategy’ or ‘back to normal’ are used by people longing for the days when we lived without the threat of this virus. But we are nowhere near back to normal. Nations must continue to seek the best balance between suppressing transmission and maintaining social and economic activities. How? By using all the tools at hand as they apply to cultures, traditions, legal frameworks and existing practices, to minimize suffering across the globe.

Nature **605**, 589 (2022)

doi: <https://doi.org/10.1038/d41586-022-01385-9>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01385-9>

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

Here in Ukraine, science continues under air raids

[Download PDF](#)

- WORLD VIEW
- 24 May 2022

Here in Ukraine, science continues under air raids



Three months after the Russian invasion, our Kyiv office is open again.

- [Yuri Moroz](#) 0

My drive to the office on 24 February, the first day of the Russian invasion of Ukraine, started as normal. I'd visited my parents the evening before and had stayed over. We spent the morning discussing the telegram messages I'd received from a colleague detailing explosions in Bровары, an eastern suburb of Kyiv. The suburb was far enough away for me to risk going into the city.

I left my parents' house in Kozelets, a town to the north of the city, got into my car and drove the 40 minutes to eastern Kyiv. There wasn't the usual traffic, but there was some. A queue of people waited for a shuttle bus for their own commutes into the city. I approached the intersection close to a military base that I'd passed hundreds of times before. I hadn't considered that it would be a target.

I slowed to wait at the traffic light and idly looked around. A rocket howled out of the sky, a blast of red and orange. It hit 150 metres away. When the traffic light turned green, no car moved. We expected a second rocket. The light turned red and then green again, and this time we started moving.

When I arrived, I began planning the evacuation of our office.

Three months on, I'm back in the office and Kyiv is back to being relatively safe — at least compared with other parts of my country. I'm a chemist and the executive director of Chemspace, which is headquartered here. We're a market-place for small molecules and biomolecules, with the largest online catalogue in the world. Pharmaceutical researchers globally use our database to find what they need for their research.

We're headquartered in Kyiv because many of the world's chemical companies are also here. Many chemists probably still don't realize that suppliers such as Life Chemicals, OTAVA and Enamine are all based here. More than 50% of the world's research chemicals are produced in Ukraine, or at least, they were before the invasion. (Much of the other half is produced in Russia, an artefact of the growth of the chemical industry after the collapse of the Soviet Union.)

I now know from experience that, once war strikes, the first reaction of many is to make sure their family, friends and colleagues are safe. The night of 24 February was spent listening to the sounds of explosions in the middle distance. In the morning, my family and I packed our bags and drove to the western part of the country.

The 5-hour drive took 12, spent in traffic with other Ukrainians, all of us taking part in the same mass exodus. My wife eventually drove our family

car — once used to drop my kids off at school — to Apetlon, Austria, where they all are now.

I could not join them, even if I wanted to — the law says no men between the ages of 18 and 60 can leave Ukraine.

Now, on-the-ground fighting is largely restricted to the east, but the conflict has left wreckage. My grandparents' village was shelled by cluster bombs; one landed on their vegetable patch and squats there in the mud, unexploded, waiting for a defusing team, or for a storm or another bomb to trigger it.

When I returned to Kyiv after spending some weeks in the west of the country, the city was dark. During the pandemic, I knew the apartments were occupied. Now, I know they're empty.

I write this from my office — and sometimes from our air-raid shelter, a small room on the ground floor, tucked under the stairs with a kettle, tea, a store of bottled water and six chairs arranged around a desk.

Today, our office is getting a little busier, as more of us return to the city. There are normally 7 or 8 of us in the office; we used to have 35 people here. Team meetings are still regularly disrupted by the sirens, and we've all become used to the sometimes far, sometimes near, sound of explosions. But we know this is good news — it means that our air defence systems are working as they should. When the sirens start, we unplug our laptops, take our personal documents and plod down the stairs.

Even work travel has resumed. Our business-development team attended Swiss Biotech Day in Basel on 2 May and is planning to attend a few more events in May and June. Sometimes people ask me why I'm working, not fighting. Sometimes I ask myself that, too. We feel it's important for us to do our part, as providers of crucial materials for scientists, but also as Ukrainian citizens.

Although we're not fighting, many of us are dying: the United Nations reports that there have been more than 3,500 confirmed civilian deaths

because of the war. The International Criminal Court has opened an investigation into war crimes and crimes against humanity committed here.

As long as our business continues to function, we can continue to pay our staff and our taxes, which support this country's military. Even as the war continues, Ukraine must keep working as a country if there is any hope of an end.

Today, I sent my wife and children a care package: clothes, shoes and swimming goggles for the summer. Ivan's goggles are red; Maria's are yellow. They left their country huddled in winter clothes.

I didn't often drive my kids to school before the invasion, but I hope we'll all be in the family car, on the way to school in Kyiv, come September.

Nature **605**, 590 (2022)

doi: <https://doi.org/10.1038/d41586-022-01431-6>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01431-6>

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

News in Focus

- **[Dark-matter-free galaxies, alarming tree deaths and the dawn of farming](#)** [25 May 2022]
News Round-Up • The latest science news, in brief.
- **[First pig kidneys transplanted into people: what scientists think](#)** [19 May 2022]
News • The genetically modified organs seemed to function for more than two days but some researchers are sceptical that the experiments had value.
- **[Home testing for syphilis gains support in wake of COVID](#)** [19 May 2022]
News • With cases of sexually transmitted infections growing, researchers hope self-tests — made popular by the pandemic — could stem the tide.
- **[Will the FDA change how it vets drugs following the Alzheimer's debacle?](#)** [13 May 2022]
News • The accelerated approval of aducanumab has triggered US lawmakers to push for more oversight from the agency.
- **[Flu vaccine could cut COVID risk](#)** [16 May 2022]
News • Health-care workers who got the influenza vaccine were also protected from COVID-19 — but the effect might not last long.
- **[Ancient tooth suggests Denisovans ventured far beyond Siberia](#)** [17 May 2022]
News • Molar found in Laos could be the first fossil evidence that the hominin species was far-ranging and able to adapt to different climates.
- **[Mammoth UK research assessment concludes as leaders eye radical shake up](#)** [12 May 2022]
News • Funding councils are considering changes to the Research Excellence Framework to improve research culture.
- **[How the revamped Large Hadron Collider will hunt for new physics](#)** [25 May 2022]
News Feature • The particle-smashing machine has fired up again — sparking fresh hope it can find unusual results.

- **COVID derailed learning for 1.6 billion students. Here's how schools can help them catch up** [25 May 2022]

News Feature • The pandemic is the largest disruption to education in history. But research has identified ways to help children make up lost ground. Will they work in classrooms around the world?

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

Dark-matter-free galaxies, alarming tree deaths and the dawn of farming
[Download PDF](#)

- NEWS ROUND-UP
- 25 May 2022

Dark-matter-free galaxies, alarming tree deaths and the dawn of farming

The latest science news, in brief.



This Hubble image captures a set of galaxies that are unusual because they seem not to have dark matter. Credit: NASA/ESA/P. van Dokkum, Yale Univ.

Galaxies without dark matter baffle astronomers

Scientists have long thought that galaxies cannot form without the gravitational pull of the mysterious material known as dark matter. But one group of astronomers thinks it might have observed a line of [11 galaxies that don't contain any of the substance](#), and could all have been created in an ancient collision ([P. van Dokkum *et al. Nature* 605, 435–439; 2022](#)).

This kind of system could be used to learn about how galaxies form, and about the nature of dark matter itself. However, some researchers are not convinced that the claim is much more than a hypothesis.

The finding centres on two galaxies, called DF2 and DF4, that were described in 2018 and 2019. Their stars moved so slowly that the pull of dark matter was not needed to explain their orbits, so the team concluded that the galaxies contained no dark matter.

In the latest research, scientists identified between three and seven new candidates for dark-matter-free galaxies in a line between DF2 and DF4, as well as strange, faint galaxies at either end.

“If proven right, this could certainly be exciting for galaxy formation. However, the jury is still out,” says Chervin Laporte, an astronomer at the University of Barcelona in Spain.

Northern Australian tree deaths double in 35 years

The rate at which trees are dying in the old-growth tropical forests of northern Australia each year has doubled since the 1980s, and [researchers say climate change is probably to blame](#).

The findings, published in *Nature* on 18 May, come from an extraordinary record of tree deaths catalogued at 24 sites in the tropical forests of northern Queensland over the past 49 years (D. Bauman *et al. Nature* <https://doi.org/hv67>; 2022).

The research team recorded that 2,305 trees across 81 key species had died since 1971. But from the mid-1980s, tree mortality risk increased from an average of 1% a year to 2% a year (see ‘Increasing death rate’). Of the 81 tree species that the team studied, 70% showed an increase in mortality risk over the study period.

The study found that the rise in death rate occurred at the same time as a long-term trend of increases in the atmospheric vapour pressure deficit, which is the difference between the amount of water vapour that the atmosphere can hold and the amount of water it does hold at a given time. The higher the deficit, the more water trees lose through their leaves, which can lead to sustained stress and eventually tree death.



Europe’s first farming populations descend mostly from farmers in the Anatolian peninsula, in what is now Turkey. Credit: Fatih Kurt/Anadolu Agency/Getty

Ancient DNA maps ‘dawn of farming’

Sometime before 12,000 years ago, nomadic hunter-gatherers in the Middle East made one of the most important transitions in human history: they began staying put and took to farming.

Two ancient-DNA studies have now [homed in on the identity of the hunter-gatherers who settled down](#).

Researchers sequenced the genomes of 15 hunter-gatherers and early farmers who lived in southwest Asia and Europe, along a key migration routes into Europe — the Danube River (N. Marchi *et al.* *Cell* <https://doi.org/gp49rr>; 2022).

The team found that ancient farmers in Anatolia — now Turkey — descended from repeated mixing between distinct hunter-gatherer groups from Europe and the Middle East. These groups first split at the height of the last Ice Age, some 25,000 years ago. Modelling suggests that the western groups nearly died out, before rebounding as the climate warmed.

Once established in Anatolia, the researchers found, early farmers moved west into Europe in a stepping-stone-like way, beginning around 8,000 years ago. They mixed occasionally — but not extensively — with local hunter-gatherers.

The findings chime with those of a similar ancient-genomics study posted on the bioRxiv preprint server this month (M. E. Allentoft. *et al.* Preprint at bioRxiv <https://doi.org/hv7g>; 2022).

Nature **605**, 595 (2022)

doi: <https://doi.org/10.1038/d41586-022-01386-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01386-8>

First pig kidneys transplanted into people: what scientists think
[Download PDF](#)

- NEWS
- 19 May 2022

First pig kidneys transplanted into people: what scientists think

The genetically modified organs seemed to function for more than two days but some researchers are sceptical that the experiments had value.

- [Sara Reardon](#)



A kidney is removed from a donor, in a standard transplant procedure. Credit: BSIP/Universal Images Group/Getty

Kidneys from pigs that had been genetically modified to have human-like immune systems worked successfully when transplanted into two patients who had recently died, the team that performed the experiments has reported¹. Although the organs seemed to function, some researchers question the value of the experiments and argue that clinical trials in living people are the only way to find out whether transplants from pigs can help to alleviate the shortage of human organs available to people who need them.

Researchers have transplanted pig organs into non-human primates with great success: one baboon lived for more than two years with a genetically modified pig heart². But baboons' immune and metabolic systems are different from humans', and certain immunosuppressive drugs routinely used in human organ transplants don't work in non-human primates, says Robert Montgomery, a transplant surgeon at New York University (NYU) in New York City who led the experiments. People who have recently died, he says, are "the closest thing we're going to get to a living human without the risk of harm".

Organ pair

In their transplant tests, which they performed in September and November 2021, Montgomery and his colleagues used pigs that had been genetically engineered to lack a gene for a protein called alpha-1,3-galactosyltransferase (α Gal). The pig version of α Gal triggers the human immune system to reject xenotransplants (organs transferred from a different species). With each kidney that the researchers transplanted, they also transplanted a pig thymus, an organ that produces immune cells and helps the body accept the foreign organs.

They tested these "thymokidneys" in two people who had been declared legally dead one to two days earlier because they did not have brain function. The researchers did not remove the patients' own kidneys, but grafted the pig kidneys onto the veins and arteries that carried blood to and from the recipients' legs. They then monitored the kidneys' function and the patients' immune responses for 54 hours — a limit imposed by NYU's ethics board and based on the amount of time typically required to harvest a person's organs for transplantation.



Pig organs are typically genetically modified before being transplanted into humans or non-human primates. Credit: Getty

In a paper published on 19 May in *The New England Journal of Medicine*, the researchers said that the patients showed no immediate immune reaction to the organs¹. Montgomery calls this “very reassuring”, although he says that a later immune reaction might have arisen if the patients had been kept on life support for months. The amount of urine the recipients produced increased after the transplant and the amount of creatinine — a bodily waste product — decreased, suggesting that the pig kidneys were working as intended.

Genetic modifications

Other researchers, however, are sceptical of the results. For one thing, the pigs used in the study were missing only one gene, despite research showing that modifying three or more genes helps the human immune system to accept the organ³. “It’s a pig that’s not relevant to what we need to know,” says David Cooper, a transplant surgeon at Massachusetts General Hospital

in Boston. Montgomery says that some of the immune response could be controlled using drugs.

Moreover, it's difficult to tell whether the pig kidneys were functioning or whether the urine and creatinine were actually coming from the patients' own kidneys. "You can't interpret the results," says Paige Porrett, a transplant surgeon at the University of Alabama at Birmingham. In January, she and her team published results from their own study⁴ in which they removed both kidneys from a person who had been without discernible brain function for five days. After testing whether the patient's antibodies would attack the pig organ, they transplanted two kidneys from a pig with 10 genetic modifications and monitored the deceased patient for 74 hours.

Like Montgomery's team, Porrett's team saw little immune reaction against the organs. But although the pig kidneys produced some urine, they did not process creatinine, suggesting that they weren't functioning properly. Porrett suspects the reason for this was that the patient's metabolic systems were shutting down because he had been dead for five days. She says that her group has transplanted pig kidneys into several more people who had been declared brain dead, and is planning to publish the results soon.

Ethics questions

Cooper is not surprised by either group's findings: research in baboons and human serum had already shown⁵ that primate immune systems won't immediately reject a pig organ that lacks αGal. The unanswered questions, he says, are whether the human immune system will attack the organ months later and whether the organ will continue to function. "I personally don't think we'll get definitive data from doing studies in brain-dead patients," he says.

Montgomery and Porrett say that it might be medically possible to extend the experiment, because some patients can survive for months after being declared brain dead. For instance, physicians have occasionally kept pregnant women on life support to allow the fetus to finish developing.

But doing this for research purposes creates ethical problems, says Rebecca Pentz, a bioethicist at Emory University in Atlanta, Georgia, because patients' bodies should be released to their families as soon as possible. According to a set of guidelines⁶ that she authored in 2005, researchers should maintain people who have irreversible loss of brain function on life support for only one day, unless there are valid scientific reasons to extend the period.

Still, Pentz says that the NYU study followed the regulations set by the researchers' ethics board. "It's a smart use of the newly dead," she says. "I do think xenotransplantation — if we can work out the science — is going to be an ethical advance because we can save more human lives."

Heart transplant

Partly because of such limitations, Cooper and others argue that it's time to start transplanting animal organs into living people — something that has been only partially successful. In January, researchers at the University of Maryland got special permission from the US Food and Drug Administration (FDA) to perform the first emergency transplant of a genetically modified pig heart into a man who was certain to die otherwise.

The man died two months later. In a 20 April webinar hosted by the American Society of Transplantation, the researchers behind the study said that they thought he had died from porcine cytomegalovirus — an animal virus that is harmless to humans but that caused his immune system to reject the organ.

Porrett points out that the infection hadn't been detected in the living pig, and that it remains unknown whether other latent animal viruses might start affecting humans long after a transplant has taken place. "We will not be able to answer that until we have patients we can follow for months or years," she says.

Porrett and Cooper's groups are applying to the FDA to start small clinical trials that would transplant genetically modified pig kidneys into humans. The kidney is the ideal organ to start with, Cooper says, because, unlike a

heart, it can be removed if problems arise and the patient can be placed on dialysis. “It should be done cautiously,” he says.

Nature **605**, 597–598 (2022)

doi: <https://doi.org/10.1038/d41586-022-01418-3>

References

1. Montgomery, R. A. *et al.* *N. Engl. J. Med.* **386**, 1889–1898 (2022).
2. Mohiuddin, M. *et al.* *Nature Commun.* **7**, 11138 (2016).
3. Estrada, J. L. *et al.* *Xenotransplantation* **22**, 194–202 (2015).
4. Porrett, P. M. *et al.* *Am. J. Transplant.* <https://doi.org/10.1111/ajt.16930> (2022).
5. Kuwaki, K. *et al.* *Nature Med.* **11**, 29–31 (2005).
6. Consensus Panel on Research with the Recently Dead. *Nature Med.* **11**, 1145–1149 (2005).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01418-3>

Home testing for syphilis gains support in wake of COVID
[Download PDF](#)

- NEWS
- 19 May 2022

Home testing for syphilis gains support in wake of COVID

With cases of sexually transmitted infections growing, researchers hope self-tests — made popular by the pandemic — could stem the tide.

- [Carrie Arnold](#)



Some at-home syphilis tests also detect multiple other infections, such as HIV and hepatitis C. Credit: Pixel-shot/Alamy

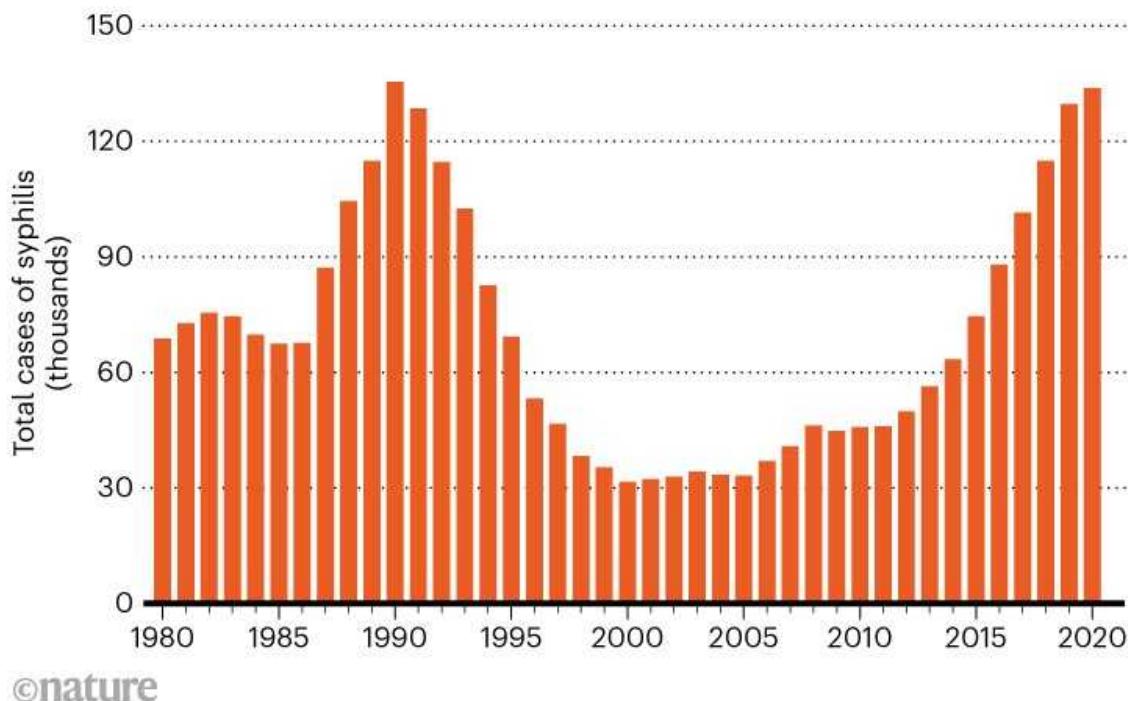
Global cases of sexually transmitted infections (STIs) have been on the rise in some countries. Even the COVID-19 pandemic — which locked down life in many ways — hasn't halted the trend. In April, the US Centers for Disease Control and Prevention [reported that](#) the first year of the pandemic saw 133,945 cases of syphilis, a 52% increase since 2016 (see 'Resurgence').

And this is probably an underestimate, the CDC says, given that health-care clinics had to limit in-person visits at the start of the pandemic, and STI surveillance programmes found their resources shifted elsewhere. The situation sparked a push for at-home tests for syphilis and other STIs.

Encouraged by the popularity of at-home tests for COVID-19, self-testing has been embraced by many, including policymakers and people in marginalized populations. In January, a California law went into effect that requires private health insurers to cover the cost of at-home STI testing. Supporters say that self-testing has the benefit of allowing individuals to collect samples in the privacy of their homes, free from any stigma that might be associated with attending clinics, and that it allows people from all backgrounds and income levels to test frequently.

RESURGENCE

After two decades of progress on reducing and then holding down US cases of syphilis, infections have climbed back to alarming levels. Even in 2020, the first year of the COVID-19 pandemic, when people were relatively isolated, case numbers rose. The reasons are complex, including funding cuts for sexual-health clinics and a rise in substance use. Other countries are also seeing increases in syphilis levels.



©nature

Source: US Centers for Disease Control and Prevention.

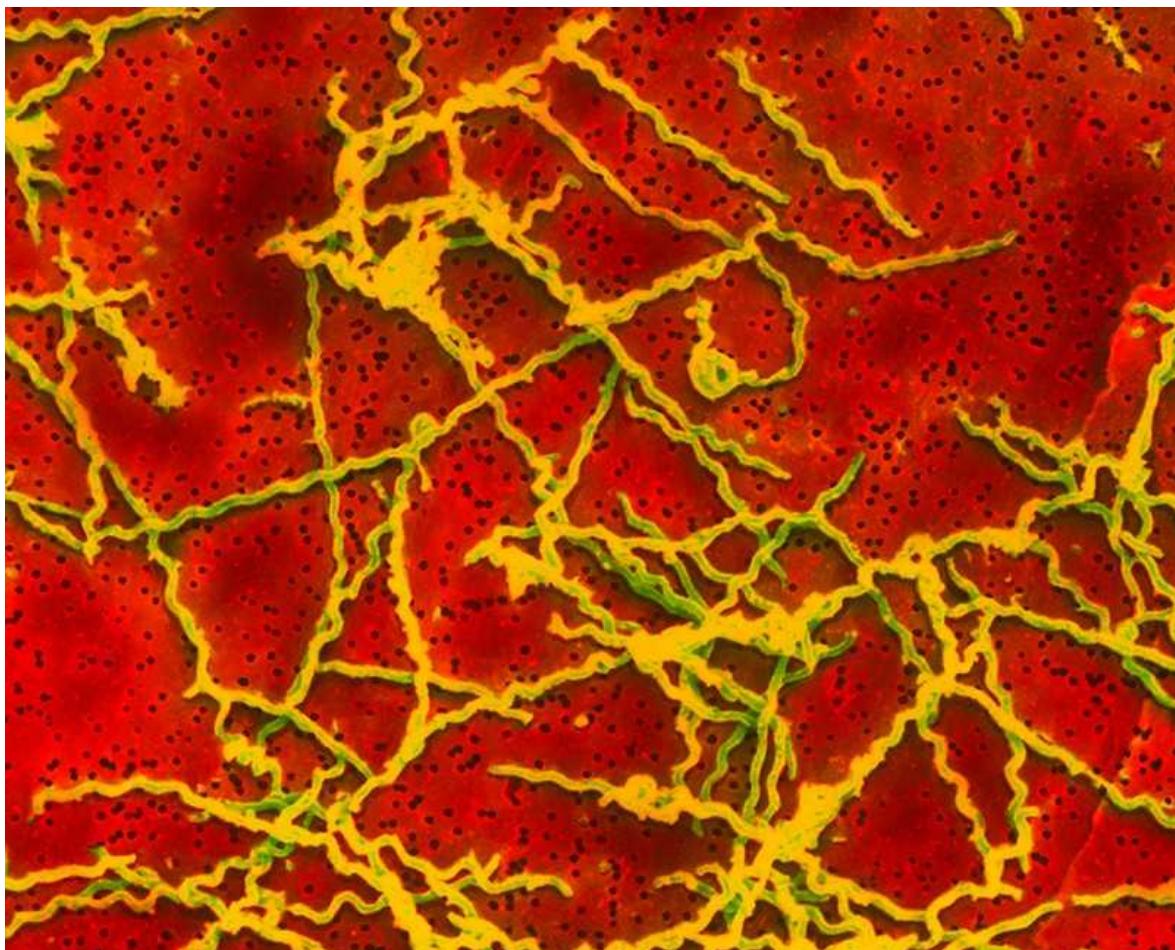
But the shift could also have trade-offs — some of them similar to those experienced for COVID-19, cautions Shweta Patel, a gynaecologist at the University of Alabama at Birmingham. With at-home tests, people do not receive the counselling that comes with in-person testing, and public-health departments might lose valuable statistics. Users must report their own results, and perform their own contact tracing to inform others that they might be in need of testing, Patel says, and this doesn't always happen.

Still, during the pandemic, self-testing for COVID-19 proved to be useful, says Natalie Cramer, deputy executive director of programmes at the National Alliance of State and Territorial AIDS Directors in Washington DC. She adds that it's past time for a similar approach to be taken for STIs including syphilis and HIV. "I think we've all become more comfortable

self-testing with COVID,” Cramer says. “It’s opened up both a need and a comfort level that folks have around it.”

Anatomy of a pathogen

For around half a millennium, syphilis has been one of the most widely recognized — and feared — STIs. It spurred the development of the first drug dubbed a magic bullet (salvarsan, an arsenic-based compound, in 1910) and, since the advent of penicillin three decades later, has largely been considered curable. With proper condom use, disease transmission can be reduced to near zero. Left untreated, however, the disease can prove deadly, causing severe neurological issues, especially if transmitted at or before birth.



Many syphilis tests detect antibodies against the bacterial family that includes the corkscrew-shaped *Treponema pallidum*. Credit: Dr Kari Lounatmaa/Science Photo Library

Detecting the disease, which is caused by the corkscrew-shaped bacterium *Treponema pallidum*, is easy enough, says Deborah Williamson, a microbiologist at the University of Melbourne in Australia, who has spent several decades studying syphilis. At a clinic, a health-care worker usually takes a blood sample and runs an assay to look for antibodies against the bacterial family that includes *T. pallidum*. If this assay is positive, the clinician runs further tests on the original sample to zero in on a diagnosis.

Only a small subset of tests, which require swabs of open sores or other regions of the body, actually look for the bacterium itself, whether searching for it under a microscope or identifying bits of its DNA or protein. Regardless of the method, however, if a person tests positive, treatment requires a physician's prescription, Patel says.

Testing goes DIY

When the COVID-19 pandemic arrived, clinic-based testing around the world took a hit. For instance, as rates of SARS-CoV-2 infection soared in New York City in the second quarter of 2020, [STI testing sharply declined](#). Clinics restricted visits, and people hesitated to see physicians in spaces where they might contract COVID-19. STI transmission began to rise. Although necessary, these changes "created a perfect storm", says Jennifer Mahn, director of clinical and sexual health at the National Coalition of STD Directors in Washington DC.

But the pandemic also spawned a boom in at-home testing. Governments began promoting at-home COVID-19 testing as a way for people to check whether they might be contagious before travelling or socializing. A variety of studies [showed that](#) frequent self-testing, even with rapid antigen tests that could detect infections only when viral load was high, could aid the public-health response^{1,2}. The move also made physicians, researchers and the public more comfortable about the accuracy and utility of home-testing kits in general.

Proponents hope that the popularity of at-home COVID-19 testing will boost the use of and demand for at-home STI tests. Research shows that these could be popular. For instance, a study of men who have sex with men in Zimbabwe, first published in April 2021³, showed that the participants were eager to self-test for syphilis.

At-home kits would go a long way towards helping these men access testing services, says study co-author Definate Nhamo, senior programmes manager at Pangaea Zimbabwe AIDS Trust in Harare. “They can get a result within 20–30 minutes and they do not need to endure some of the health-care providers’ negative attitudes.”

Test accuracy

In 2012, the US Food and Drug Administration approved its first at-home test for HIV, which detects antibodies against HIV-1 and HIV-2 in saliva. Users could get results in minutes, without having to send samples to a laboratory. The test was not as popular as first expected, and its price was high at US\$40–\$50. However, it inspired similar initiatives for other STIs: syphilis self-tests began appearing in the late 2010s.

Like their clinic-based counterparts, at-home syphilis tests detect antibodies against the family of bacteria that includes *T. pallidum* in a blood sample. But they require only a drop of blood, rather than a tubeful. They function like a hybrid between a blood-glucose check and a home pregnancy test.

Other self-administered syphilis tests ask users to take separate swabs of their mouth, anus and genitals, and send the samples to a clinical lab for analysis. These kits directly test for pieces of bacterial protein from *T. pallidum*, or for substances secreted by cells during syphilis infection. When used correctly, some tests can have sensitivities (true positive rates) and specificities (true negative rates) comparable to those of clinic-based tests — above 98%. Results, however, can take several days to receive, which makes them less useful for testing before every sexual encounter, says Weiming Tang, an infectious-disease epidemiologist at the University of North Carolina at Chapel Hill.

One drawback is that kits are often pricey, starting at \$50, with some options costing over \$100 — although some include tests for multiple STIs. Another drawback for the at-home tests that provide results within minutes is that clinical trials have found variabilities across different lots and users.

Self-test challenges remain

The push for syphilis self-testing is also creating public-health challenges that echo those from the COVID-19 pandemic. The accuracy of any at-home test depends on how well the user can collect the sample, Cramer says, whether it's a nasopharyngeal swab for SARS-CoV-2 or a genital swab for syphilis. Perhaps most concerning to some is that home tests lack the counselling and contact tracing that come with clinic-based tests, especially for STIs such as syphilis⁴.

Cramer says that although the results of self-testing aren't automatically uploaded to public-health databases, that might not cause such a big drop in reporting as expected. Many people whom she has seen use at-home tests wouldn't go to a clinic for testing anyway, so there would be no net loss in data where they are concerned. Furthermore, people still need to engage with the health-care system to receive treatment, which provides another opportunity to track infections and trace contacts. Some self-testing programmes send antibiotic prescriptions to people who test positive, without the need for a formal clinic visit. Overall, society is better off with at-home tests than without them, says Jen Hecht, executive director at Springboard HealthLab in Richmond, California.

Initiatives to promote the use of at-home syphilis test kits have picked up steam during the pandemic. In the United States, Hecht is trialling a programme to send free STI kits to people who sign up. Importantly, the programme also provides phone counselling and connects users with local public-health departments for follow-up and contact tracing.

“We’re seeing quite a number of positive cases, which is indicating to us that we’re reaching a group of people who really needed this kind of service,” Hecht says. “We want to be able to make it as easy as possible for folks to access.”

Nature **605**, 598–599 (2022)

doi: <https://doi.org/10.1038/d41586-022-01363-1>

References

1. Chin, E. T. *et al.* *Clin. Infect. Dis.* **73**, e3127–e3129 (2021).
2. Rader, B. *et al.* *Morb. Mortal. Wkly Rep.* **71**, 489–494 (2022).
3. Sri-Pathmanathan, C. *et al.* *Sex. Transm. Infect.* **98**, 197–202 (2022).
4. Vaulet, L. G. *et al.* *Sci. Rep.* **8**, 7542 (2018).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01363-1>

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

Will the FDA change how it vets drugs following the Alzheimer's debacle?
[Download PDF](#)

- NEWS
- 13 May 2022

Will the FDA change how it vets drugs following the Alzheimer's debacle?

The accelerated approval of aducanumab has triggered US lawmakers to push for more oversight from the agency.

- [Max Kozlov](#)



FDA commissioner Robert Califf has pledged to reform the agency's accelerated drug-approval programme. Credit: Bill Clark/CQ-Roll Call, Inc via Getty

Nearly a year after the US Food and Drug Administration (FDA) gave the green light to a controversial [drug to treat Alzheimer's disease](#), lawmakers are attempting to amend the process that led to its approval.

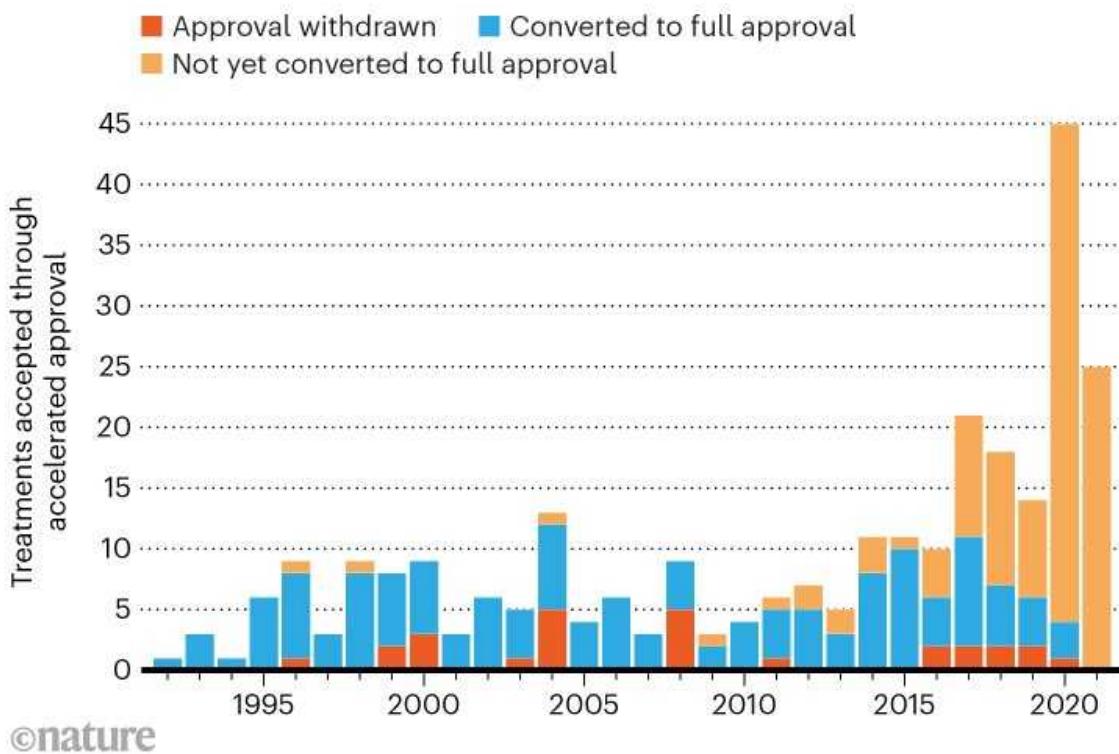
The House Committee on Energy and Commerce, which oversees drug safety and biomedical research, announced last week that it hopes to grant the FDA greater authority to rescind accelerated approvals if a company fails to complete follow-up studies on the treatment in a reasonable amount of time.

The provision, which was introduced as part of an FDA funding reauthorization bill, likely to be passed before September, comes on the heels of the agency's 2021 approval of aducanumab, an antibody drug shown to reduce the accumulation of plaques in the brain associated with the progression of Alzheimer's. Despite a nearly unanimous vote against the approval by an independent panel of experts, the agency fast-tracked the drug, which was developed by Biogen, a biotechnology company based in Cambridge, Massachusetts. Three advisory-panel members resigned in protest against the decision, and the approval is the subject of multiple investigations by federal regulators.

Aducanumab is not the only reason that this drug-approval pathway is coming under fire: since its inception, the programme has led to 279 treatments reaching the market, with nearly two-thirds in the past decade alone (see 'Growing momentum for accelerated approval'). The programme's increasing popularity signals a shift away from its original intent, says Diana Zuckerman, president of the National Center for Health Research, a non-profit organization in Washington DC. "Accelerated approval started out as a special programme for a small number of drugs, and now most cancer drugs are going through accelerated or some other expedited pathway," she says.

GROWING MOMENTUM FOR ACCELERATED APPROVAL

Nearly 300 treatments have been approved through the US Food and Drug Administration's accelerated-approval programme. With additional data, these can go on to receive full approval. Some approvals have been withdrawn because of a lack of follow-up data, but full approval is still pending for many.



Source: FDA

Companies, moreover, have been slow to produce the follow-up studies promised as part of the approval process. The FDA has limited power to compel them to provide the data, but the legislative proposal — which could still change significantly as it wends its way through the House of Representatives and the Senate — could grant it more authority to do so.

Days before his appointment in February, [FDA commissioner Robert Califf](#) pledged to make accelerated-approval reform a priority for the agency. Researchers who spoke to *Nature* agree that reforms are needed to protect the integrity of the programme, and that the proposed legislation is a good start. But they also recommended more agency oversight and other changes that would further prevent pharmaceutical firms from abusing this route to the market.

“Instead of the drug companies living up to and working to ensure that they are employing the accelerated-approval pathway as intended, we have too many that are willing to take advantage of the loopholes where they can find them,” says David Mitchell, president of Patients for Affordable Drugs, a non-profit organization in Washington DC, who serves as a consumer representative on the independent panel that reviews cancer drugs for the FDA.

The need for speed

The FDA created the accelerated-approval pathway in 1992, largely in response to the HIV–AIDS crisis, to get urgently needed drugs to the market without delay. Instead of demonstrating efficacy through clinically-meaningful endpoints, such as patient survival or reduction of symptoms, drug candidates reviewed under this pathway often rely on what are known as surrogate endpoints, which may be faster or easier to track than conventional clinical-trial endpoints. For example, tumour shrinkage is a common surrogate used in cancer-drug clinical trials, but this metric is not necessarily linked to a direct benefit to patients.

Gregg Gonsalves, an epidemiologist and global health specialist at Yale University in New Haven, Connecticut, was among the group that persuaded the FDA to adopt this programme. “We pushed for this accelerated approval pathway because people were dying,” he says. “I’m HIV positive, so I get the desperation and need for hope.”

The pathway has turbocharged the number of immunotherapies and cancer treatments on the market. But some of these drugs cost hundreds of thousands of dollars per year, despite, in many cases, limited data showing their clinical utility. Gonsalves argues that the programme has been co-opted by the pharmaceutical industry to speed approvals. Cancer treatments approved through the pathway have made it to market on average about three years earlier than they would through standard routes. And a single study using surrogate endpoints could be enough to get a treatment on the market.

Part of the problem, says Caleb Alexander, an internal-medicine specialist and epidemiologist at the Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland, is that drug companies aren't upholding their end of the bargain with timely post-market studies confirming the benefits of the drug. Some researchers question whether companies are given too much time to produce such data. A [2021 analysis](#) found that 13% of drugs granted accelerated approval between 1992 and 2016 hadn't been converted to full approval within five years — and remained on the market for a median of 9.5 years without the data needed for conversion.



Clinical trials on the Alzheimer's treatment aducanumab are continuing after the drug's fast-track approval. Credit: Charles Krupa/AP/Shutterstock

The FDA granted Biogen nine years to complete its confirmatory trial on aducanumab — a timeline that Alexander calls “frankly offensive”. Biogen has since said it will complete the trial in four years, and a spokesperson for the company says that data from the trial — not yet published or peer reviewed — show a slowing in clinical decline for some people who are taking the drug.

Post-market trials can take a long time, especially for slowly-progressing conditions [such as neurodegenerative diseases](#), says a spokesperson for the Rare Disease Company Coalition, an organization in Washington DC that represents 21 pharmaceutical firms.

It is also difficult for companies to recruit participants, because people would much rather be guaranteed an approved medicine than risk getting a placebo. Instead of demanding that a company stop selling a drug that hasn't been converted to full approval, says Zuckerman, the agency often requests that the company voluntarily withdraw it from the market. "The FDA loses an enormous amount of leverage once a product is approved," says Alexander.

For example, in 2011, the FDA revoked its accelerated approval of [the antibody drug bevacizumab](#) to treat breast cancer, citing a lack of clinical benefit. This resulted in public backlash as people with breast cancer gave emotional testimonies in an attempt to keep the approval in place. For many people, accelerated approval offers a "valuable source of hope", noted the National Organization for Rare Disorders, a non-profit organization in Danbury, Connecticut, [in a 2021 report](#). But Mitchell, who has multiple myeloma, argues: "It is not the FDA's job to give me hope. Hope is not what keeps me alive. It's drugs that are safe and effective."

Jeremy Kahn, a spokesperson for the FDA, affirmed in an e-mail that the agency is committed to ensuring the integrity of the accelerated-approval programme and noted that the agency believes people who lack treatment options for serious diseases are willing to "accept some uncertainty" in clinical benefit for new treatments. Clinical benefit has been verified in the vast majority of accelerated approvals, he added.

Changing rules, same problems

The accelerated-approval programme has served as a model for other countries. But adopters including the European Union and Japan have given drug regulators more authority. Regulators can require companies to submit data from confirmatory trials in a set time period; if they do not, their approvals can be withdrawn.

How effective the proposed rule changes for the US FDA would be is unclear. Although they would make it easier for the agency to withdraw approval, they would also lengthen the bureaucratic process of rescinding approvals. This defangs the provision, Zuckerman says. She would have preferred to stick with an earlier proposal, which would have automatically revoked approvals once confirmatory trials were one year overdue.

Zuckerman also recommends that the FDA commissioner's office create a separate independent advisory group to review agency approvals that go against advisory panel recommendations — as happened for aducanumab. "The vast majority of advisory-committee votes recommend approval, so when they don't recommend approval, there's usually a really good reason," she says.

Alexander suggests using health-care coverage as leverage. The US Centers for Medicare & Medicaid Services (CMS) in Baltimore, for example, decides which treatments will be funded for tens of millions of US residents. Earlier this year, concerned about the efficacy of aducanumab, the CMS stated that it would cover the annual US\$28,800 cost of the drug only for people enrolled in clinical trials.

Although that decision is nearly unprecedented, Alexander thinks that the CMS should consider a lower reimbursement rate for other accelerated-approval treatments that have not yet gained full approval. Such a move could "light a fire underneath manufacturers" to complete their trials, he says. "Why should taxpayers be on the hook for paying the full price of a drug when we don't know the full scope of its safety and effectiveness?" he asks.

But Mitchell worries that cutting reimbursement would remove companies' incentive to produce potentially life-saving medication. He credits three accelerated-approval drugs with keeping him alive after his myeloma diagnosis.

Mitchell considers the aducanumab debacle a "blip" in the grand scheme of accelerated approvals, but says it has brought more attention to the need for confirmatory trials for some drugs. Reform won't be simple. Once a

medicine enters the market, Mitchell says, “drug companies aren’t anxious to find a reason to take it off”.

Still, many researchers and drug-safety advocates are eager to see change. “We started out trying to fix a pendulum that was too far in one direction,” says Zuckerman, “and look how far we’ve come in this direction now.”

Nature **605**, 600-601 (2022)

doi: <https://doi.org/10.1038/d41586-022-01316-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01316-8>

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

Flu vaccine could cut COVID risk

[Download PDF](#)

- NEWS
- 16 May 2022

Flu vaccine could cut COVID risk

Health-care workers who got the influenza vaccine were also protected from COVID-19 — but the effect might not last long.

- [Ewen Callaway](#)



People in Santiago are vaccinated against influenza. Credit: Ivan Alvarado/Reuters/Alamy

Influenza vaccines have a surprising health benefit: they might also prevent COVID-19, particularly in its most severe forms¹.

A study of more than 30,000 health-care workers in Qatar found that those who got a flu jab were nearly 90% less likely to develop [severe COVID-19](#) over the next few months, compared with those who hadn't been recently vaccinated against flu.

The study, which was conducted in late 2020, before [the roll-out of COVID-19 vaccines](#), is in line with previous work suggesting that ramping up the immune system using influenza vaccines and other jabs could help the body to fend off the coronavirus SARS-CoV-2.

Collateral benefit

In the early months of the pandemic — while COVID-19 vaccines were still in development — researchers were intensely interested in the possibility that existing vaccines might provide some protection against SARS-CoV-2. But collecting strong evidence for such an effect is difficult, because people who seek vaccination for diseases other than COVID-19 might also make other choices that reduce their risk of being infected with SARS-CoV-2.

To minimize the impact of this ‘healthy-user effect’, a team led by Laith Jamal Abu-Raddad, an infectious-disease epidemiologist at Weill Cornell Medicine–Qatar in Doha, analysed the health records of 30,774 medical workers in the country. There is probably less variation in health-related behaviour among such workers than in the general population, reducing — but probably not eliminating — bias, Abu-Raddad says.

The researchers tracked 518 workers who tested positive for SARS-CoV-2 and matched them to more than 2,000 study participants who had tested negative for the virus. Those who had received an influenza vaccine that season were 30% less likely to test positive for SARS-CoV-2, and 89% less likely to develop severe COVID-19, compared with workers who had not (although the number of severe cases was small in both groups). The study was posted on the medRxiv preprint server on 10 May.

Günther Fink, an epidemiologist at the University of Basel in Switzerland, says the Qatar analysis reduces the odds that other studies that uncovered the same link were a fluke. His team reported that flu vaccines were associated with a reduced risk of death in people hospitalized with COVID-19 in Brazil².

“This is an important piece of evidence,” says Mihai Netea, an infectious-disease specialist at Radboud University Medical Center in Nijmegen, the Netherlands. The observation that influenza vaccines are linked to a reduction in not just SARS-CoV-2 infections, but also disease severity, strongly suggests that the protection is genuine, he adds.

Time limit

How long this protection lasts is unclear. Among those in the Qatar study who had the flu jab and later contracted COVID-19, Abu-Raddad’s team recorded SARS-CoV-2 infections occurring, on average, about six weeks after vaccination. “I don’t expect to see this effect lasting long at all,” he says. Netea guesses that the benefits last for between six months and two years.

It’s not fully clear why flu vaccines — which are composed of killed influenza viruses — would also protect against COVID-19. Vaccines train the immune system to recognize specific pathogens, but they also rev up broad-acting antiviral defences, says Netea, who has found signs of such responses in flu-vaccine recipients³.

Netea’s team is also working to better quantify the benefits of vaccines targeting influenza and other diseases against COVID-19. To fully rule out healthy-user effects, his team has launched a randomized, placebo-controlled trial in Brazil that will test whether influenza and measles–mumps–rubella vaccines can protect against COVID-19.

Knowing that vaccines for flu and other diseases can offer protection against COVID-19, even if only partial and for a limited period, could limit the damage caused by a future pandemic before a vaccine for that disease is

developed, Netea argues. “If you have something in the beginning, you could save millions of lives.”

Nature **605**, 602 (2022)

doi: <https://doi.org/10.1038/d41586-022-01315-9>

References

1. Tayar, E. *et al.* Preprint at medRxiv
<https://doi.org/10.1101/2022.05.09.22274802> (2022).
2. Fink, G. *et al.* *BMJ Evid. Based Med.* **26**, 192–193 (2020).
3. Debisarun, P. A. *et al.* *PLOS Pathog.* **17**, e1009928 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01315-9>

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

Ancient tooth suggests Denisovans ventured far beyond Siberia
[Download PDF](#)

- NEWS
- 17 May 2022

Ancient tooth suggests Denisovans ventured far beyond Siberia

Molar found in Laos could be the first fossil evidence that the hominin species was far-ranging and able to adapt to different climates.

- [Freda Kreier](#)



The fossilized molar, seen here from several angles, is thought to have belonged to a young Denisovan girl that died between 164,000 and 131,000 years ago. Credit: F. Demeter

A fossilized tooth unearthed in a cave in northern Laos might have belonged to a young Denisovan girl that died between 164,000 and 131,000 years ago. If confirmed, it would be the first fossil evidence that Denisovans — an extinct hominin species that co-existed with Neanderthals and modern humans — lived in southeast Asia.

The molar, described in *Nature Communications* on 17 May¹, is only the second Denisovan fossil to be found outside Siberia. Its presence in Laos supports the idea that the species had a much broader geographic range than the fossil record previously indicated.

“We’ve always assumed that Denisovans were in this part of the world, but we’ve never had the physical evidence,” says study co-author Laura Shackelford, a palaeoanthropologist at the University of Illinois Urbana-Champaign. “This is one little piece of evidence that they were really there.”

Expanded range

Denisovans were first identified in 2010, when scientists sequenced DNA from a fingertip bone found in Denisova cave in Siberia, and showed that it belonged to a previously unknown species of ancient human². Subsequent genetic studies^{3,4} have revealed that millions of people from Asia, Oceania and the Pacific Islands carry traces of Denisovan DNA.

This suggests that the species ranged far beyond Siberia — but the fossil evidence has been sparse. The entire fossil record for Denisovans so far boils down to a handful of teeth, bone shards and a jawbone found in Tibet. Aside from the latter, every specimen (including a piece of bone that belonged to a [half-Denisovan girl](#) whose mother was a Neanderthal) has come from Denisova cave.

That’s partially because fossils have a better chance of surviving in cold, dry conditions than in warm, humid ones. But in 2018, Shackelford and her colleagues were looking for potential dig sites in northern Laos when they came across a cave “just filled with teeth”. These belonged to a mixture of species, including giant tapirs, deer, pigs and ancient relatives of modern elephants. The collection was probably amassed by porcupines collecting

bones to sharpen their teeth and extract nutrients, says Shackelford. Among the first batch of fossils to come out of the cave was a small, underdeveloped hominin tooth.



The tooth was found at this site in Laos. Credit: F. Demeter

Dating of the cave's rock and animal teeth revealed that the tooth pre-dated the arrival of modern humans in the area. "It was just a huge surprise," says Shackelford, who says the team wasn't expecting to find ancient-human remains. At first, the researchers thought the tooth might belong to *Homo erectus* — an ancient-human species that lived in Asia between around 2 million and 100,000 years ago. But the molar is "too complex" to belong to *H. erectus*, the researchers say, and although it shares some characteristics with Neanderthal teeth, it is also "large, and kind of weird", says Bence Viola, a palaeoanthropologist at the University of Toronto in Canada.

The molar has the greatest resemblance to teeth found in the Denisovan jawbone from Tibet. "Denisovans have absolutely gigantic teeth," Viola says. "So it seems like a good assumption that this is likely a Denisovan."

The tooth's roots are not fully developed, so it probably belonged to a child, the researchers say. They also found that it lacked certain peptides in its

enamel that are associated with the Y chromosome — a possible indication that its owner was female.

Right place, right time

Reconstructing the identity of a person whose bones have been degraded by thousands of years of tropical conditions is challenging, says Katerina Douka, an archaeological scientist at the University of Vienna. Without more fossils or DNA analysis, “the reality is that we cannot know whether this single and badly preserved molar belonged to a Denisovan”, she says.

But Viola says that the molar is in the “right place and right time” to belong to a Denisovan. If this were confirmed, it would reveal that the species was able to adapt to different environmental conditions. At the time the tooth’s owner died, more than 131,000 years ago, the area would have been lightly wooded and temperate — completely different from the frigid temperatures faced by Denisovans in Siberia and Tibet. The ability to live in a wide range of climates would set the Denisovans apart from Neanderthals — whose bodies were adapted for colder places — and make them more similar to our own species.

Even with the uncertainty, the discovery is likely to encourage other researchers to look for ancient-human fossils in southeast Asia, says Viola.

“When we started looking in Laos, everyone thought we were crazy,” says Shackelford. “But if we can find things like this tooth — which we weren’t even anticipating — then there are probably more hominin fossils to be found.”

Nature **605**, 602-603 (2022)

doi: <https://doi.org/10.1038/d41586-022-01372-0>

References

1. Demeter, F. *et al. Nature Commun.* **13**, 2557 (2022).

2. Krause, J. *et al.* *Nature* **464**, 894–897 (2010).
 3. Browning, S. R., Browning, B. L., Zhou, Y., Tucci, S. & Akey, J. M. *Cell* **173**, 53–61.E9 (2018).
 4. Vernot, B. *et al.* *Science* **352**, 235–239 (2016).
-

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01372-0>

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

Mammoth UK research assessment concludes as leaders eye radical shake up

[Download PDF](#)

- NEWS
- 12 May 2022

Mammoth UK research assessment concludes as leaders eye radical shake up

Funding councils are considering changes to the Research Excellence Framework to improve research culture.

- [Holly Else](#)



The Research Excellence Framework dictates how much government funding universities will get over the coming years. Credit: Lewis Houghton/Science Photo Library

Researchers across the United Kingdom are celebrating or commiserating this week as universities receive the results of a years-long research-assessment exercise that dictates how much government funding they will get over the coming years.

The Research Excellence Framework (REF) is one of the world's most comprehensive research assessments — many nations look to it as an example of how to review research. But its leaders are already looking at how they might change the process.

The 2021 REF results suggest that more than 40% of UK research is world-leading. In an [analysis of the results](#) by the *Times Higher Education* (THE), the country's biggest research universities rank among the highest-scoring institutions overall. The top five are all located in southern England (see 'Institution rankings top 10').

Almost all UK universities took part in the mammoth exercise, together submitting 185,000 pieces of research from more than 76,000 researchers for assessment.

The funders in charge of the REF are reviewing how the next assessment might change. Currently, research is given a star rating by 34 expert panels that fall into four broad categories: medicine, health and life sciences; physical sciences, engineering and mathematics; social sciences; and arts and humanities.

Reviewers also judge institutions on the impact of their work in the wider world and the standard of their research environment. Scores for each element carry a different weight in a formula that dictates the size of each institution's share of the multibillion-pound pot for public research funding. Currently, the score for research outputs holds the biggest sway, accounting for 60% of the final mark.

The UK government is yet to announce how much money will be up for grabs, and how it will be divided between different institutions. The results of the previous exercise, released in 2014, guided £14 billion (US\$17.1 billion) of university research funding.

The results “represent an exceptional achievement for UK university research and demonstrate the huge return on public investment in research”, says Steven Hill, director of research at the funding body Research England.

Changing priorities

Those who administer the research-assessment exercise — the UK’s higher education funding councils — have a history of changing the rules each time to reflect priorities and to help stop institutions from gaming the process to boost their scores. Ahead of the 2014 results, the exercise was broadened to include a measure of research impact. Researchers now submit case studies to demonstrate the economic, social and policy contributions of their work — a move that has been copied by other countries seeking to widen their approach to research assessment.

The latest change included a rule that institutions must submit for assessment the work of everyone who does research as part of their job. Previously, some institutions put forward only top performers in an attempt to skew ratings in their favour. As a result of the change, the latest exercise saw a 46% increase in the number of staff submitted for assessment compared with the previous one.

Institution rankings top 10

Ranking	Institution	THE grade-point average score
1	Imperial College London	3.63

Ranking	Institution	THE grade-point average score
2	Institute of Cancer Research	3.58
=3	University of Cambridge	3.53
=3	London School of Economics and Political Science	3.53
5	University of Bristol	3.51
6	University College London	3.5
7	University of Oxford	3.49
8	University of Manchester	3.47
9	King's College London	3.46
=10	University of York	3.45
=10	London School of Hygiene and Tropical Medicine	3.45

Source: *Times Higher Education*. Grade-point average score is [calculated](#) using REF quality classifications for each institution's research output, impact and environment.

Last year, funders began to look at how the process could be used to recognize and reward institutions fostering a positive research culture. So the rules of the next assessment, the date of which has yet to be confirmed, could look significantly different.

“The research system itself is under unsustainable pressure in terms of careers and livelihoods. If we want to improve research culture, then the REF is potentially a powerful ally in that effort because it gives you that reach across the entire system,” says James Wilsdon, a science-policy researcher at the University of Sheffield, UK. “I do think the stars are now aligning to support a more radical overhaul of the exercise than at any point in the past 20–25 years.”

Catriona Firth, the associate director for research environment at Research England — one of the four funding councils that administers the REF — agrees that a radical shake up could be on the cards. One aspect of the current review process, known as the Future Research Assessment Programme (FRAP), looks at how the framework can be used to recognize and reward positive research culture. “The current REF has been quite focused on the research end points, and not focusing so much on the inputs or the research process,” says Firth.

Because research outputs are assessed for their originality, institutions do not submit review articles, negative results or replication studies as part of the exercise, which are all important for research, she says. “There are wider consequences of focusing on excellence. What institutes think is going to be valuable in the REF is what they encourage staff to do and what they invest in.”

The tricky part, says Firth, will be balancing the continued drive for excellence with rewarding healthy research culture, without placing a disproportionate administrative burden on institutions. The REF already has many critics, who claim it is bureaucratic and expensive (the 2014 exercise cost £246 million to run).

The results of the REF review will be very important for the rest of the world, says Lidia Borrell-Damián, who is head of research and innovation at policy group Science Europe and sits on the FRAP international advisory board. “It will be full of insights on what to consider when reforming the assessment,” she adds.

Nature **605**, 603 (2022)

doi: <https://doi.org/10.1038/d41586-022-01310-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01310-0>

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

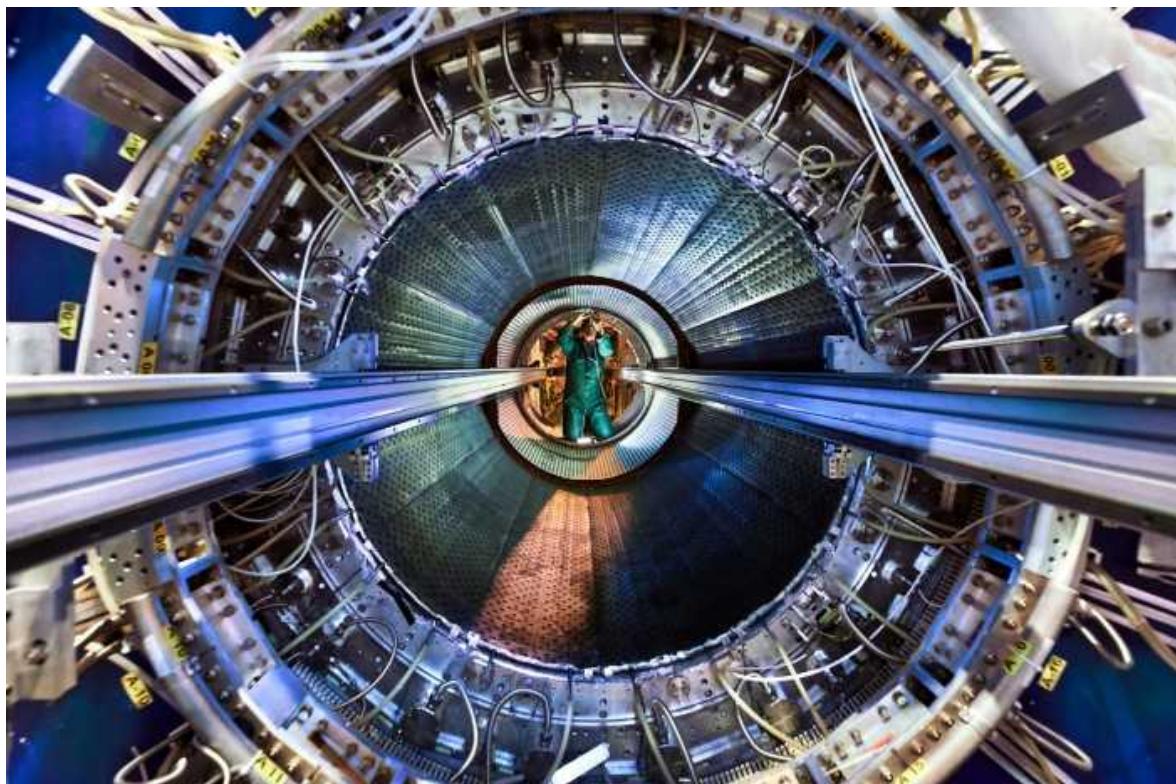
How the revamped Large Hadron Collider will hunt for new physics
[Download PDF](#)

- NEWS FEATURE
- 25 May 2022

How the revamped Large Hadron Collider will hunt for new physics

The particle-smashing machine has fired up again — sparking fresh hope it can find unusual results.

- [Elizabeth Gibney](#)



Detectors at the ALICE experiment were revamped during the Large Hadron Collider's 2018–22 shutdown. Credit: Maximilien Brice, Julien Marius

Ordan/CERN

The hunt for new physics is back on. The world's most powerful machine for smashing high-energy particles together, the Large Hadron Collider (LHC), has fired up after a shutdown of more than three years. Beams of protons are once again whizzing around its 27-kilometre loop at CERN, Europe's particle-physics laboratory near Geneva. By July, physicists will be able to switch on their experiments and watch bunches of particles collide.

In its first two stints, in 2009–13 and 2015–18, the LHC explored the known physics world. All of that work — including the triumphant 2012 discovery of the Higgs boson — reaffirmed physicists' current best description of the particles and forces that make up the Universe: the standard model. But scientists sifting through the detritus of quadrillions of high-energy collisions have yet to find proof of any surprising new particles or anything else completely unknown.

This time could be different. The LHC has so far cost US\$9.2 billion to build, including the latest upgrades: version three comes with more data, better detectors and innovative ways to search for new physics. What's more, scientists start with a tantalizing shopping list of anomalous results — many more than at the start of the last run — that hint at where to look for particles outside the standard model.

“We’re really starting with adrenaline up,” says Isabel Pedraza, a particle physicist at the Meritorious Autonomous University of Puebla (BUAP) in Mexico. “I’m sure we will see something in run 3.”

Higher energy and more data

After renovations to its particle accelerators, the third version of the LHC will collide protons at 13.6 trillion electron volts (TeV) — slightly higher than in run 2, which reached 13 TeV. The more-energetic smashes should increase the chances that collisions will create particles in high-energy regions where some theories suggest new physics could lie, says Rende Steerenberg, who leads beam operations at CERN. The machine’s beams will also deliver more-compact bunches of particles, increasing the

probability of collisions. This will allow the LHC to maintain its peak rate of collisions for longer, ultimately allowing experiments to record as many data as in the first two runs combined.

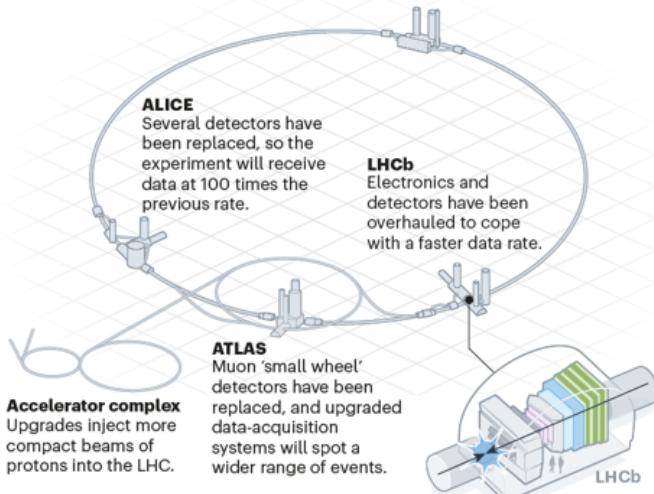
To deal with the flood, the machine's detectors — layers of sensors that capture particles that spray from collisions and measure their energy, momentum and other properties — have been upgraded to make them more efficient and precise (see 'Data boost').

DATA BOOST

In the third run of the Large Hadron Collider (LHC), researchers expect to record as many data on particle collisions as in the machine's first two runs combined. That's in part owing to more collisions, produced by stabler, more tightly packed streams of particles circulating around the ring. It is also because of upgrades to the machines four main detectors: CMS, ATLAS, ALICE and LHCb. The LHCb has been completely revamped.

CMS

Prototype detector layers will test the capture of muons, heavy cousins of electrons, at higher rates. Upgraded electronics will improve energy measurements for hadrons, particles made up of multiple quarks (such as protons and neutrons).



Better detectors

The LHCb experiment has been comprehensively upgraded for the LHC's third run. This top-down view shows how it detects particles spraying out of collisions, and how its electronics and detectors have been improved.

Pinpointing Bs

Millimetres from where collisions occur, short-lived B hadrons decay into other particles. A new 'vertex locator' will measure this point with greater precision.

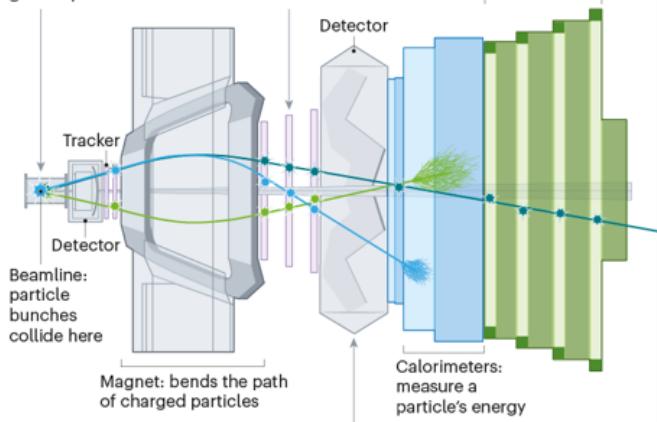
Tracking particles

'Trackers' trace particle paths. Two new detectors will better separate nearly identical paths and cut out the noise that mimics real tracks.

Charged hadron (such as a proton or pion)

- Green line
- Blue line
- Green dot

Muon detector



Electronics

Renovated electronics mean that LHCb can now use software to scan through 40 million events per second. Previously, a coarser hardware filter first triaged these to identify one million to be scanned.

Identifying particles

Other detectors measure the velocity of charged particles; combining this information with a particle's path reveals its identity. Upgraded detectors are more sensitive to velocity and can cope with higher data rates.

Nik Spencer/*Nature*; Source: CERN

A major challenge for LHC researchers has always been that so little of the collision data can be stored. The machine collides bunches 40 million times per second, and each proton–proton collision, or ‘event’, can spew out hundreds of particles. ‘Trigger’ systems must weed out the most interesting of these events and throw the bulk of the data away. For example, at CMS — one of the LHC’s four main experiments — a trigger built into the hardware makes a rough cut of around 100,000 events per second on the basis of assessments of properties such as the particles’ energies, before software picks out around 1,000 to reconstruct in full for analysis.

With more data, the trigger systems must triage even more events. One improvement comes from a trial of chips originally designed for video games, called GPUs (graphics processing units). These can reconstruct particle histories more quickly than conventional processors can, so the software will be able to scan faster and across more criteria each second. That will allow it to potentially spot strange collisions that might previously have been missed.

In particular, the LHCb experiment has revamped its detector electronics so that it will use only software to scan events for interesting physics. Improvements across the experiment mean that it should collect four times more data in run 3 than it did in run 2. It is “almost like a brand new detector”, says Yasmine Amhis, a physicist at the Laboratory of the Irène-Joliot Curie Physics of the Two Infinities Lab in Orsay, France, and member of the LHCb collaboration.



The LHCb's 'vertex locator', placed close to the LHC's beamline to see short-lived particles.Credit: Maximilien Brice, Julien Marius Ordan/CERN

Spotting anomalies

Run 3 will also give physicists more precision in their measurements of known particles, such as the Higgs boson, says Ludovico Pontecorvo, a physicist with the ATLAS experiment. This alone could produce results that conflict with known physics — for instance, when measuring it more precisely shrinks the error bars enough to put it outside the standard model's predictions.

But physicists also want to know whether a host of odd recent results are genuine anomalies, which might help to fill some gaps in understanding about the Universe. The standard model is incomplete: it cannot account for phenomena such as dark matter, for instance. And findings that jar with the

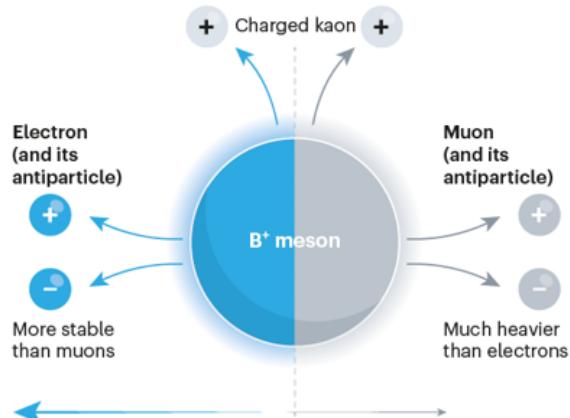
model — but are not firm enough to claim as a definite discrepancy — have popped up many times in the past two years (see ‘Hints of new physics?’).

HINTS OF NEW PHYSICS?

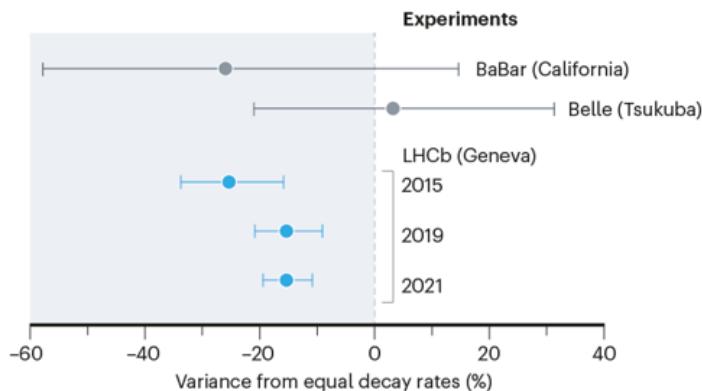
Some recent anomalies clash with physics' standard model — but they are not yet firm enough to claim as discoveries.

The B-meson anomaly

The B^+ meson — a transient particle — can decay in two ways that should be equally rare.

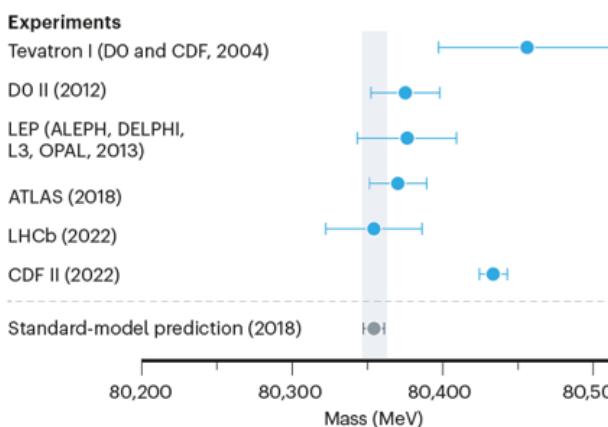


The LHCb detector has seen the electron decay pathway 15% more often than the muon one. That suggests the influence of particles beyond the standard model. Here's how the LHCb results compare with those from other experiments.



The W boson puzzle

The latest analysis of data from the CDF detector at the Tevatron — a US collider — suggests that the mass of the W boson is higher than the standard model predicts. Most other experiments disagree.



©nature

Nik Spencer/*Nature*; Source: CERN

The most recent is from the Tevatron collider at the Fermi National Accelerator Laboratory (Fermilab) in Batavia, Illinois, which shut down in 2011. Researchers have spent the past decade poring through data from the Tevatron's CDF experiment. In April, they reported¹ that the mass of the W boson, a fundamental particle that carries the weak nuclear force involved in radioactive decay, is [significantly higher than the standard model predicts](#).

That doesn't chime with LHC data: measurements at ATLAS and LHCb disagree with the CDF data, although they are less precise. Physicists at CMS are now working on their own measurement, using data from the machine's second run. Data from run 3 could provide a definitive answer, although not immediately, because the mass of the W boson is notoriously difficult to measure.

B-meson confusion

The LHC's data have hinted at other anomalies. In particular, evidence has been building for almost a decade of odd behaviour in particles called B mesons. These transient particles, which quickly decay into others, are so named because they contain pairs of fundamental particles that include a 'bottom' or 'beauty' quark. LHCb analyses suggest that B-meson decays tend to produce electrons more often than they produce their heavier cousins, muons². The standard model predicts that nature should not prefer one over the other, says Tara Shears, a particle physicist at the University of Liverpool, UK, and a member of the LHCb collaboration. "Muons are being produced about 15% less often than electrons, and it's utterly bizarre," she says.

The result differs from the predictions of the standard model with a significance of around 3 sigma, or 3 standard deviations from what's expected — which translates to a 3 in 1,000 chance that random noise could have produced the apparent bias. Only more data can confirm whether the effect is real or a statistical fluke. Experimentalists might have misunderstood something in their data or machine, but now that many of the relevant LHCb detectors have been replaced, the next phase of data-

gathering should provide a cross-check, Shears says. “We will be crushed if [the anomaly] goes away. But that’s life as a scientist, that can happen.”

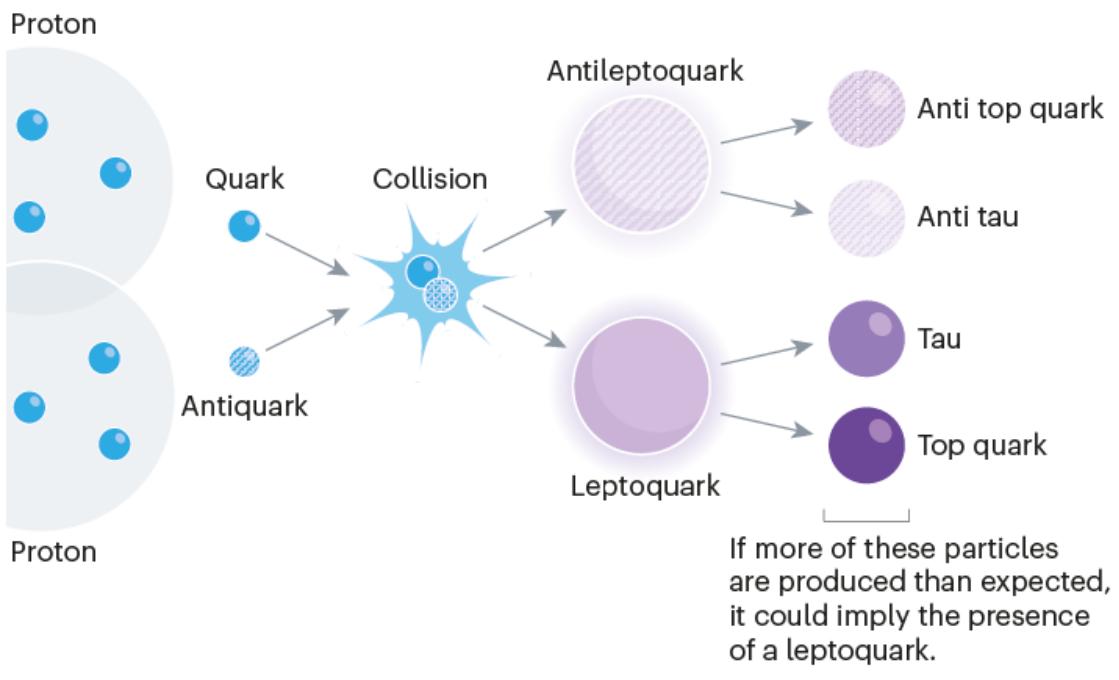
The anomaly is backed up by similar subtle discrepancies that LHCb has seen in other decays involving bottom quarks; experiments at colliders in Japan and the United States have also seen [hints of this odd result](#). This kind of work is LHCb’s métier: its detectors were designed to study in detail the decays of particles that contain heavy quarks, allowing the experiment to gather indirect hints of phenomena that might influence these particles’ behaviour. CMS and ATLAS are more general-purpose experiments, but experimenters there are now checking to see whether they can spot more of the events that are sensitive to the anomalies, says Florencia Canelli, an experimental particle physicist at the University of Zurich in Switzerland and member of the CMS collaboration.

Hunt for the leptoquark

CMS and ATLAS will also do what LHCb cannot: comb collision data to look directly for the exotic particles that theorists suggest could be causing the still-unconfirmed anomalies. One such hypothetical particle has been dubbed the leptoquark, because it would, at high energies, take on properties of two otherwise distinct families of particles — leptons, such as electrons and muons, and quarks (see ‘Decoding decays’). This hybrid particle comes from theories that seek to unite the electromagnetic, weak and strong fundamental forces as aspects of the same force, and could explain the LHCb results. The leptoquark — or a complex version of it — also fits with another tantalizing anomaly; a measurement last year³, from the Muon $g - 2$ experiment at Fermilab, that [muons are more magnetic than expected](#).

DECODING DECAYS

Physicists might be able to infer the presence of short-lived heavy particles by seeing the more stable particles they decay into. For instance, here is one way that the ‘leptoquark’ — a hypothesized transient particle that takes on properties of both leptons (such as electrons) and quarks — might be produced and decay.



©nature

Nik Spencer/*Nature*

At the Moriond particle-physics conference in La Thuile, Italy, in March, CMS researchers presented results of a search that found [intriguing hints of a beyond-standard-model lepton](#). This particle would interact with leptoquarks and is predicted by some leptoquark theories. Physicists saw a slight excess of the particles that the proposed lepton could decay into, bottom quarks and taus (heavier cousins of the muon), but the finding’s significance is only 2.8 sigma. “Those are very exciting results, as LHCb is also seeing something similar,” says Pedraza. CMS physicists presented hints of other new phenomena at the conference: [two possible particles that might decay into two taus](#), and a potential high-energy particle that, through a theorized but unproven decay route, would [turn into distinctive particle cascades termed jets](#).

Another intriguing result comes from ATLAS, where Ismet Siral at the University of Oregon in Eugene and his colleagues looked for hypothetical heavy, long-lived charged particles. In trillions of collisions from 3 years of data they found 7 candidates at around 1.4 TeV, around 8 times the energy of the heaviest known particle⁴. Those results are 3.3 sigma, and the identity of the candidate particles remains a mystery. “We don’t know if this is real, we need more data. That’s where run 3 comes in,” says Siral.



CERN’s 86-metre long Linac4 accelerator, which produces proton beams for the Large Hadron Collider. Credit: Robert Hradil, Monika Majer/ProStudio22.ch/CERN

Another LHC experiment, ALICE, will explore its own surprising finding: that the extreme conditions created in collisions between lead ions (which the LHC smashes together when not working with protons) might crop up elsewhere. ALICE is designed to study quark–gluon plasma, a hot, dense soup of fundamental particles created in collisions of heavy ions that is thought to have existed just after the Big Bang. Analyses of the first two runs found that particles in proton–proton and proton–lead ion collisions show some traits of this state of matter, such as paths that are correlated

rather than random. “It’s an extremely interesting, unexpected phenomenon,” says Barbara Erazmus, deputy spokesperson for ALICE at CERN.

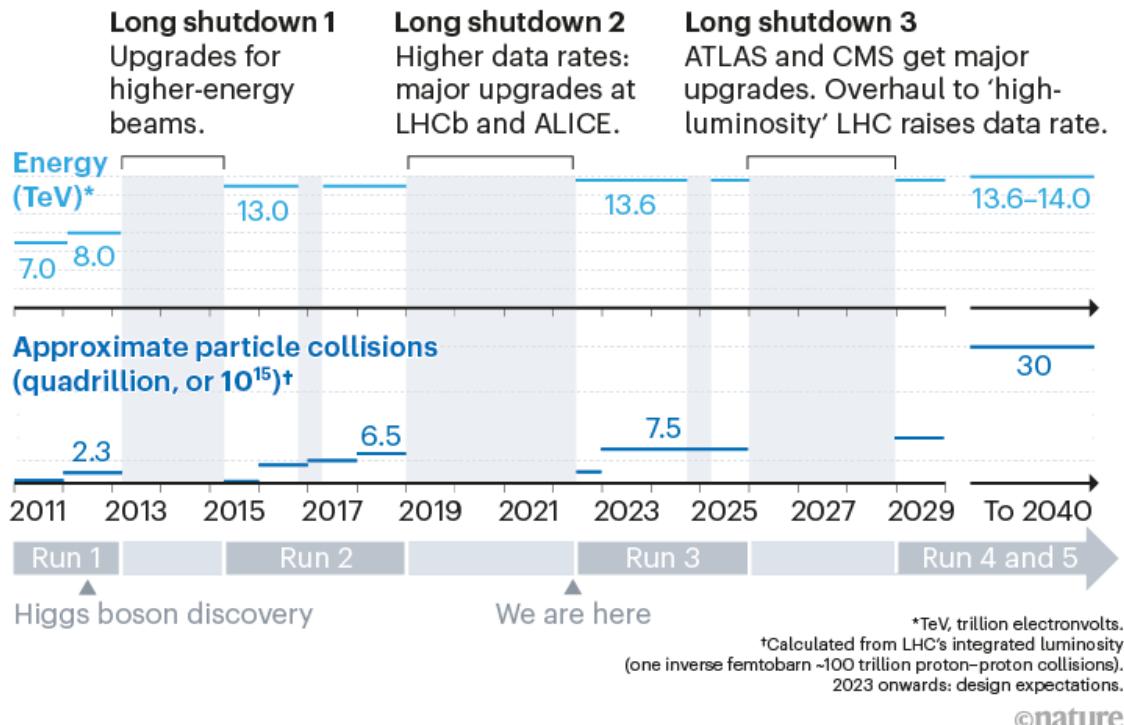
Like LHCb, ALICE has had a major upgrade, including updated electronics to provide it with a faster software-only trigger system. The experiment, which will probe the temperature of the plasma as well as precisely measuring particles that contain charm and beauty quarks, will be able to collect 100 times more events this time than in its previous two runs, thanks to improvements across its detectors.

Machine learning aids the search

Run 3 will see also entirely new experiments. FASER, half a kilometre from ATLAS, will hunt for light and weakly interacting particles including neutrinos and new phenomena that could explain dark matter. (These particles can’t be spotted by ATLAS, because they would fly out of collisions on a trajectory that hugs close to the LHC’s beamline and evades the detectors). Meanwhile, the ATLAS and CMS experiments now have improved detectors but will not receive major hardware upgrades until the next long shutdown, in 2026. At this point, the LHC will be overhauled to create more focused ‘high-luminosity’ beams, which will start up in 2029 (see ‘LHC timeline’). This will allow scientists in the following runs to collect 10 times more collision data than in runs 1 to 3 combined. For now, CMS and ATLAS have got prototype technology to help them prepare.

LHC TIMELINE

The Large Hadron Collider (LHC) will be further upgraded from 2026 to 2029 to conduct even more particle collisions, at higher energies. It is then scheduled to run for another decade.



Nik Spencer/*Nature*; Source: CERN

As well as collecting more events, physicists such as Siral are keen to change the way in which LHC experiments hunt for particles. So far, much of the LHC's research has involved testing specific predictions (such as searching for the Higgs where physicists expected to see it) or hunting for particular hypotheses of new physics.

Scientists thought this would be a fruitful strategy, because they had a good steer on where to look. Many expected to find new heavy particles, such as those predicted by a group of theories known as supersymmetry, soon after the LHC started. That they have seen none rules out all but the most convoluted versions of supersymmetry. Today, few theoretical extensions of the standard model seem any more likely to be true than others.

Experimentalists are now shifting to search strategies that are less constrained by expectations. Both ATLAS and CMS are going to search for

long-lived particles that could linger across two collisions, for instance. New search strategies often mean writing analysis software that rejects the usual assumptions, says Siral.

Machine learning is likely to help, too. Many LHC experiments already use this technique to distinguish particular sought-for collisions from the background noise. This is ‘supervised’ learning: the algorithm is given a pattern to hunt for. But researchers are increasingly using ‘unsupervised’ machine-learning algorithms that can scan widely for anomalies, without expectations. For example, a neural network can compare events against a learned simulation of the standard model. If the simulation can’t recreate the event, that’s an anomaly. Although this kind of approach is not yet used systematically, “I do think this is the direction people will go in,” says Sascha Caron of Radboud University Nijmegen in the Netherlands, who works on applying these techniques to ATLAS data.

In making searches less biased, the triggers that decide which events are interesting to look at are crucial, so it helps that the new GPUs will be able to scour candidate events with wider criteria. CMS will also use an approach called ‘scouting’: analysing rough reconstructions of all the 100,000 or so events initially selected but not saved in full detail. “It’s the equivalent of 10 years more of running your detector, but in one year,” says Andrea Massironi, a physicist with the CMS experiment.



The detector at the Large Hadron Collider’s CMS experiment, pictured during the machine’s shutdown. Credit: Samuel Joseph Hertzog, Julien Marius Ordan/CERN

The triggers themselves could also soon rely on machine learning to make their choices. Katya Govorkova, a particle physicist at CERN, and her colleagues have come up with a high-speed proof-of-principle algorithm that uses machine learning to select which of the collider’s 40 million events per second to save, according to their fit with the standard model⁵. In run 3, researchers plan to train and test the algorithm on CMS collisions, alongside the experiment’s conventional trigger. A challenge will be knowing how to analyse events that the algorithm labels as anomalous, because it cannot yet point to exactly why an event is anomalous, says Govorkova.

Physicists must keep an open mind about where they might find the thread that will lead them to a theory beyond the standard model, says Amhis. Although the current crop of anomalies is exciting, even previous oddities seen by multiple experiments turned out to be [statistical flukes that faded away](#) when more data were gathered. “It’s important that we continue to

push all of the physics programme,” she says. “It’s a matter of not putting all your eggs in one basket.”

Nature **605**, 604–607 (2022)

doi: <https://doi.org/10.1038/d41586-022-01388-6>

References

1. CDF Collaboration. *Science* **376**, 170–176 (2022).
2. LHCb collaboration. *Nature Phys.* **18**, 277–282 (2022).
3. Abi, B. *et al.* *Phys. Rev. Lett.* **126**, 141801 (2021).
4. ATLAS Collaboration. Preprint at arXiv
<https://doi.org/10.48550/arXiv.2205.06013> (2022).
5. Govorkova, E. *et al.* *Nature Mach. Int.* **4**, 154–161 (2022).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01388-6>

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

COVID derailed learning for 1.6 billion students. Here's how schools can help them catch up

[Download PDF](#)

- NEWS FEATURE
- 25 May 2022

COVID derailed learning for 1.6 billion students. Here's how schools can help them catch up

The pandemic is the largest disruption to education in history. But research has identified ways to help children make up lost ground. Will they work in classrooms around the world?

- [Helen Pearson](#)



Students in India in March; schools there were closed for months earlier in the pandemic. Credit: Saqib Majeed/SOPA Images/LightRocket/Getty

By October last year, Meg Brydon could see the terrible toll the pandemic had taken on children at her school. Brydon was a teacher at Ashwood High School, in the suburbs of Melbourne, Australia — the city that has spent more time in COVID-19 lockdowns than any other in the world. The school had been closed, on and off, for about seven months.

Before the pandemic, around 10% of children who joined Ashwood at the age of 12 would be below the expected national standard. But in the latest cohort, Brydon could see that a shocking 30% of them were behind. And the damage ran even deeper. So many children had behavioural or psychological problems after lockdowns that some were getting violent, and the school hired a full-time psychologist to help. “The number of referrals to her was astronomical,” Brydon says.

Similar scenarios have played out in classrooms around the world. By February this year, [schools globally had been closed because of COVID-19](#) for an average of 4.5 months, affecting an estimated 1.6 billion students and

creating what the United Nations has called the largest disruption to education in history. Even 2 years into the pandemic, 48 countries had not yet fully reopened their schools, according to the UN cultural organization UNESCO.

The consequences of these closures follow a sad but predictable course. In rich countries, disadvantaged and vulnerable children have fallen behind the most. Those in poorer countries have been the hardest hit, and millions will never go back to school at all. UNESCO estimates that today's generation of students could lose US\$17 trillion in lifetime earnings at current values because of missed learning and skills. "We're really talking about a generational loss," says Margarete Sachs-Israel, who leads the Inclusive Quality Education Section at UNESCO in Bangkok.

Now, governments and schools need to know the best approach to help children catch up — and research could show the way. Over the past 20–30 years, researchers in education, economics and international development have built substantial bodies of evidence, including banks of randomized controlled trials, showing strategies that are effective at boosting school attendance and learning. They reveal, for example, that tutoring is one of the most cost-effective ways to help children to make up lost ground. And some countries are drawing on this evidence in their COVID-19 responses, putting a focus on tutoring and other programmes that educational studies have shown to be effective.

But experts point to a number of concerns. The true extent of learning losses in the pandemic is not yet clear; educational research rarely provides simple answers about what to do; and nations might not use this opportunity to make much-needed systemic change. "Every single time there's been a calamity in the world, we've rushed back to the old normal fast," says John Hattie, an educational researcher at the University of Melbourne. "The biggest travesty of COVID is if we learn nothing."

What's more, the scale of the task ahead is immense. Researchers and education experts are concerned that the amounts being invested are laughably insufficient, given the number of students who need help. "It's a real test for the global community," says Kenneth Russell, an education

specialist at the UN children's charity UNICEF in New York. "And I don't think the magnitude of the response matches the magnitude of the need."

Even so, the pandemic could eventually drive some transformative changes in education — ones that both improve practices and reach more students, researchers say. "I do think it has thrown into the air many of the assumptions that we make about education," says Lee Elliot Major, who studies social mobility at the University of Exeter, UK.

Tough sell

The concept of using research in education has been a long, tough sell. "The fundamental issue is that many practitioners do not believe it will ever be a science," says Andreas Schleicher, who heads the directorate for education and skills at the Organisation for Economic Co-operation and Development (OECD) in Paris. Teachers are not expected to browse academic journals, and educational policies are often set by the ideology of bureaucrats rather than by research showing what actually works. "Many of them use evidence to confirm what they want to do," Schleicher says.

Some researchers and educators have been trying to change that view for decades. They want education to operate more like medicine, where a drug typically has to be proven effective in randomized controlled trials before it's used. Advocates of evidence-informed education argue that teaching and learning methods should also be shown to work by research — rather than being used because of tradition, opinion or the latest fad. But they acknowledge that testing whether a method improves educational outcomes is often more complex than testing whether a drug improves health.



A student in Tokyo receives one-to-one tutoring through a transparent barrier in August 2020.Credit: Toru Hanai/Bloomberg/Getty

In late 2010, evidence-informed education got one of its biggest boosts when the UK government invested £125 million (US\$156 million) to raise standards in schools. This gave rise to the Education Endowment Foundation (EEF), a non-profit organization in London that has since become a leader in educational research. It has funded at least 160 randomized controlled trials in education, probably more than any other organization in the world. Around half of English schools have taken part in these trials. The investment in the EEF “had a ripple effect around the world”, says Annette Boaz, who studies evidence and policy at the London School of Hygiene & Tropical Medicine.

Other databases of educational research have flowered, too. Hattie led an early, pioneering project to synthesize evidence from around the globe on what influences learning¹. And, the US Department of Education’s Institute of Education Sciences in Washington DC maintains the What Works Clearinghouse, a source of information on educational programmes that have been shown to be effective through rigorous research. Hattie argues

that with databases such as these, the field doesn't need more evidence — the challenge lies in getting the information used by governments and schools. "We're hopeless at that," he says.

The pandemic could, in theory, help to bridge that gap. Countries worldwide want to know the best way to invest in educational recovery, and billions of dollars are already pouring into schools. "This moment in time really is a unique one, for changing the conversation about evidence in education," says Nancy Madden, a psychologist and researcher at Johns Hopkins University School of Education in Baltimore, Maryland. "People want something that works, they aren't just doing business as usual."

Dismantling dogma

The crown jewel at the EEF is its Teaching and Learning Toolkit, which is based on systematic reviews and meta-analyses of studies, such as randomized controlled trials, that have tested 30 educational approaches. The toolkit translates findings into an easy-to-understand metric: the number of months of additional progress achieved over a year, on average, by children who receive an intervention, compared with similar children who do not. It also displays the strength of the underlying evidence and the intervention's cost (see 'Which educational techniques get top grades?' and go.nature.com/3nbhdzm).

WHICH EDUCATIONAL TECHNIQUES GET TOP GRADES?

The Education Endowment Foundation, a UK charity, has systematically reviewed evidence supporting many educational approaches. Its Teaching and Learning Toolkit for schools rates each technique on the basis of its cost, how much it improves student achievement and the strength of evidence supporting it.

★ Impact Q Evidence £ Cost



Better

High impact,
extensive evidence,
low cost



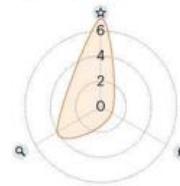
Worse

Low impact, limited
evidence, high cost

See below for definitions.

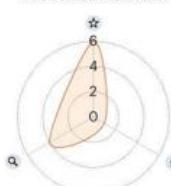
Metacognition and self-regulation

Learning how to learn, such as through planning and evaluation



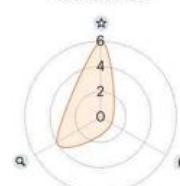
Oral language interventions

Focus on speaking
and verbal interaction



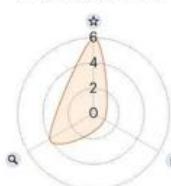
Reading comprehension strategies

Improving understanding
of written text



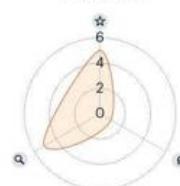
Feedback

Providing meaningful
information about a
student's performance



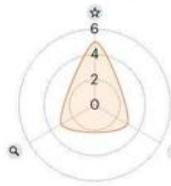
Phonics

Knowledge of the relationship
between written symbols
and sounds



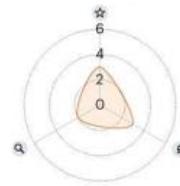
One-to-one tutoring

Intensive individual
support for pupils



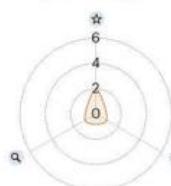
Extending school time

Adding hours or days
of learning



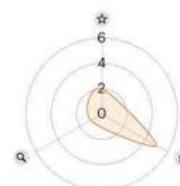
Within-class attainment grouping

Organizing students by
level in a class



Reducing class size

Lowering the student:teacher
ratio; most studies looked at
cuts of 8–10 students



Impact: Number of months of progress made over a year, on average, by children who received the intervention, compared to similar children who did not.

Evidence: The robustness of the evidence, based on the number of studies supporting each intervention and their rigour. 1 = at least 10 studies (very limited evidence); 2 = 11–24 (limited); 3 = 25–44 (moderate); 4 = 45–69 (extensive); 5 = 70 or more (very extensive; see go.nature.com/3nbhdzm for full details).

Cost: Estimated cost of each intervention per student per year. 1 = less than £80 (US\$99); 2 = up to £200; 3 = up to £720; 4 = up to £1,200, 5 = more than £1,200.

©nature

Source: EEF

The toolkit dismantles many common beliefs by showing that modest reductions in class size (from 30 to 20 students, for example), wearing school uniforms and grouping children according to attainment level have little if any effect, on the basis of the evidence so far. The most effective strategies include ones that help children to understand what they read; giving them meaningful feedback; and approaches that improve meta-cognition — the ability of students to think about, plan and evaluate their own learning. These each give children six or seven months of progress, on average.

More than 70% of secondary-school leaders in England now use the toolkit when making decisions about how to spend funding. The EEF has partnered with groups to adapt it for use in Australia and parts of Latin America, the Middle East and Africa.

Long before the pandemic, it was clear that one of the most cost-effective approaches is tutoring, either in small groups or one-to-one. The toolkit says this can buy four to five months of additional progress at relatively low cost. And, unlike some other effective methods, tutoring programmes can be ramped up and implemented quickly. So, in 2020, the EEF rapidly reviewed evidence on the possible impacts of the United Kingdom's nationwide school closures² and highlighted that tutoring was likely to be a particularly effective way to help children to catch up. At the time, “tutoring seemed such a plausible response”, says Becky Francis, an education researcher who is chief executive of the EEF. The recommendation “landed in a void at the time and was seized upon eagerly by policymakers”, she says.

In June 2020, the UK government announced a £350-million National Tutoring Programme as part of its wider £1-billion catch-up funding for children. (The EEF was one of several partners that ran the programme for the first year; the Dutch company Randstad took it over in the second year.) But the tutoring programme has been widely criticized for drastically failing to reach enough children, in particular those who stand to gain most from it. “I think it hasn’t targeted the most disadvantaged pupils properly. It hasn’t won over teachers,” says Elliot Major. “And partly that’s because there’s some scepticism about variation in the quality of the tutors.”

This March, the government ended Randstad's contract and announced that funding for tutoring would go directly to schools in the 2022–23 academic year. The National Foundation for Educational Research in Slough, UK, is conducting independent evaluations of the tutoring programme's impact on student attainment.

Both the EEF and Randstad say they are proud of what they achieved with the tutoring programme. In statements to *Nature*, the EEF said that 60% of secondary schools had accessed tutoring by July 2021, and Randstad said it had tripled the number of students in the tutoring programme.

Case studies

Another evidence-backed programme has been widely, and less controversially, put in place in England. The Nuffield Early Language Intervention (NELI) has been shown in randomized controlled trials to boost language skills in children aged 4–5 through a series of teaching sessions in small groups (see go.nature.com/39xtgsk). NELI is now being used in two-thirds of English primary schools to help make up for learning missed during the pandemic, and its results are being independently evaluated. “Although it’s had a tremendous reach, it’s flown almost entirely under the radar,” says Francis.

Some researchers point to the Netherlands as having taken an exemplary approach to education recovery based on evidence. There, the government handed €4.2 billion (US\$4.4 billion) of funding to schools to support students, and required that they spend it by picking from a ‘menu card’ of evidence-based approaches largely based on the EEF’s toolkit. “We want to make sure as much as possible that schools will base their decisions on knowledge that’s available on effective approaches,” says Femke Bink, senior adviser in the Department for Secondary Education at the Ministry of Education, Culture and Science in The Hague.



Two children attend school remotely using a smartphone in Dehradun, India, in late January. By that point in the COVID-19 pandemic, they had been to school in person for only one month since March 2020, according to their mother. Credit: Atul Loke/The New York Times/Redux/eyevine

And in Panama, where schools were fully closed for more than a year, the Ministry of Education in April launched resources and training for teachers showing how to implement evidence-based practices, including feedback to students. “Teachers are tired and stressed, so we’re trying to say to them, ‘we want to channel your efforts into what really works’,” says Javier González, director of SUMMA in Santiago, Chile. SUMMA aims to improve education systems in Latin America and the Caribbean using research, and helped to develop the training.

The United States, too, has put some emphasis on evidence in its recovery plans. In 2021, a giant stimulus bill channelled \$122 billion to schools. The law requires that at least 20% of funds received by districts must be used on evidence-based measures to help students’ academic, social and emotional needs. In practice, however, it’s hard to know how this money is being used, says Mike Petrilli, president of the Thomas B. Fordham Institute, an

educational foundation in Washington DC. “Based on past experience, we should expect that much of the money will not be spent in the best way.”

Another complication is that tutoring comes in many styles: one-to-one or small groups; online or in person; delivered by human teachers or digital ones. There is no guarantee that a particular programme will be effective, or that it will be successful in a particular school or for a certain child. “It’s not just hiring some people that call themselves tutors and putting them in the room with some kids — you can waste a lot of money that way,” Madden says.

In Melbourne, Brydon saw the challenges of putting a tutoring programme in place. Her school was able to place an extra teacher in some classrooms to help children who have fallen behind, using money it received as part of a catch-up programme from the government. But the school is struggling to find teachers to fill positions, she says, because exhausted colleagues are quitting their jobs. “We need upwards of ten substitute teachers every day just to keep the school running,” she says.

Global problems

Things are looking even grimmer elsewhere in the world. UNESCO estimates that, by April 2020, more than 1.2 billion children in the highly populous Asia Pacific region had been affected by school closures. And, whereas schools closed in Japan and Singapore for only a month or so, those in Bangladesh and the Philippines have experienced some of the worst disruptions in the world, with schools fully shut for more than 13 months.

Even before COVID-19, there was a learning crisis in the region, Sachs-Israel says, because so many children did not achieve expected proficiency levels at school. An estimated 10 million children in the Asia Pacific region will not go back to school, and the expectation is that early or forced marriages and child labour are expected to soar.

The scale of this problem is not one that extra tutoring alone can address. With many schools still closed, the obvious top priority, say education specialists, is for classes to reopen so that children can return — even if

COVID-19 cases start rising again. Sachs-Israel says schools have to be welcoming and safe, and need to overcome any fears that parents, teachers and children might have about infection risks.

According to a 2020 report³ from an international group called the Global Education Evidence Advisory Panel, one cost-effective approach for schools is to target teaching to a child's learning level rather than to their age. And education researchers say that schools should assess each returning student.

This is the strategy behind an evidence-based programme called Teaching at the Right Level, run by the learning organization Pratham in New Delhi. The organization's chief executive, Rukmini Banerji, says it is working with several state governments in India and other countries, and has observed that children are making progress in basic literacy and numeracy in just a few weeks. "We feel that is what is really needed across the world," she says.

Questioning the evidence

Even with all the support for use of evidence in education, there have been some long-standing concerns about how reliable some of that evidence is.

In 2019, a pair of researchers examined 141 large randomized controlled trials commissioned by the EEF and the US-based National Center for Educational Evaluation and Regional Assistance. They concluded that 40% of the trials were uninformative because their effects were small or imprecise⁴.

"So at the beginning, you didn't know whether the intervention works or not. But at the end, we're still unsure whether it works," says study author Hugo Lortie-Forgues, who studies mathematics education at Loughborough University, UK. This could be because early, promising research on an approach turned out to be misleading, a method was hard to scale up or the trial was poorly designed, he says.

This was no big surprise to researchers who conduct such studies. Just as most new drugs prove ineffective in large clinical trials, most bright ideas for improving learning show little effect when they are put to the test. And

whereas in medicine, physicians start with someone who is ill and try to make them measurably better, in education, many countries are starting with a fairly healthy education system — so any new method is likely to produce only marginal gains. “It’s perhaps a little naive to assume that teachers haven’t discovered, over time, some of the approaches that are more likely to be successful,” says Steve Higgins at Durham University, UK, who has led work on the EEF’s toolkit.

With data still rolling in, there are some suggestions that school closures might have had a smaller impact on some children’s achievement than many doom-laden headlines suggest — or that students might bounce back quickly.

When Hattie examined the effects of school closures in Victoria, Australia, where schools (including Brydon’s) had been closed for extended periods, he concluded that it was surprising that learning trajectories had only marginally decreased (see go.nature.com/3mtxucq). One possible reason is that some students working alone were able to be more efficient than at school. Schleicher adds that technology also became more accepted, teachers rallied to support children socially and emotionally, and parents became more involved in their children’s education. Looking at the overall impacts of the pandemic on education, he says, “the balance sheet has pluses and minuses”.

Teacher training

In the longer term, a key way to get research used in education more routinely will be to weave it into teachers’ training and continuing professional development. One model comes from Japan, where teachers have for decades conducted ‘lesson study’. This is a form of research in which they develop a goal — to improve understanding of fractions, say — then write a detailed lesson plan, observe the lesson in action and discuss what they learnt. Schools draw on external research and often consult an academic in the process. This type of ongoing professional development is unusual, says mathematics education specialist Toshiakira Fujii at Tokyo Gakugei University. Teachers develop a deep understanding of teaching materials “but more importantly they learn how to learn as a teacher”.

Other countries are starting to integrate evidence into teacher training, too. The EEF and SUMMA are working with the University of West Indies at Cave Hill, Barbados, to train teachers in evidence-based practices such as giving effective feedback to students. And starting this year, all 650 students enrolled in the master's in education at Harvard Graduate School of Education in Cambridge, Massachusetts, will have to take a course on evidence, says Carrie Conaway, who is a senior lecturer there. "The idea is that we have a generation of leaders who understand the value of this as part of their decision-making," she says.

Brydon says she was taught almost nothing about using research evidence during her training — "you get exposed to a couple of major theorists and then that's really it". But she is now part of Q Project, an effort in Australia to improve the use of evidence in schools. She thinks that the biggest barrier, however, is a lack of time. "We're so swamped, and when you have to decide between getting your year-12 essays marked or reading some research evidence, I know which one I'm going to choose every day of the week."

Right now, Brydon and her colleagues are still battling to help children to catch up, amid simmering concerns that the next coronavirus variant could shut schools all over again. When people used to ask Brydon about her work, she'd tell them that teaching is the greatest job in the world. But now, for the first time, she has a different response. "There are some parts that I really love," she says, "but other parts that are making it really hard to do the job."

Nature **605**, 608-611 (2022)

doi: <https://doi.org/10.1038/d41586-022-01387-7>

References

1. Hattie, J. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement* (Routledge, 2008).
2. Education Endowment Foundation. *Impact of School Closures on the Attainment Gap: Rapid Evidence Assessment* (EEF, 2020).

3. Global Education Evidence Advisory Panel. *Cost-Effective Approaches to Improve Global Learning* (World Bank, 2020).
4. Lortie-Forgues, H. & Inglis, M. *Educ. Res.* **8**, 158–166 (2019).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01387-7>

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

Books & Arts

- **[Santiago Ramón y Cajal: art, politics and neuroscience revolution](#)** [24 May 2022]
Book Review • A biography charts the rise of the father of modern neuroscience.
- **[Ukrainian nuclear historian chronicles six disasters](#)** [23 May 2022]
Book Review • From explosions to meltdowns, history provides key context in the search for alternative energy sources amid decarbonization.

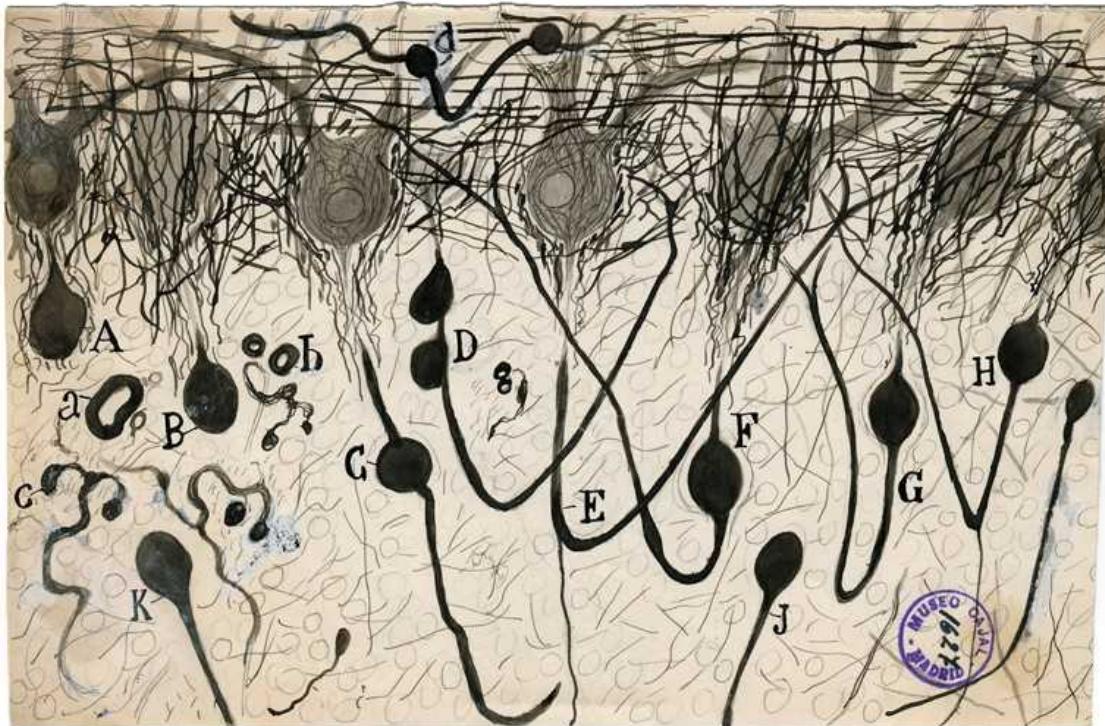
Santiago Ramón y Cajal: art, politics and neuroscience revolution
[Download PDF](#)

- BOOK REVIEW
- 24 May 2022

Santiago Ramón y Cajal: art, politics and neuroscience revolution

A biography charts the rise of the father of modern neuroscience.

- [Alison Abbott](#) 



A drawing of neurons in the cerebellum by Santiago Ramón y Cajal. Credit: Cajal Institute (CSIC), Madrid

The Brain in Search of Itself: Santiago Ramón y Cajal and the Story of the Neuron *Benjamin Ehrlich* Farrar, Straus and Giroux (2022)

Is the brain's grey matter a continuous network of fibres? At the end of the nineteenth century, this was a contentious question, with high-stakes implications: an opposing theory that the grey matter was made of separate cells had gained traction by 1906, when the Nobel prize honoured two giants of neuroscience for work on the structure of the nervous system. The laureates embodied the hostile extremes of the dispute, a feud that has gone down in the annals of science history.

Camillo Golgi, who clung to the continuous-web theory, abused his Nobel acceptance speech to attack his younger co-laureate, Santiago Ramón y Cajal. Cajal behaved himself at the ceremony, but elsewhere, he had described Golgi's stubbornness as "a typical case of the crippling influence of theoretical prejudice".

Cajal was proved right. His delicate tissue preparations clearly showed tiny gaps between individual brain cells. His work has been compared to that of Andreas Vesalius, the Renaissance physician whose exquisite drawings of his own anatomical dissections revolutionized medicine. Cajal's immense artistic skills similarly led him to profound insights into brain development and function, most of which have stood the test of time.

He began his life as a juvenile delinquent in the remote Pyrenean mountains of northern Spain, relates Benjamin Ehrlich in his deeply researched book *The Brain in Search of Itself*. The first major English-language biography of Cajal, it maps out his scientific work in broad strokes and describes his politically turbulent times. Such is the wealth of detail, however, that the narrative sometimes struggles to maintain its thread.

Cajal's domineering father was determined that his son should follow in his footsteps as a physician. But Cajal was a reluctant pupil: he found rote learning difficult and kicked against discipline, resulting in beatings from his father. When he declared that he wanted to be a professional artist — he drew and painted obsessively — his father confiscated his brushes and sent him to a remote and harsh Jesuit secondary school, from which he absconded after a few months. At his next school, one teacher described him

as “inattentive, lazy, disobedient, and annoying”. Another said he would end up in jail “if they do not hang him first”. Once, he taught himself to make gunpowder, constructed a cannon from junk and fired it at a neighbour’s house. That escapade earnt him a few days in jail.

Somehow, Cajal scraped his baccalaureate, and enrolled in medical school in Zaragoza, the region’s largest town. There, he came across the cell theory of German pathologist Rudolf Virchow, which held that the body, at least outside the brain, is composed of individual basic units of life. Excited, Cajal sought out the university’s only microscope and looked down at a frog whose cells were rendered visible by a pink stain derived from cochineal insects.

Revolution was rumbling during the years of Cajal’s education, and Spain’s colonies, particularly Cuba, were struggling for their independence. In 1873, when he graduated, the first Spanish republic was declared and Cajal was conscripted into the army’s medical corps. Dispatched to Cuba, he became so ill with malaria that he requested discharge.

Thin and in poor health, he taught at his alma mater and prepared for his doctorate. Browsing a medical supply shop while in Madrid for an exam, he saw the most advanced microscope then available. It cost more than half his yearly income; he used his army discharge money to help buy it.

Microscopy methods

Cajal shut himself in the attic of the family home, learning how to prepare tissue for microscopy and drawing or photographing his preparations. At 31, he landed his first faculty position, in Valencia. He wrote a textbook of histology, which surveyed all tissues of the human body except the brain. That was too difficult to stain. On another fateful visit to Madrid, he was introduced to a staining technique called *la reazione nera* — the black reaction — developed by Golgi some years earlier. It displayed nerve cells beautifully.

Cajal refined the highly unreliable method. Now in Barcelona, he isolated himself with his microscope and brain preparations and saw individual cells

wherever he looked. He founded a journal, *The Trimonthly Review of Normal and Pathological Histology*. Its inaugural issue, in May 1888, described the first unequivocal evidence that the brain was made of individual cells. Cajal sent copies to leading scientists abroad; he got no response. The next year, at an international congress in Berlin, his work was finally recognized. He shot to fame. Golgi was enraged that his own method had been used, as he saw it, against him.

Cajal was a prolific, flowery writer who published popular science and novels as well as innumerable academic papers. He developed new stains for nerve cells and described the development and evolution of the nervous systems in other species. He elaborated all this in the two-volume *Textura del sistema nervioso del hombre y los vertebrados* (*Texture of the Nervous System of Man and the Vertebrates*, 1899–1904) — 2,000 pages of text with nearly 1,000 of his own illustrations.

Ehrlich paints a vivid picture of a strong, driven character, anti-authoritarian, anti-elite and brash. But he doesn't develop how and why Cajal's discoveries have been so fundamental to modern neuroscience. None of Cajal's drawings is included, an incomprehensible omission. There is an occasional misinterpretation; for example, he implies that Cajal had a dig at Golgi in his Nobel acceptance speech, but the passage he quotes refers to other scientists.

Still, the narrative is charming. In his late sixties, Cajal found himself behaving like Golgi. A younger colleague developed a new stain that enabled him to see even finer structures than Cajal had been able to see. He used it to show two distinct types of glial cell — non-neuronal brain cells. Cajal had judged these cells, which he called the “third element”, to be homogeneous. Incensed, Cajal dismissed the discovery. Unlike Golgi, he later reflected on his behaviour and apologized.

Nature **605**, 613–614 (2022)

doi: <https://doi.org/10.1038/d41586-022-01389-5>

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

Ukrainian nuclear historian chronicles six disasters

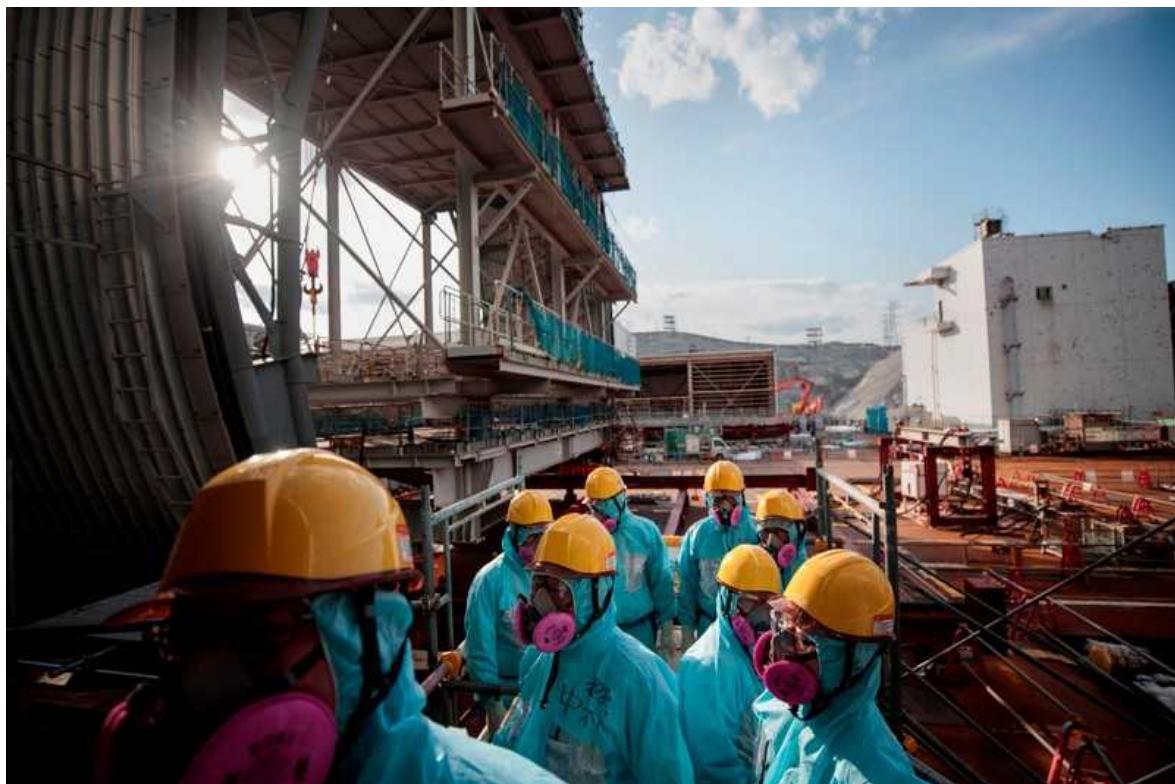
[Download PDF](#)

- BOOK REVIEW
- 23 May 2022

Ukrainian nuclear historian chronicles six disasters

From explosions to meltdowns, history provides key context in the search for alternative energy sources amid decarbonization.

- [Alexandra Witze](#)



Workers gather outside one of the reactors that melted down in the 2011 Fukushima disaster. Credit: Behrouz Mehri/AFP/Getty

Atoms and Ashes: A Global History of Nuclear Disasters *Serhii Plokhy*
W. W. Norton (2022)

In February, soon after Russian forces invaded Ukraine, they reportedly dug trenches in the radioactive soil at Chernobyl and drove heavy vehicles in the area, kicking up contaminated dust. Thirty-six years after a reactor core exploded at the Chernobyl nuclear power plant, fallout from the world's worst nuclear accident still permeates the environment. Amid all the atrocities committed by Russian troops during the war in Ukraine, ignorance of this history does not rank high. But it underscores the lasting, dangerous and frequently unforeseen consequences of nuclear disasters.

One of the leading chroniclers of this nuclear legacy is Serhii Plokhy, a historian of Ukraine and author of *Chernobyl* (2018), a definitive account of that fateful day in April 1986. In *Atoms and Ashes*, he places that disaster in a broader, more global history of six nuclear accidents — from explosions at plutonium production plants to meltdowns at nuclear power stations. The result is a revealing tour of some of the most terrifying experiences involving nuclear power. It is also useful context for today's discussions about whether nuclear energy deserves a major role in the push to decarbonize the global economy.

Atoms for war

Nuclear power is a child of war, birthed in the world's first artificial self-sustaining nuclear reaction inside a stack of graphite and uranium bricks in Chicago, Illinois, in December 1942. That was a key early milestone in the Manhattan Project, which developed the atomic bombs that devastated Hiroshima and Nagasaki in Japan, and brought an end to the Second World War — and instigated the deadly competition between the United States and the Soviet Union as they built up their nuclear arsenals.

While global superpowers battled, remote communities with little geopolitical power often suffered the public-health consequences. Before the 1963 treaty that banned atmospheric nuclear tests, the United States, Soviet Union and United Kingdom regularly showered radioactivity into the air during test blasts. One of the most shocking was a March 1954 test in the

Marshall Islands in the Pacific Ocean. Nicknamed Castle Bravo, it was the first US test of a hydrogen bomb using a new type of fuel; its designers badly miscalculated how big the blast would be. The bomb detonated over Bikini Atoll with a yield of 15 megatons, 2.5 times what was expected. And weather forecasters misjudged where and how strongly the winds were blowing.

Radioactivity spread over inhabited islands including Rongelap, 157 kilometres away. People reported a whitish substance resembling snow falling on them. It was irradiated coral, vaporized in the blast. Given no warning or advice by the US government, people stayed outside until their skin began to burn and itch. Hundreds were unknowingly exposed to radiation.

This theme of distance — of radioactivity traveling far beyond the supposedly secure confines of a nuclear testing ground or power plant — is part of what makes nuclear accidents so terrifying. Everything seems to be under control, until it's not. That was what happened in two plutonium-production accidents in 1957.

At Kyshtym, a plutonium-processing facility in the Soviet Union's Ural Mountains, a tank storing nuclear waste exploded and sent a plume of radioactivity drifting widely. Days later, the nuclear facility of Windscale in the United Kingdom saw thousands of tonnes of graphite ignite and release radioactive material. In both cases, delays to maintenance tasks involving complicated engineering, combined with stressed workers, led to cascading failures. In both cases, national governments covered up the scale and scope of the disasters.

Atoms for peace

These disasters all involved military “atoms for war” projects. Similar mistakes were repeated in civilian “atoms for peace” projects, starting in the 1950s, in which nuclear power was repurposed to provide energy. “Looking closely at what led to these accidents and the ways in which the industry and governments dealt with them,” writes Plokhy, “is the most effective way of understanding the perils associated with reliance on nuclear energy.”

At the Three Mile Island nuclear power plant in Pennsylvania, a minor failure in a valve in 1979 led to a loss of coolant and the start of a reactor meltdown. Ultimately, plant operators vented radioactive gases into the air, at much lower levels than at Kyshtym but enough to cause widespread panic in local communities. (Here, as elsewhere, Plokhy valiantly attempts to be clear about public risk while juggling the many units of radioactivity exposure, from rems to roentgens to sieverts.) Then-US president Jimmy Carter, a Navy veteran who had worked on a damaged reactor in Chalk River, Canada, tried to take charge of the Three Mile Island response and messaging, but ultimately did not hold the US nuclear regulatory agency fully to account.

Plokhy excels in unpacking the human and systemic factors that contribute to nuclear disasters. At Chernobyl, it was a long-overdue test, conducted by exhausted workers around midnight, that led to the catastrophic explosion of reactor number four. Even that accident did not stem the rise of nuclear power in countries that had made it a strategic part of their energy portfolio, such as Japan.

And so, on the afternoon of 11 March 2011, reactors at the Fukushima Daiichi nuclear power plant were operating when a magnitude-9 earthquake rocked the eastern coast of Japan. Emergency generators kicked in to keep coolant flowing, but they were swamped when an enormous tsunami arrived. The plant's designers had anticipated an earthquake but not a tsunami. Three reactors melted down.

Nuclear power is one of the most complex systems that humanity has tried to harness. What looks like a straightforward science and engineering project turns out to be an intricate interplay of commercial and government interests, with plenty of opportunities for bad decisions. Collectively, these six nuclear accidents are a cautionary tale for any Prometheus looking to play with nuclear fire.

They also raise questions about what part nuclear power should play in slashing emissions. Around 10% of the world's electricity comes from approximately 440 nuclear plants around the globe. Industry leaders and supporters point to the long record of reliability for most nuclear power

plants, and the importance of generating energy domestically instead of importing fossil fuels from nations such as, for example, Russia.

Yet Plokhy notes correlations between the fate of the nuclear industry and the history of nuclear disasters. In 1979, the year of Three Mile Island, the number of reactors under construction peaked. The year before Chernobyl, 1985, saw the peak of reactors starting up. And the industry has been in a slide ever since Fukushima.

Nature **605**, 614-615 (2022)

doi: <https://doi.org/10.1038/d41586-022-01390-y>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01390-y>.

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

Opinion

- **[Artificial intelligence is breaking patent law](#)** [24 May 2022]
Comment • The patent system assumes that inventors are human. Inventions devised by machines require their own intellectual property law and an international treaty.
- **[Open access in low-income countries — open letter on equity](#)** [24 May 2022]
Correspondence •

Artificial intelligence is breaking patent law

[Download PDF](#)

- COMMENT
- 24 May 2022

Artificial intelligence is breaking patent law

The patent system assumes that inventors are human. Inventions devised by machines require their own intellectual property law and an international treaty.

- [Alexandra George](#) ⁰ &
- [Toby Walsh](#) ¹

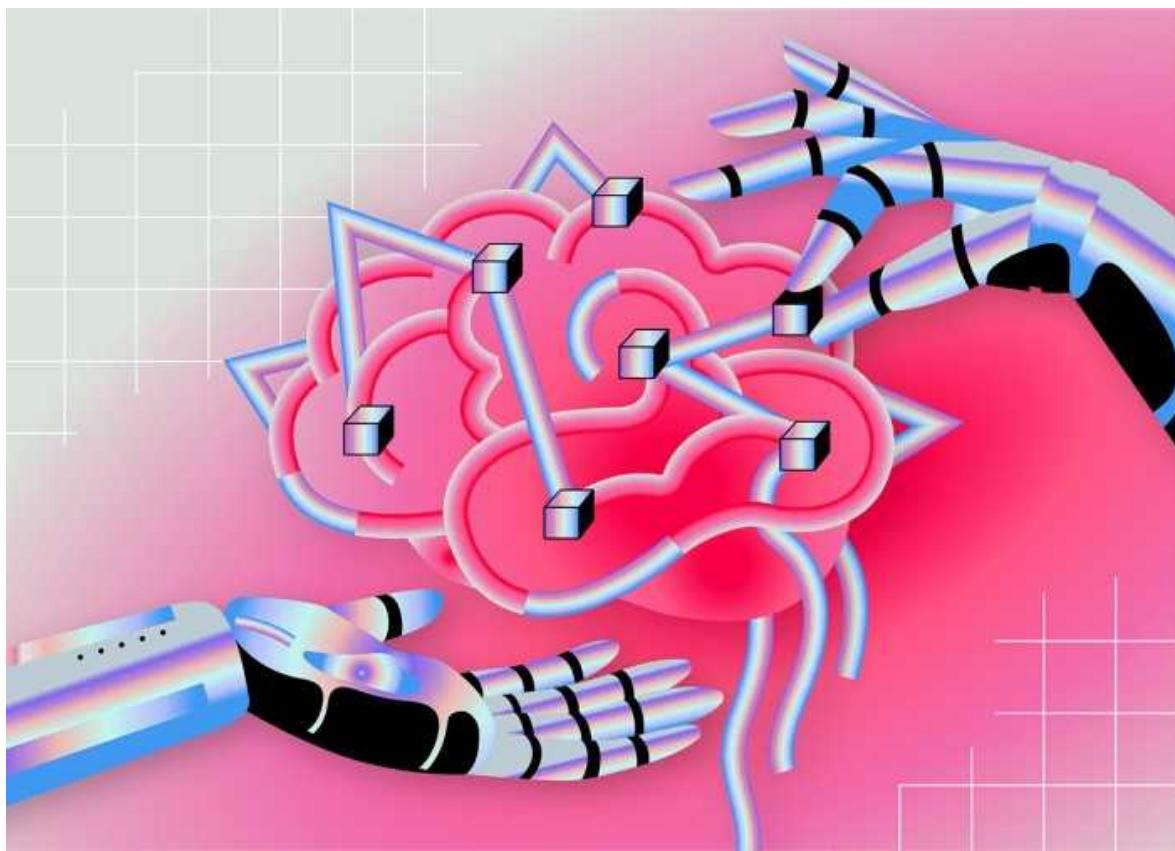


Illustration by Ana Kova

In 2020, a machine-learning algorithm helped researchers to develop a potent antibiotic that works against many pathogens (see [Nature https://doi.org/ggm2p4](https://doi.org/ggm2p4); 2020). Artificial intelligence (AI) is also being used to aid vaccine development, drug design, materials discovery, space technology and ship design. Within a few years, numerous inventions could involve AI. This is creating one of the biggest threats patent systems have faced.

Patent law is based on the assumption that inventors are human; it currently struggles to deal with an inventor that is a machine. Courts around the world are wrestling with this problem now as patent applications naming an AI system as the inventor have been lodged in more than 100 countries¹. Several groups are conducting public consultations on AI and intellectual property (IP) law, including in the United States, United Kingdom and Europe.

If courts and governments decide that AI-made inventions cannot be patented, the implications could be huge. Funders and businesses would be less incentivized to pursue useful research using AI inventors when a return on their investment could be limited. Society could miss out on the development of worthwhile and life-saving inventions.

Rather than forcing old patent laws to accommodate new technology, we propose that national governments design bespoke IP law — AI-IP — that protects AI-generated inventions. Nations should also create an international treaty to ensure that these laws follow standardized principles, and that any disputes can be resolved efficiently. Researchers need to inform both steps.

Who, not what

Machines that are able to invent were not a consideration for drafters of the world's first patent legislation, the Venetian Patent Statute of 1474. Nor were they contemplated in the 1883 Paris Convention for the Protection of Industrial Property, which established the foundations of the international patent system. Even by 1994, AI-generated inventions were still almost unheard of when the World Trade Organization finalized its Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS). The 1883 and 1994 treaties mandate international patent standards today.

The TRIPS agreement protects “any inventions, whether products or processes, in all fields of technology, provided that they are new, involve an inventive step and are capable of industrial application” (see go.nature.com/3n4khc2). In its wording, ‘inventions’, ‘new’, ‘inventive step’ and ‘capable of industrial application’ are terms of art, each with a legal definition. In essence, an object is not patentable if any of these requirements is not met (see ‘What is patentable?’).

What is patentable?

Generally, an invention must meet each of the following requirements before it can be patented.

- **An invention** made by one or more inventors. This includes products, processes or methods in almost all fields of technology.
- **Novel.** The invention does not already exist.
- **Inventive step or non-obvious.** The invention would not be obvious to a ‘person skilled in the art’ who has ‘common general knowledge’ in that field.
- **Capable of industrial application or utility.** The invention can be made or used in industry, does as is claimed and/or has economic significance.

All 164 World Trade Organization members must comply with these principles, standardized by the 1994 Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS).

New technologies have challenged the system before. High-profile cases have tested whether genetic sequences, human-made living organisms and other objects could be patented. The central legal question in these cases was whether they were inventions at all. For example, after a years-long court battle between the US Association for Molecular Pathology (among others) and molecular-diagnostics firm Myriad Genetics in Salt Lake City, Utah, the US Supreme Court concluded in 2013 that isolated human gene sequences were unpatentable because genetic information is a product of nature rather than a human invention².

Inventions generated by AI challenge the patent system in a new way because the issue is about ‘who’ did the inventing, rather than ‘what’ was invented. The first and most pressing question that patent registration offices have faced with such inventions has been whether the inventor has to be human³. If not, one fear is that AIs might soon be so prolific that their inventions could overwhelm the patent system with applications.

Another challenge is even more fundamental. An ‘inventive step’ occurs when an invention is deemed ‘non-obvious’ to a ‘person skilled in the art’. This notional person has the average level of skill and general knowledge of an ordinary expert in the relevant technical field. If a patent examiner

concludes that the invention would not have been obvious to this hypothetical person, the invention is a step closer to being patented.

But if AIs become more knowledgeable and skilled than all people in a field, it is unclear how a human patent examiner could assess whether an AI's invention was obvious. An AI system built to review all information published about an area of technology before it invents would possess a much larger body of knowledge than any human could. Assessed against all knowledge, almost everything would seem obvious⁴. If everyone has access to such AI tools in future, then the 'inventive step' criterion of patentability would be close to impossible to achieve, and almost nothing would be patentable. A complete rethink would be required.

Test case

These issues have been brought into focus by an AI system called DABUS (Device for the Autonomous Bootstrapping of Unified Sentience), created by Stephen Thaler, president and chief executive of US-based AI firm Imagination Engines. Thaler claims that DABUS invented a new type of food container and a flashing light for attracting attention in emergencies.

The inventions are not remarkable. The fallout from them is. In 2018, Thaler's international legal team, led by academic Ryan Abbott at the University of Surrey in Guildford, UK, started submitting applications to patent offices around the world, naming DABUS as the inventor. These cases are thought to be the first to test whether an AI system can be recognized as an inventor under existing laws. Patent offices and courts have had to rule on this question, and have started to flag gaps in the law⁵.

Patent registration offices have so far rejected the applications in the United Kingdom, United States, Europe (in both the European Patent Office and Germany), South Korea, Taiwan, New Zealand and Australia. Challenges to these decisions have for the most part failed, with courts concluding that inventors are presumed to be human (see, for example, go.nature.com/3fjwd9n). In Germany, a court accepted that the inventions could potentially be patented if Thaler was named as the inventor who prompted DABUS to create the inventions — a compromise that

acknowledged the AI system's input. But at this point, the tide of judicial opinion is running almost entirely against recognizing AI systems as inventors for patent purposes.

In the absence of clear laws setting out how to assess AI-generated inventions, patent registries and judges currently have to interpret and apply existing law as best they can. This is far from ideal. It would be better for governments to create legislation explicitly tailored to AI inventiveness. We propose three steps to achieving this goal.

Listen and learn

First, national governments and multilateral bodies involved in patent policy (such as the World Trade Organization) should undertake a systematic investigation of the issues, evidence and viewpoints. They should consult stakeholders including patent registration offices; professional bodies that represent scientists and engineers; consumer and patient advocacy groups; bodies for business development and commercialization, and professionals in IP law. Previous inquiries of this nature have led to changes in many countries' IP laws in response to the development of the Internet and the digital economy. For example, the Australian government's public consultations on online piracy during 2014–18 resulted in laws that allow courts to block access to websites that infringe copyright.

Several countries have already begun preliminary investigations of this kind relating to AI-generated inventions, as has the European Union (go.nature.com/3j6qgu3) and the World Intellectual Property Organization (go.nature.com/3nc79cr). This is a good start, and one that patent systems in every jurisdiction should emulate.

These inquiries must go back to basics and assess whether protecting AI-generated inventions as IP incentivizes the production of useful inventions for society, as it does for other patentable goods. Programmers of AI systems can already obtain some IP protection through copyright in the computer code and patents over the functionality of the software they write. Some people, for political or pragmatic reasons, might prefer to leave the output of

AI-generated inventions in the public domain, free for all to use. Others are calling for IP protection to be extended (see, for example, refs 6–8).

AI-IP law

Tinkering with existing legal protections risks leaving grey areas, so more-comprehensive law reform is preferable. An ideal solution would be for governments to design a bespoke form of IP known as a *sui generis* law. Such custom-built laws are designed to cover types of creative output not addressed by the ‘big four’ IP doctrines of copyright, industrial designs, trademarks and patents. They already incentivize and protect investment in circuit layouts, new varieties of plants and, in some jurisdictions, databases.

Some critics might object to the mushrooming of topic-specific forms of IP. But a distinct AI-IP doctrine has the advantage that it could be tailored to meet the specific conditions in which AI creativity occurs. For example, lawmakers might decide that, if AI-IP is easier and faster to develop, it should be protected for a shorter period than the conventional 20-year term of standard patents. This would encourage others to build on inventions as soon as the patent term expires. And whereas patents are typically awarded to the inventor, lawmakers could decide to distribute the rewards from an AI-generated invention differently — perhaps between the AI developer, the person directing the AI and the owner of the data used to train it⁹.

International treaty

Countries where AI-IP is put in place would be likely to attract investment in research and development. On the flip side, royalties attached to the use of an invention could make it less available. The same situation occurs today with drugs and vaccines: patents can attract the investment that allows them to be developed, but people lose out in countries that are unable to afford the products or unable to pay royalties to manufacture them. For instance, patent restrictions on the manufacture of COVID-19 vaccines are among the reasons why, some 18 months after the vaccines first became available, only around 16% of people in low-income countries have received at least one dose. Meanwhile, some higher-income nations are offering fourth doses (see

Nature **603**, 764; 2022). In designing AI-IP, a balance must be found to avoid reproducing this sort of inequity.

A country that tends to import inventions might see benefits in not protecting AI-generated ones. It could then provide its population with cheap copies of an expensive new drug that an AI had invented elsewhere, instead of paying royalties. However, it might also miss out on attracting industry-building investment.

Patent treaties address this ‘free-rider problem’. The TRIPS agreement was set up partly to make it more difficult for countries to opt out of providing IP protection in the form of patents. International conventions also govern the use of designs, trademarks, copyright and various other areas of IP.

We think that an international treaty is essential for AI-generated inventions, too. It would set out uniform principles to protect AI-generated inventions in multiple jurisdictions. This could be done by negotiating a new treaty or adding those rules into an existing international IP agreement.

Establishing such a treaty would be an ambitious, long-term plan. Critics could object to nations relinquishing the freedom to make domestic policy about AI-generated inventions, especially when the full potential of AI is still unclear. We feel that the global benefits of an international agreement would make the time and cost of negotiating it worthwhile, because it would avoid uncertainties and disputes down the line.

An AI-IP treaty should also avoid a key limitation of the patent system. Patents are registered separately in each jurisdiction, and enforcement disputes must normally be resolved by the legal system of the country of registration. This can result in patent holders running similar legal cases in many countries — as when the technology firms Apple and Samsung spent 7 years battling more than 50 lawsuits about phone and tablet design and functionality (see go.nature.com/3lfzpej). Cumbersome and expensive, this system means that some people cannot afford to enforce or defend their rights under patent law.

Instead, an AI-IP treaty could include dispute-resolution mechanisms, perhaps adjudicated by a specialist international court. Inspiration can be

found in Europe's new Unified Patent Court, which is due to commence in the next year (www.unified-patent-court.org), and various arbitration courts around the world.

Creating bespoke law and an international treaty will not be easy, but not creating them will be worse. AI is changing the way that science is done and inventions are made. We need fit-for-purpose IP law to ensure it serves the public good.

Nature **605**, 616–618 (2022)

doi: <https://doi.org/10.1038/d41586-022-01391-x>

References

1. DABUS. Food Container and Devices and Methods for Attracting Enhanced Attention. Patent Cooperation Treaty international application PCT/IB2019/057809 (Filed 17 September 2019).
2. *Association for Molecular Pathology v. Myriad Genetics, Inc.* [2007] 569 US 576.
3. Stanková, E. *Cambridge Law J.* **80**, 338–365 (2021).
4. Abbott, R. *UCLA Law Rev.* **66**, 2–52 (2019).
5. *Commissioner of Patents v. Thaler* [2022] Federal Court of Australia (Full Court) 62, paras 119–120.
6. Abbott, R. *Washburn Law J.* **60**, 429–444 (2020).
7. Afshar, M. S. *Hastings Sci. Technol. Law J.* **13**, 55–72 (2022).
8. Kim, D. *GRUR Int.* **69**, 443–456 (2020).
9. Schuster, W. M. *Washington Lee Law Rev.* **75**, 1945–2004 (2018).

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

Open access in low-income countries — open letter on equity
[Download PDF](#)

- CORRESPONDENCE
- 24 May 2022

Open access in low-income countries — open letter on equity

- [Franco M. Cabrerizo](#) 0

As chair of the World Academy of Sciences Young Affiliates Network (TYAN), I am one of the organizers of an open letter describing the adverse impact on researchers in developing countries of article-processing charges for open-access publications (see go.nature.com/3jn1k6s). We call for a multilateral solution to the problem that will help the entire global community. By mid-May, the letter's signatory list included 17 Nobel laureates and more than 30 international organizations and academies.

Article-processing fees are deepening the inequalities between scientists from developed and developing countries in sharing scientific advances (see, for example, [T. Ross-Hellauer *Nature* 603, 363; 2022](#)). The international community must exercise multilateral governance and academic cooperation to ensure that open-access publication models promote equal opportunities for researchers worldwide.

TYAN outlines some first steps towards achieving this goal, and suggests how they could be initiated by a purpose-built committee.

Nature 605, 620 (2022)

doi: <https://doi.org/10.1038/d41586-022-01414-7>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01414-7>

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

Work

- **PhD students face cash crisis with wages that don't cover living costs** [23 May 2022]
Career Feature • As inflation rates soar, new data on the finances of US graduate students spark calls for action.
- **The overlooked variable in animal studies: why diet makes a difference** [23 May 2022]
Technology Feature • Careful consideration and documentation of laboratory animals' diets will boost the reproducibility of experiments.
- **Timbre in the timber: how I date ancient violins** [23 May 2022]
Where I Work • Forest ecologist Paolo Cherubini once helped to settle a legal dispute involving a viola supposedly made in the sixteenth century.

PhD students face cash crisis with wages that don't cover living costs
[Download PDF](#)

- CAREER FEATURE
- 23 May 2022

PhD students face cash crisis with wages that don't cover living costs

As inflation rates soar, new data on the finances of US graduate students spark calls for action.

- [Chris Woolston](#) ⁰

[Find a new job](#)



Student volunteers and workers stock the shelves and take an inventory at the CARE food pantry at Kennesaw State University in Georgia. Credit: Melissa Golden/Redux/eyevine

Salaries for PhD students in the biological sciences fall well below the basic cost of living at almost every institution and department in the United States, according to data collected by two PhD students.

The crowdsourced findings, submitted by students, faculty members and administrators and presented on an interactive dashboard, provide fresh ammunition for graduate students in negotiations for higher salaries as economies across the world grapple with rising inflation.

As this article went to press, just 2% of the 178 institutions and departments in the data set guaranteed graduate students salaries that exceed the cost of living. The researchers used the living-wage calculator maintained by the Cambridge-based Massachusetts Institute of Technology (see go.nature.com/3pkzjde), a widely used benchmark that estimates basic expenses for a given city, such as the costs of food, health care, housing and transport.

Most institutions fall far short of that standard. At the University of Florida in Gainesville, for example, the basic stipend for biology PhD students is around US\$18,650 for a 9-month appointment, about \$16,000 less than the annual living wage for a single adult in the city with no dependents. At a handful of institutions — including the University of Southern Mississippi in Hattiesburg and the University of South Dakota in Vermillion — the guaranteed minimum stipend is less than \$15,000 for 9-month appointments.

With US annual inflation now exceeding 8%, stipends haven't been keeping pace, says Michelle Gaynor, a fourth-year PhD student in evolutionary biology at the University of Florida. Gaynor created the dashboard along with Rhett Rautsaw, a fifth-year PhD student in evolutionary biology at Clemson University in South Carolina.

Costs are rising so quickly that some salaries that qualified as a living wage in early April had fallen below that mark by the end of the month. "It would be amazing if students could focus on the science and not whether they can

afford an apartment,” Gaynor says. “Not every grad student has that privilege.”

Salary shortfalls

Gaynor and Rautsaw started collecting the data after a study in March showed wide gaps between stipends and the cost of living for entomology students in 20 PhD programmes and 19 master’s degree programmes in the United States ([M. Kirchner & J. Petzoldt *Am. Entomol.* **68**, 22–23; 2022](#)). The study found that although students at three universities in the northeast generally earned living wages, students in other regions fared much worse. In the south, PhD students were nearly \$8,000 short of a living wage, on average. The average shortfall nationally was just over \$4,500.

Gaynor shared a plot of their initial results on Twitter, and the post quickly went viral. “I feel like I opened the floodgates,” she says. “Graduate students are telling me about their experiences. They’re struggling.”

The United States isn’t the only country where graduate stipends aren’t keeping pace with inflation. UK Research and Innovation, for example, Britain’s largest public funder of research, is set to increase the current minimum PhD stipend of £15,609 (US\$19,315) by 2.9% for the 2022–23 academic year. That’s less than half the UK rate of inflation, which currently stands at 9%.

Basic minimum stipends — essentially a guaranteed salary for a graduate student — are only one source of remuneration. Some students earn significantly more through fellowships. Gaynor notes, however, that US fellowships generally cover four years of training, yet it often takes at least five years to earn a PhD. When fellowships run out, a student might be forced to live on a guaranteed minimum wage that doesn’t come close to meeting needs.

The current state of PhD salaries threatens to undermine US efforts to increase diversity in science, Gaynor says. “If programmes aren’t meeting the basic cost of living, who are we selecting for? People who have financial support or external fellowships.”

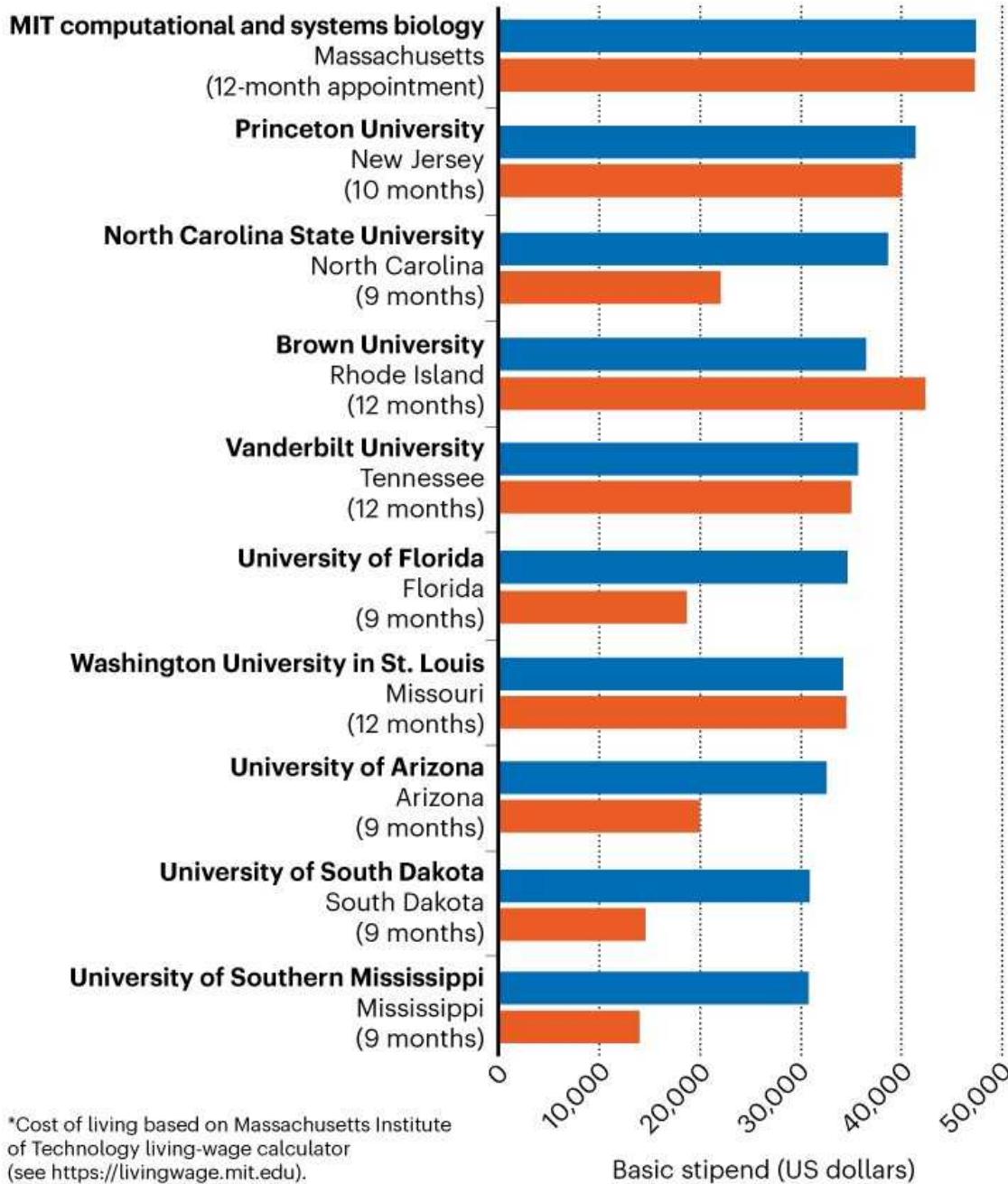
Some faculty members are expressing concerns about the current generation of students. “It was difficult when I was a grad student, but it’s far worse now,” says Akito Kawahara, an evolutionary biologist at the University of Florida. “Students can’t survive these days with the salaries that are out there. Students are leaving universities or not coming to universities because the base salaries are so low.”

The universities that do offer a living wage are true outliers. At Brown University in Providence, Rhode Island, for example, the basic annual stipend for biology PhDs exceeds \$42,000, well above the local cost of living of about \$36,000. Other institutions that pay a living wage (or close to it) include Princeton University in New Jersey; Washington University in St. Louis, Missouri; Vanderbilt University in Nashville, Tennessee; and the Massachusetts Institute of Technology’s department of computational and systems biology in Cambridge (see ‘Where US biology PhD students can — and can’t — get by’). In January, Princeton University announced that it would establish a minimum 10-month graduate-student stipend of \$40,000 — close to the local living wage — for all university departments for the 2022–23 academic year.

WHERE US BIOLOGY PHD STUDENTS CAN — AND CAN'T — GET BY

Some US universities guarantee stipends that cover the cost of living, but others fall well short.

- Cost of living
(1 adult, no children)*
- Guaranteed
minimum salary



*Cost of living based on Massachusetts Institute of Technology living-wage calculator (see <https://livingwage.mit.edu>).

©nature

Source: <https://rhettrautsaw.app/shiny/BiologyPhDStipends/>

About one-third of the stipends included in Gaynor and Rautsaw's database are for 9-month appointments and about 5% are 10-month appointments. Gaynor says that although some students have sources of income for the months not covered by their stipend, others have to stretch those funds to last the whole year. "Faculty often try their hardest to make sure you have a salary [for the entire year], but it's not guaranteed," she says. "I know people at other universities who were told they were going to have a summer salary and then they didn't."

Added value

Some administrators argue that students are receiving valuable training that will pay great dividends later, but such low wages are unfair and unsustainable, says Jane Petzoldt, a third-year master's student in entomology at North Carolina State University (NC State) in Raleigh, and a co-author of the entomology stipend study. "The value of our work far exceeds the stipend and graduate-support packages," she says.

Co-author Michelle Kirchner, a third-year PhD student in entomology at NC State, says their graduate-student contract prohibits taking on a second job to make ends meet. "How am I supposed to support myself?" she asks. "The university should be required to pay the cost of living."

Last December, to better understand the situation, Petzoldt helped to conduct a survey of 35 students in the entomology and plant pathology department at NC State. About one-quarter of respondents said they had broken the rules of their contract by taking a second job, and 15% reported trouble in affording groceries (see go.nature.com/3wey9gj). "The financial stress is cutting into our ability to perform as graduate students," Petzoldt says.

A spokesperson for NC State noted that many US universities prohibit outside employment for graduate students, adding: "Historically, these graduate-student appointments are akin to scholarships that provide graduate students a stipend and cover the costs of tuition and health insurance. Although many assistantships at NC State reach and exceed a living wage, some are lower than we'd prefer."

Some universities are keenly aware of the financial straits their trainees are in. In a survey published last September by the Graduate College of the University of Arizona in Tucson, more than 80% of the nearly 2,500 respondents, representing departments across the university, reported feeling stressed about their financial situation. One-quarter reported using a food bank, and 36% reported taking out unanticipated loans since starting their graduate programme (see go.nature.com/3jvpg4v).

In comments, many respondents highlighted the gap between stipends and the cost of living. One wrote that “graduate students are needlessly suffering and the university can do far more than it is doing to lessen the burden”. Another wrote: “Pay your graduate students a liveable salary! Getting a doctorate is hard enough … [worrying] about finances makes this experience loathsome and diminishes my opinion of this university.” Gaynor and Rautsaw’s tracker shows the minimum stipend for biology PhDs at the University of Arizona is \$20,000 for a 9-month appointment, which is about \$12,500 less than the estimated cost of living for a single adult in Tucson. The university conducted its 2021 survey after hearing widespread complaints from graduate students, says Andrew Carnie, a vice-provost at the university and a co-author of the survey. “Financial health is so clearly tied to student health, productivity, retention and recruitment,” he says. “It is something that no university can ignore. It requires critical action.”

Addressing the shortfalls

Carnie thinks that student stipends need to be higher, but takes some issue with the figures in the tracker, notably that the \$20,000 stipend is for a 9-month appointment. “Most biological-sciences students are actually on 12-month stipends, for which the \$20,000 rate is prorated to about \$27,000,” he says. He adds that the university also reduces the cost of living by providing health insurance to graduate students.



Signs for home rentals are posted on the campus at the University of California, Berkeley. Credit: Justin Sullivan/Getty

The University of Arizona announced last year that every department will pay a minimum stipend of \$20,000, starting in the 2022–23 academic year. . Previously, says Carnie, the average base stipend for a 9-month appointment was around \$17,000. He adds that the university encourages departments to pay beyond the minimum when possible. The base level is now tied to a government cost index, so students shouldn't fall farther behind as prices rise. "With the current rates of inflation, I expect that this minimum will likely increase substantially over the next few years," he says.

Carnie says the university is taking other steps to help students cope with costs. "We are currently advocating at the state level for improved health-care coverage, including dental, spousal and dependent care," he says. "We are also working with our institutional partners to create new opportunities for grad student and family housing, and working to find low-cost childcare. These are priority issues for us."

Pay negotiations

Last November, NC State students in the department of entomology and plant pathology successfully negotiated for a rise, partly with the help of the data collected by Petzoldt and Kirchner. After accounting for fees, the new minimum take-home pay will be \$28,400, a \$2,400 increase from the current base stipend of \$26,000. The current stipend is about \$12,600 below the living wage for the area . Details of the rise, which will go into effect in July, are not yet publicly available, but Petzoldt says it's \$1,000 more than the university originally offered. The current stipend is \$26,000, which is about \$12,600 below the living wage for the area. Petzoldt says they've heard from entomology students at several other institutions who plan to use the results in their negotiations.

In April, graduate students at the University of Florida started negotiations for a rise of their own. On 29 April, the university offered to boost the base 9-month salary for all departments from \$16,000 to \$17,000. That offer, which the university called “final”, has not yet been accepted, and negotiations continue. A spokesperson told *Nature* that the university couldn’t comment on graduate-student stipends because discussions are still under way.

Kawahara joined an early bargaining session at which, he says, there were roughly 100 students present, but he didn’t notice any other faculty members. “I’ve had students who also worked for DoorDash [a food-delivery company],” he says.

The negotiations were “eye-opening”, Kawahara says. “It was basically the students versus the university, and the university was very rigid.”

Kawahara says that individual faculty members can take steps to support PhD students, such as by letting them know about available fellowships. But such piecemeal efforts can lead to disparities, he says. “You can have an unfair situation where one student is making \$16,000 and another student makes close to \$30,000,” he says. In his view, the only way to avoid such discrepancies is to make liveable wages standard across the university.

Gaynor is keen to get salary details for more US biology departments. “The more input we get, the better the data set becomes,” she says.

Nature **605**, 775-777 (2022)

doi: <https://doi.org/10.1038/d41586-022-01392-w>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01392-w>

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

The overlooked variable in animal studies: why diet makes a difference
[Download PDF](#)

- TECHNOLOGY FEATURE
- 23 May 2022

The overlooked variable in animal studies: why diet makes a difference

Careful consideration and documentation of laboratory animals' diets will boost the reproducibility of experiments.

- [Jyoti Madhusoodanan](#) ⁰



The behaviour of cuttlefish and other cephalopods can be influenced by their diets. Credit: Adam Glanzman/The Washington Post/Getty

In the 1980s, two groups of researchers asked a seemingly simple question: can animals live longer by eating less? A team at the US National Institutes of Health (NIH) in Bethesda, Maryland, and another at the University of Wisconsin–Madison each fed rhesus monkeys (*Macaca mulatta*) diets with 30% fewer calories than given to their control animals. The Wisconsin-based team found that caloric restriction helped the monkeys to live longer, healthier lives^{1,2}; the NIH team observed no such effect³.

To resolve the disparity, the researchers looked into the experimental set-ups. They found that even though both studies looked at the same level of caloric restriction, the specifics of the animals' diets could have contributed to the differing outcomes⁴. "Diet is just another variable," says Kristin Gribble, a molecular biologist at the Marine Biological Laboratory in Woods Hole, Massachusetts. "If it's different between experiments, it's an additional variable to be considered when explaining the results."

Researchers in the past often overlooked the importance of an animal's diet when designing experiments, says Stephen Watts, an aquatic-nutrition researcher at the University of Alabama at Birmingham. "If the animals looked happy and seemed okay, [the researchers] were okay with it," he says.

That started to change in 1977, when a report from the American Institute for Nutrition in Rockville, Maryland, set out guidelines for ways to eliminate confounding dietary factors in medical research⁵. Scientists then went on to develop multiple standardized diets for animal facilities and laboratories, leading to the production of various standardized pellet foods for laboratory rats and mice. "It became very clear that nutrition was going to be one of the key components in improving rigour and reproducibility," Watts says.

But those diets still encompass a lot of variation. And they don't exist for many common animal models. Several were developed for reasons of cost and practicality, rather than to mimic an animal's natural habits. So careful documentation of the experimental set-up is crucial to improving the replicability of experiments, says David Allison, a biostatistician at Indiana University Bloomington who worked on teasing out the differences between the NIH and Wisconsin primate studies. Including details about the diet used

in a study can help future researchers to “decide whether they want to replicate it and how to interpret the results”, he says.

Shortfalls of standardization

Across various species, researchers have found that variations in food can yield unexpected results, ranging from small deviations that “don’t really inform the science” to data with profound implications for the conclusions, says Allison. For instance, in one long-running study Allison was involved in, rats developed bladder stones owing to an adulterant in some batches of food bought from a commercial supplier. That meant that they didn’t have animals from which to draw statistically significant conclusions and had to go back to the drawing board to work out why it happened and fix the problem before continuing the experiment. And in another, they had a similar situation when the animals developed skin lesions because of excess vitamin A. Although these issues didn’t affect Allison’s conclusions, the observations “made our experiments not as pristine as we wished”, he says. “It’s an issue that often comes up in longevity studies, because the animal is being fed the same foods for so long.”

Even in short-term studies — and even with standardized diets — experimental complications crop up. In 2018, dietitian Caroline Tuck at La Trobe University in Melbourne, Australia, was studying short-chain fermentable carbohydrates known as FODMAPs, which can exacerbate conditions such as irritable bowel syndrome in humans. When researching diets with different FODMAP levels in mouse models, she realized that standardized laboratory diets vary widely in their levels of protein and micro- and macronutrients. “Standard chow can be so many different things,” says team member Giada De Palma, a microbiome researcher at McMaster University in Hamilton, Canada. To assess the impact of the various foods, the team fed four groups of mice different commercial foods for three weeks, and found significant differences in the microbial diversity in the caecum⁶, the main intestinal site of food fermentation in mice.



The specific composition of mouse chow can affect the rodents' behaviour and physiology. Credit: Hemant Mishra/Mint/Getty

Tuck and De Palma emphasize that no one diet was better than the others. But, given the differences, they recommend that scientists using pelleted feeds consider whether the composition might affect their research question. “It’s really about considering it at the design phase of an experiment,” says Tuck, “rather than as an afterthought or just using whatever is routinely used in the lab.”

Furthermore, Tuck says, researchers should report the specifics in their studies. “When publishing, we talk about when the animals had access to food and water, but actually listing exactly what the diet was is really important as well.”

Careful documentation also helped to solve the question of whether caloric restriction helped rhesus monkeys to live longer. When retracing their steps, researchers at the NIH and at Wisconsin found that the contrasting findings could have been caused, in part, by the quality of the food, the timing of administration and the animals’ own choices⁴. For example, monkeys at the

NIH had access to food throughout the night, whereas those in Wisconsin had their afternoon fruit snack removed at the end of the day.

Embracing variety

Another route to improving reproducibility is to focus on animal welfare and encompass more variety in the diet, says Robyn Crook, a behavioural neuroscientist at San Francisco State University in California. “There’s a lot of value that comes from not having a standardized diet,” she says.

Consistent data from animals that “are homogeneously extremely stressed and deprived” are not useful to research. “There’s only so much you can standardize about the life of a rat or a mouse before you start to negatively impact welfare. And then you negatively impact science as well,” she says. Gribble agrees. “Some lab diets are not even close to what those animals would be eating in the wild,” she says. “It’s not the nutritional profile that those organisms have adapted to, so that alone could change results.”

For example, researchers typically feed the nematode *Caenorhabditis elegans* with the bacterium *Escherichia coli*. But that’s because *E. coli* is easy to grow, not what the worms prefer to eat. In a 2020 study⁷, researchers routinely spotted *C. elegans* eating spontaneously occurring bacterial contaminants, and found that those species also occur in the nematodes’ natural environment. They also found that raising *C. elegans* on those species altered processes such as gene expression, development and reproduction when compared with feeding them on standard *E. coli* strains. When offered a buffet of six bacterial strains, the worms actively avoided some and gravitated towards others.

In her lab, Gribble feeds microscopic aquatic animals known as rotifers algae and phytoplankton. The rotifers’ traits vary depending on how their food was grown, she says: phytoplankton grown in high-light conditions tend to produce more lipid and less protein than do those grown in low-light conditions. These variations can influence the experimental results.

Similar changes in shrimp and other small aquatic species can cause variations in the animals that feed on them, such as octopus and squid. Crook says that when she keeps squid in the lab for extended periods of

time, she also needs to maintain cultures of shrimp to feed the cephalopod. It's possible to standardize the squid's diet by controlling what its prey are fed, she says. But many cephalopods are tough to rear in labs, so some researchers rely on wild-caught animals — and wild-caught food for those animals, she adds. When running shorter experiments with octopus, her team uses wild-caught foods or live fish for reasons of cost and practicality. Crook's research focuses on the neuroscience of pain, not on food or on animal husbandry, "but you can't really work with cephalopods without engaging with those questions", she says. "They're fundamental to animal health." Crook encourages researchers to think about an animal's natural history when designing experiments. One of the central things about good animal welfare is allowing the animal to have some control, she says. "So why not give choices in diet?"

And a focus on animal welfare can boost scientific rigour, Allison says. Then, he says, "it's much more likely that we'll find effects that hold up under a broad range of circumstances."

Nature **605**, 778–779 (2022)

doi: <https://doi.org/10.1038/d41586-022-01393-9>

References

1. Colman, R. J. *et al.* *Science* **325**, 201–204 (2009).
2. Colman, R. J. *et al.* *Nature Commun.* **5**, 3557 (2014).
3. Mattison, J. A. *et al.* *Nature* **489**, 318–321 (2012).
4. Mattison, J. A. *et al.* *Nature Commun.* **8**, 14063 (2017).
5. American Institute of Nutrition Ad Hoc Committee on Standards for Nutritional Studies. *J. Nutr.* **107**, 1340–1348 (1977).
6. Tuck, C. J. *et al.* *Sci. Rep.* **10**, 17784 (2020).
7. Stuhr, N. L. & Curran, S. P. *Commun. Biol.* **3**, 653 (2020).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01393-9>

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

Timbre in the timber: how I date ancient violins

[Download PDF](#)

- WHERE I WORK
- 23 May 2022

Timbre in the timber: how I date ancient violins

Forest ecologist Paolo Cherubini once helped to settle a legal dispute involving a viola supposedly made in the sixteenth century.

- [Nic Fleming](#)



Paolo Cherubini is a senior scientist at the Swiss Federal Institute for Forest, Snow and Landscape Research in Zurich, Switzerland. Credit: Elisabetta Zavoli for *Nature*

Stringed instruments can be worth millions of dollars, particularly those made in northern Italy in the seventeenth and eighteenth centuries. There are also many copies and forgeries.

As a forest ecologist, I use dendrochronology — or tree-ring dating — to understand how trees grow, as well as to investigate historical environmental conditions. The widths of tree rings vary according to meteorological conditions, so samples can be dated by cross-referencing against databases of ring-width series.

In 2010, I was an expert witness in a legal case about a viola supposedly made in the sixteenth century. I agreed with two laboratories that had independently concluded that it could not have been made before the late eighteenth century.

Dendrochronology cannot precisely date when an instrument was made, but it can identify the most recent year that the wood it was made from was part of a growing tree. Tree rings give probabilities and levels of confidence in a date according to the availability of appropriate reference series.

Dendrochronological analyses can produce conflicting dates, creating confusion about the method's reliability. The most famous example is the 'Messiah' violin made by Antonio Stradivari in 1716. In 1999, doubts were raised about its authenticity when, among other reasons, tree-ring analysis found that the instrument could not have been made before 1738, a year after Stradivari's death. The finding was based on examining a photograph of the instrument, and was later retracted. The controversy highlighted the limitations of the technique for verifying the authenticity of musical instruments.

I also use dendrochronology to understand the impacts of droughts, and I'm fascinated by how trees grow, how water is transported in them and whether species differences are evolutionary adaptations.

Nature **605**, 782 (2022)

doi: <https://doi.org/10.1038/d41586-022-01394-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-01394-8>

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

Research

- **Defining the risk of SARS-CoV-2 variants on immune protection** [31 March 2022]

Perspective • The SARS-CoV-2 Assessment of Viral Evolution (SAVE) programme provides a real-time risk assessment of SARS-CoV-2 variants with the potential to affect transmission, virulence and resistance to infection- and vaccine-induced immunity.

- **In situ recording of Mars soundscape** [01 April 2022]

Article • Using data gathered from the microphones of the Perseverance rover, the first characterization of the acoustic environment on Mars is presented, showing two distinct values for the speed of sound in CO₂-dominated atmosphere.

- **Qubit teleportation between non-neighbouring nodes in a quantum network** [25 May 2022]

Article • A quantum network formed by three optically connected nodes comprising solid-state qubits demonstrates the teleportation of quantum information between two non-neighbouring nodes, negating the need for a direct connection between them.

- **Enhanced silica export in a future ocean triggers global diatom decline** [25 May 2022]

Article • Mesocosm experiments in different biomes show that future ocean acidification will slow down the dissolution of biogenic silica, decreasing silicic acid availability in the surface ocean and triggering a global decline of diatoms as revealed by Earth system model simulations.

- **Extricating human tumour immune alterations from tissue inflammation** [11 May 2022]

Article • Complementary single-cell approaches show that a population of regulatory T cells co-expressing ICOS and IL-1 receptor type 1 is highly enriched in tumours but not in site-matched inflamed non-malignant tissue.

- **Intermittent PI3K δ inhibition sustains anti-tumour immunity and curbs irAEs** [04 May 2022]

Article • A study using a mouse solid tumour model suggests that adjusting the dosing frequency of the PI3K δ inhibitor AMG319 in the treatment of human cancers could decrease tumour growth with fewer adverse effects.

[Download PDF](#)

- Perspective
- [Published: 31 March 2022](#)

Defining the risk of SARS-CoV-2 variants on immune protection

- [Marciela M. DeGrace](#)^{1,2 na1},
- [Elodie Ghedin](#) ORCID: orcid.org/0000-0002-1515-725X^{1,3 na1},
- [Matthew B. Frieman](#) ORCID: orcid.org/0000-0003-0107-0775^{4 na1},
- [Florian Krammer](#) ORCID: orcid.org/0000-0003-4121-776X^{5,6 na1},
- [Alba Grifoni](#) ORCID: orcid.org/0000-0002-2209-5966^{7 na1},
- [Arghavan Alisoltani](#)⁸,
- [Galit Alter](#) ORCID: orcid.org/0000-0002-7680-9215⁹,
- [Rama R. Amara](#) ORCID: orcid.org/0000-0002-6309-6797¹⁰,
- [Ralph S. Baric](#) ORCID: orcid.org/0000-0001-6827-8701^{11,12},
- [Dan H. Barouch](#) ORCID: orcid.org/0000-0001-5127-4659¹³,
- [Jesse D. Bloom](#) ORCID: orcid.org/0000-0003-1267-3408¹⁴,
- [Louis-Marie Bloyet](#) ORCID: orcid.org/0000-0002-5648-3190¹⁵,
- [Gaston Bonenfant](#)¹⁶,
- [Adrianus C. M. Boon](#) ORCID: orcid.org/0000-0002-4700-8224¹⁷,
- [Eli A. Boritz](#)^{1,2,18},
- [Debbie L. Bratt](#) ORCID: orcid.org/0000-0002-5822-5558^{1,2,19},
- [Traci L. Bricker](#)¹⁷,
- [Liliana Brown](#)^{1,2},
- [William J. Buchser](#) ORCID: orcid.org/0000-0002-6675-6359²⁰,
- [Juan Manuel Carreño](#)⁵,
- [Liel Cohen-Lavi](#) ORCID: orcid.org/0000-0001-6909-4779²¹,
- [Tamarand L. Darling](#)¹⁷,
- [Meredith E. Davis-Gardner](#)²²,
- [Bethany L. Dearlove](#) ORCID: orcid.org/0000-0003-3653-4592²³,

- [Han Di](#)¹⁶,
- [Meike Dittmann](#)²⁴,
- [Nicole A. Doria-Rose](#)^{1,18},
- [Daniel C. Douek](#)^{1,18},
- [Christian Drosten](#) [ORCID: orcid.org/0000-0001-7923-0519](#)²⁵,
- [Venkata-Viswanadh Edara](#) [ORCID: orcid.org/0000-0001-9321-7839](#)²²,
- [Ali Ellebedy](#) [ORCID: orcid.org/0000-0002-6129-2532](#)²⁶,
- [Thomas P. Fabrizio](#) [ORCID: orcid.org/0000-0002-8960-0728](#)²⁷,
- [Guido Ferrari](#)²⁸,
- [Will M. Fischer](#)⁴¹,
- [William C. Florence](#)^{1,2},
- [Ron A. M. Fouchier](#) [ORCID: orcid.org/0000-0001-8095-2869](#)²⁹,
- [John Franks](#)²⁷,
- [Adolfo García-Sastre](#) [ORCID: orcid.org/0000-0002-6551-1827](#)^{5,6,30,31,32},
- [Adam Godzik](#) [ORCID: orcid.org/0000-0002-2425-852X](#)⁸,
- [Ana Silvia Gonzalez-Reiche](#) [ORCID: orcid.org/0000-0003-3583-4497](#)³³,
- [Aubree Gordon](#) [ORCID: orcid.org/0000-0002-9352-7877](#)³⁴,
- [Bart L. Haagmans](#) [ORCID: orcid.org/0000-0001-6221-2015](#)²⁹,
- [Peter J. Halfmann](#)³⁵,
- [David D. Ho](#) [ORCID: orcid.org/0000-0003-1627-149X](#)³⁶,
- [Michael R. Holbrook](#) [ORCID: orcid.org/0000-0002-0824-2667](#)³⁷,
- [Yaoxing Huang](#) [ORCID: orcid.org/0000-0001-6270-1644](#)³⁶,
- [Sarah L. James](#) [ORCID: orcid.org/0000-0002-6969-1167](#)³⁸,
- [Lukasz Jaroszewski](#)⁸,
- [Trushar Jeevan](#)²⁷,
- [Robert M. Johnson](#) [ORCID: orcid.org/0000-0002-1976-7688](#)⁴,
- [Terry C. Jones](#) [ORCID: orcid.org/0000-0003-1120-9531](#)^{25,38},
- [Astha Joshi](#) [ORCID: orcid.org/0000-0003-4914-8228](#)¹⁷,
- [Yoshihiro Kawaoka](#) [ORCID: orcid.org/0000-0001-5061-8296](#)^{35,39,40},
- [Lisa Kercher](#) [ORCID: orcid.org/0000-0001-6300-0452](#)²⁷,
- [Marion P. G. Koopmans](#) [ORCID: orcid.org/0000-0002-5204-2312](#)²⁹,
- [Bette Korber](#) [ORCID: orcid.org/0000-0002-2026-5757](#)⁴¹,
- [Eilay Koren](#)^{21,42},

- [Richard A. Koup^{1,18}](#),
- [Eric B. LeGresley](#) [ORCID: orcid.org/0000-0002-5286-5693³⁸](#),
- [Jacob E. Lemieux⁴³](#),
- [Mariel J. Liebeskind](#) [ORCID: orcid.org/0000-0003-4595-0651²⁰](#),
- [Zhuoming Liu](#) [ORCID: orcid.org/0000-0001-8198-0976¹⁵](#),
- [Brandi Livingston²⁷](#),
- [James P. Logue](#) [ORCID: orcid.org/0000-0002-7410-9741⁴](#),
- [Yang Luo](#) [ORCID: orcid.org/0000-0003-3277-8792³⁶](#),
- [Adrian B. McDermott](#) [ORCID: orcid.org/0000-0003-0616-9117^{1,18}](#),
- [Margaret J. McElrath⁴⁴](#),
- [Victoria A. Meliopoulos](#) [ORCID: orcid.org/0000-0003-1442-9177²⁷](#),
- [Vineet D. Menachery](#) [ORCID: orcid.org/0000-0001-8803-7606⁴⁵](#),
- [David C. Montefiori⁴⁶](#),
- [Barbara Mühlemann^{25,38}](#),
- [Vincent J. Munster](#) [ORCID: orcid.org/0000-0002-2288-3196⁴⁷](#),
- [Jenny E. Munt¹¹](#),
- [Manoj S. Nair](#) [ORCID: orcid.org/0000-0002-5994-3957³⁶](#),
- [Antonia Netzl](#) [ORCID: orcid.org/0000-0001-8034-2382³⁸](#),
- [Anna M. Niewiadomska⁴⁸](#),
- [Sijy O'Dell^{1,18}](#),
- [Andrew Pekosz](#) [ORCID: orcid.org/0000-0003-3248-1761⁴⁹](#),
- [Stanley Perlman](#) [ORCID: orcid.org/0000-0003-4213-2354⁵⁰](#),
- [Marjorie C. Pontelli¹⁵](#),
- [Barry Rockx](#) [ORCID: orcid.org/0000-0003-2463-027X²⁹](#),
- [Morgane Rolland](#) [ORCID: orcid.org/0000-0003-3650-8490²³](#),
- [Paul W. Rothlauf](#) [ORCID: orcid.org/0000-0002-0941-4467¹⁵](#),
- [Sinai Sacharen^{21,42}](#),
- [Richard H. Scheuermann](#) [ORCID: orcid.org/0000-0003-1355-892X⁴⁸](#),
- [Stephen D. Schmidt^{1,18}](#),
- [Michael Schotsaert](#) [ORCID: orcid.org/0000-0003-3156-3132^{5,32}](#),
- [Stacey Schultz-Cherry](#) [ORCID: orcid.org/0000-0002-2021-727X²⁷](#),
- [Robert A. Seder](#) [ORCID: orcid.org/0000-0003-3133-0849^{1,18}](#),
- [Mayya Sedova⁸](#),
- [Alessandro Sette^{7,51}](#),

- [Reed S. Shabman](#) [ORCID: orcid.org/0000-0003-3272-3484^{1,2}](#),
- [Xiaoying Shen²⁸](#),
- [Pei-Yong Shi](#) [ORCID: orcid.org/0000-0001-5553-1616⁵²](#),
- [Maulik Shukla^{53,54}](#),
- [Viviana Simon](#) [ORCID: orcid.org/0000-0002-6416-5096^{5,6,30,32}](#),
- [Spencer Stumpf¹⁵](#),
- [Nancy J. Sullivan^{1,18}](#),
- [Larissa B. Thackray](#) [ORCID: orcid.org/0000-0002-9380-6569¹⁷](#),
- [James Theiler⁴¹](#),
- [Paul G. Thomas](#) [ORCID: orcid.org/0000-0001-7955-0256⁵⁵](#),
- [Sanja Trifkovic](#) [ORCID: orcid.org/0000-0002-0710-9514²⁷](#),
- [Sina Türeli](#) [ORCID: orcid.org/0000-0001-7342-9295³⁸](#),
- [Samuel A. Turner³⁸](#),
- [Maria A. Vakaki²⁰](#),
- [Harm van Bakel](#) [ORCID: orcid.org/0000-0002-1376-6916³³](#),
- [Laura A. VanBlargan](#) [ORCID: orcid.org/0000-0002-8922-8946¹⁷](#),
- [Leah R. Vincent](#) [ORCID: orcid.org/0000-0001-9262-1813^{1,2}](#),
- [Zachary S. Wallace](#) [ORCID: orcid.org/0000-0003-0237-501X^{48,56}](#),
- [Li Wang¹⁶](#),
- [Maple Wang³⁶](#),
- [Pengfei Wang](#) [ORCID: orcid.org/0000-0003-2454-7652³⁶](#),
- [Wei Wang⁵⁷](#),
- [Scott C. Weaver](#) [ORCID: orcid.org/0000-0001-8016-8556⁴⁵](#),
- [Richard J. Webby](#) [ORCID: orcid.org/0000-0002-4397-7132²⁷](#),
- [Carol D. Weiss](#) [ORCID: orcid.org/0000-0002-9965-1289⁵⁷](#),
- [David E. Wentworth](#) [ORCID: orcid.org/0000-0002-5190-980X¹⁶](#),
- [Stuart M. Weston](#) [ORCID: orcid.org/0000-0001-9840-2953⁴](#),
- [Sean P. J. Whelan](#) [ORCID: orcid.org/0000-0003-1564-8590¹⁵](#),
- [Bradley M. Whitener](#) [ORCID: orcid.org/0000-0001-6652-0701¹⁷](#),
- [Samuel H. Wilks³⁸](#),
- [Xuping Xie](#) [ORCID: orcid.org/0000-0003-0918-016X⁵²](#),
- [Baoling Ying¹⁷](#),
- [Hyejin Yoon](#) [ORCID: orcid.org/0000-0002-3344-9096⁴¹](#),
- [Bin Zhou¹⁶](#),

- [Tomer Hertz](#) ORCID: [orcid.org/0000-0002-0561-1578⁵⁸](https://orcid.org/0000-0002-0561-1578),
- [Derek J. Smith](#) ORCID: [orcid.org/0000-0002-2393-1890³⁸](https://orcid.org/0000-0002-2393-1890),
- [Michael S. Diamond](#) ORCID: [orcid.org/0000-0002-8791-3165^{17,59,60,61}](https://orcid.org/0000-0002-8791-3165),
- [Diane J. Post](#) ORCID: [orcid.org/0000-0003-3890-9116^{1,2}](https://orcid.org/0000-0003-3890-9116) &
- [Mehul S. Suthar](#) ORCID: [orcid.org/0000-0002-2686-8380²²](https://orcid.org/0000-0002-2686-8380)

Nature volume **605**, pages 640–652 (2022)

- 34k Accesses
- 2 Citations
- 316 Altmetric
- [Metrics details](#)

Subjects

- [Vaccines](#)
- [Viral infection](#)

Abstract

The global emergence of many severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants jeopardizes the protective antiviral immunity induced after infection or vaccination. To address the public health threat caused by the increasing SARS-CoV-2 genomic diversity, the National Institute of Allergy and Infectious Diseases within the National Institutes of Health established the SARS-CoV-2 Assessment of Viral Evolution (SAVE) programme. This effort was designed to provide a real-time risk assessment of SARS-CoV-2 variants that could potentially affect the transmission, virulence, and resistance to infection- and vaccine-induced immunity. The SAVE programme is a critical data-generating component of the US Government SARS-CoV-2 Interagency Group to assess implications of SARS-CoV-2 variants on diagnostics, vaccines and therapeutics, and for communicating public health risk. Here we describe the coordinated

approach used to identify and curate data about emerging variants, their impact on immunity and effects on vaccine protection using animal models. We report the development of reagents, methodologies, models and notable findings facilitated by this collaborative approach and identify future challenges. This programme is a template for the response to rapidly evolving pathogens with pandemic potential by monitoring viral evolution in the human population to identify variants that could reduce the effectiveness of countermeasures.

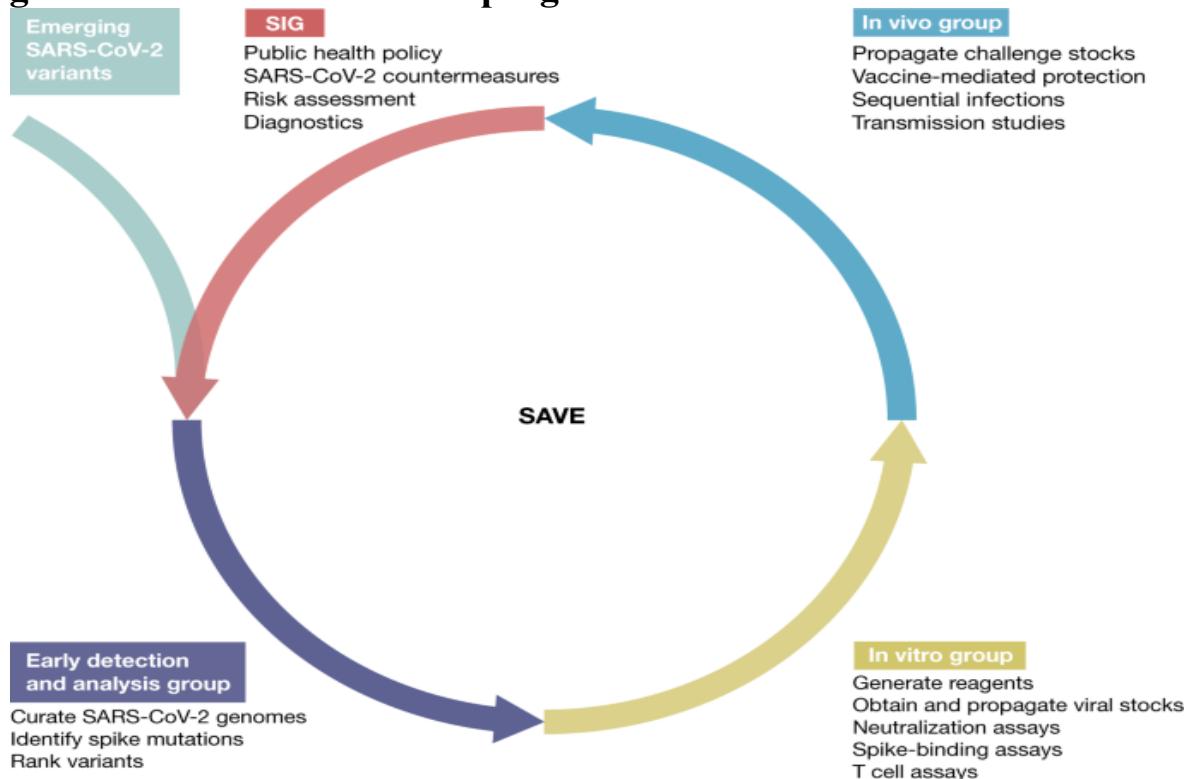
Main

SARS-CoV-2, the aetiological agent of coronavirus disease 2019 (COVID-19), has caused a devastating pandemic resulting in more than 6 million deaths worldwide (<https://covid19.who.int>). With continuous transmission cycles occurring around the world, SARS-CoV-2 variants have arisen with mutations throughout its genome, including in the spike protein gene, the principal antigenic target of all SARS-CoV-2 vaccines currently in use^{1,2}. The rapid emergence of variants—the latest being Omicron in November 2021—has raised concerns about how new mutations affect virus replication, infectivity, transmission and infection, and vaccine-induced immunity. This rapid genetic evolution of SARS-CoV-2 created an immediate need to monitor and characterize variants for potential resistance to medical countermeasures.

The US Department of Health and Human Services established the SARS-CoV-2 Interagency Group (SIG) to maximize coordination between the Centers for Disease Control and Prevention, the National Institutes of Health (NIH), the Food and Drug Administration, the Biomedical Advanced Research and Development Authority and Department of Defense for the US public health response to the COVID-19 pandemic³. The National Institute of Allergy and Infectious Diseases (NIAID) formed the SAVE consortium in January 2021 as a critical data-generating component for the SIG and to facilitate rapid data sharing with global partners and the scientific community (Fig. 1). The SAVE programme provides a comprehensive real-time risk assessment of emerging mutations in SARS-CoV-2 strains that could affect transmissibility, virulence, and infection- or vaccine-induced immunity. SAVE was constructed as a rational, structured and iterative risk-

assessment pipeline with a goal of providing critical data to support SIG actions and ensure the effectiveness of countermeasures against emerging variants.

Fig. 1: Overview of the SAVE programme.



The SAVE programme is divided into three working groups to provide real-time risk assessments of SARS-CoV-2 variants on infection and vaccine-induced immunity. The early-detection and analysis group curates and prioritizes emerging SARS-CoV-2 variants. The in vitro group evaluates the effect of SARS-CoV-2 variants on humoral and cell-mediated immune responses. The in vivo group uses animal models to test vaccine efficacy, transmission, and define immune mechanisms and correlates of protection. These data are fed into the SIG, which coordinates between different US government agencies to assess the impact of variants on critical SARS-CoV-2 countermeasures, including vaccines, therapeutics and diagnostics. This iterative approach allows for information flow between the SAVE programme and the SIG to continue prioritizing and testing SARS-CoV-2 variants.

The SAVE programme is composed of an international team of scientists with expertise in virology, immunology, vaccinology, structural biology, bioinformatics, viral genetics and evolution. Each team member is responsible for key contributions ranging from curation of viral mutations, bioinformatics analysis, development of new reagents, assay development and testing, in vitro characterization, and in vivo model development and countermeasure testing. The SAVE programme is divided into three working groups: (1) the early-detection and analysis group; (2) the in vitro group; and (3) the in vivo group. The early-detection group uses public databases and analysis tools to curate and prioritize emerging SARS-CoV-2 variants. The in vitro group evaluates the impact of SARS-CoV-2 variants on humoral and cell-mediated immune responses using in vitro assays. The in vivo group uses small and large animal models to test vaccine efficacy, transmission, and define immune mechanisms and correlates of protection. A common theme across these subgroups is the integration of orthogonal experimental and computational approaches to validate findings and strengthen the evidence for recommendations. Collaborative efforts between the early-detection geneticists and evolutionary biologists, and the in vitro group virologists/immunologists enable the rapid determination of relationships between viral evolution and neutralization sensitivity. In turn, these results enable the in vivo team to assess and evaluate vaccine protection in animal studies. The SAVE programme has regularly scheduled (usually weekly) meetings that include individual subgroup meetings and an all-hands meeting, which serves as an opportunity to share key information across groups and align priorities for the most urgent experimental questions. NIAID programme staff and intramural and extramural scientists share leadership responsibilities. Collaboration within and across these groups has accelerated research and discovery due to the immediate and open sharing of ideas, reagents, protocols and data^{4,5,6,7,8,9,10,11,12}. The SAVE group routinely invites scientists from international sites to present a real-time assessment of SARS-CoV-2 variants and infections within their region. The SAVE group coordinates with the Biodefense and Emerging Infections (BEI) Research Resources Repository, the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA) and the World Health Organization (WHO) to distribute SARS-CoV-2 isolates, proteins and plasmids. The SAVE group also has an open-face sharing policy in which findings are quickly disseminated through preprint servers while manuscripts undergo

formal peer review. The head-to-head comparison, review and discussion of unpublished data has yielded real-time peer review that would otherwise take months to achieve.

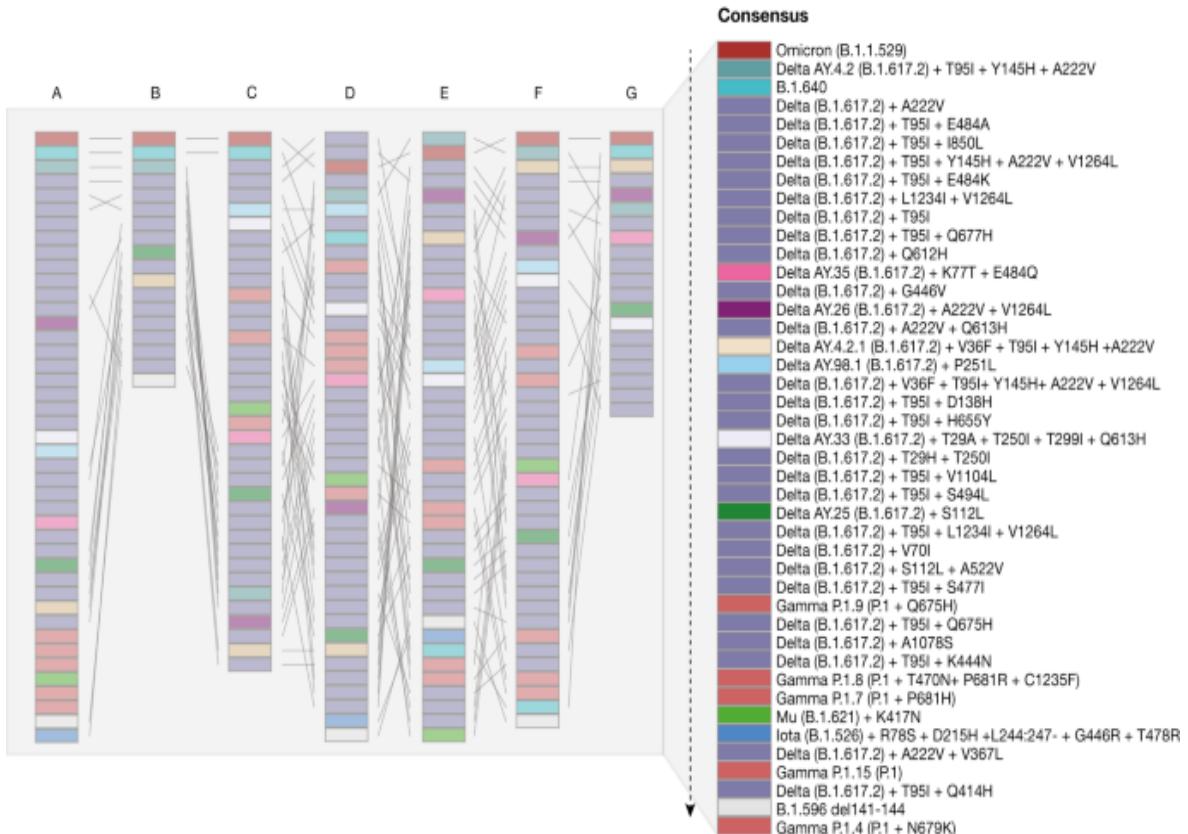
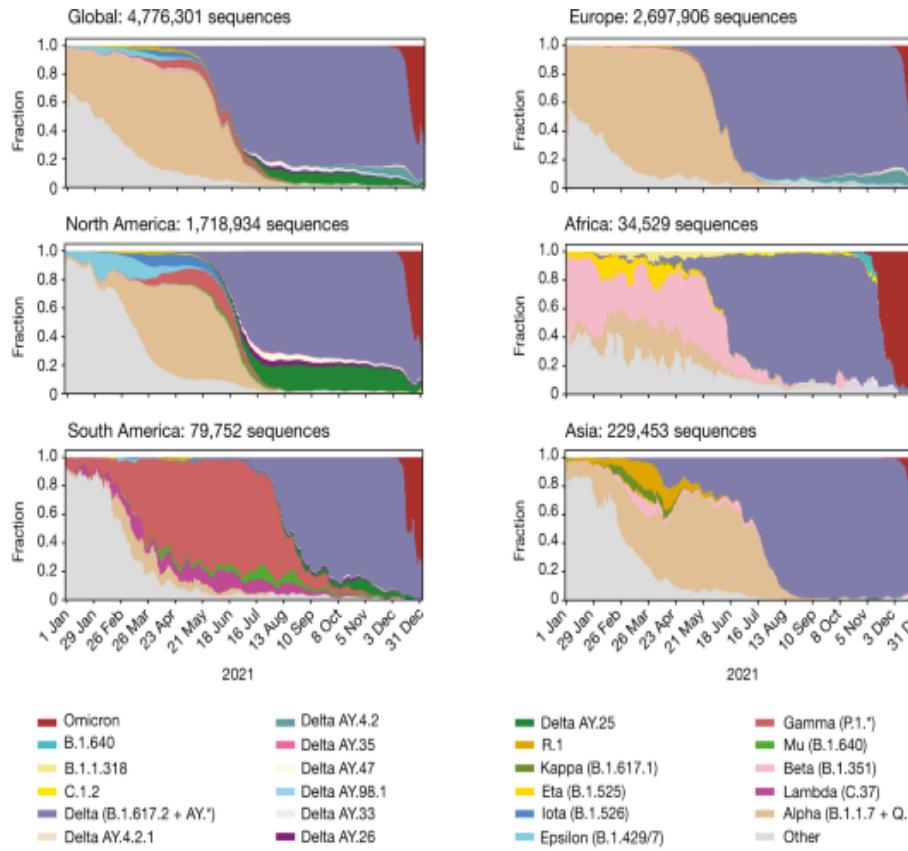
The early-detection and analysis group

SARS-CoV-2 genome sequencing data have been shared in public databases. As of December 2021, GISAID—the most widely used database for SARS-CoV-2—has more than 6.5 million sequences deposited with more than 150,000 sequences added weekly. This depth and rate of growth of genetic information for an emerging virus is unprecedented, providing a unique resource to track virus evolution. From late 2020, the emergence of variants of concern (VOCs) with an increased risk to global public health prompted scientists to establish variant detection and tracking pipelines (such as Outbreak.info¹³, CDC SARS-CoV-2 Variant Classifications and Definitions¹⁴ and the BV-BRC SARS-CoV-2 Real-time Tracking and Early Warning System for Variants and Lineages of Concern (<https://www.bv-brc.org>)). The early-detection and analysis group was assembled to establish a systematic approach to identify and predict SARS-CoV-2 variants that might increase virus replication, transmission and/or escape immunity. The team's main goal is to select and prioritize variants for development of key experimental reagents (for example, spike proteins for binding assays and pseudoviruses (PSVs) for neutralization assays) and viruses for challenge studies, as well as to inform the in vitro and in vivo groups about predicted variant properties to guide their experiments. The initial and primary focus has been on variants with mutations in the spike protein that might lead to antibody escape, with subsequent analyses considering T cell escape, infectivity and transmission. Other important characteristics—such as replication fitness and virulence—and genomic regions outside of the spike gene are also evaluated. The process is collaborative and iterative, with seven teams using independent models and methodologies to prioritize mutations and lineages as well as rank importance for downstream testing. Although the focus is on human infections, the early-detection group also monitors variants circulating in animal populations, such as mink and deer, as they represent a potential reservoir source.

Methodology

Genomic surveillance consists of weekly downloads of SARS-CoV-2 genomes from GISAID/GENBANK, quality filtering, alignment, and the identification of variant or co-variant substitutions. The main focus has been on potential antibody escape to identify mutations in key epitopes in the receptor-binding domain (RBD) and the N-terminal domain (NTD) supersite, but regions proximal to the furin-cleavage site or experiencing convergent/parallel evolution are also considered. The dynamics of these spike substitutions, as a function of time and geographical spread, are evaluated considering sequence prevalence and viral population growth rate, including comparative analyses to other variants co-circulating in a given geographical location (Fig. 2a). One example of recurrent substitutions with phenotypic relevance are those near to the furin-cleavage site, which result in enhanced spike cleavage and infectivity^{15,16}. These mutations have been identified in different variants and in newly expanding lineages. Some teams take into account vaccine coverage when prioritizing an emerging lineage for analysis.

Fig. 2: Prioritization of variants by the early-detection and analysis group.



a, The trajectory of SARS-CoV-2 variant sequence prevalence over a one-year period, 1 January 2021 to 31 December 31 2021, tracking frequencies of weekly counts based on PANGO lineage designations. The data in the graphs are based on the 4.8 million SARS-CoV-2 sequences sampled in 2021 and made available through the GISAID Initiative. Updated graphs can be found online (<https://cov.lanl.gov>; the tracking tool is called Embers). Global summary and status of five continents. Europe and North America remain the most highly sampled regions of the world, biasing the global sampling. **b**, Tangle plots for comparative prioritization of circulating variants across subgroups. The list of variants to prioritize was built collectively by the whole group and prioritized by individual teams to arrive at a consensus list. Each column graph refers to the prioritization order made by each subteam for circulating variants in December 2021 (top, highest priority; bottom, lowest priority): A, Cambridge University; B, LANL; C, ISMMS; D, JCVI/BV-BRC; E, UCR SOM; F, Broad Institute; G, WRAIR. The final consensus ranking of the 43 variants was produced by ordering the lineages by their mean rank across the different teams, who also have the option to defer from ranking a lineage or to assign multiple lineages a tied ranking and, after discussion with the group, determine priority categories. The dashed arrow indicates the order of priority. The colours refer to each PANGO lineage tracked, but blocks of the same colour can also refer to different variants within a PANGO lineage. For example, in addition to the coloured Delta AY.* sublineages indicated, Delta has 26 subvariants (purple) with different combinations of mutations that are being prioritized for analysis.

The rankings are split into two broadly distinct methodologies, each with slight variations: one is based on convergent evolution as the main signal for selection and functional impact of mutations (that is, the Cambridge and Walter Reed Army Institute of Research (WRAIR) teams); whereas the other is anchored on prevalence and growth patterns of mutations and defined lineages (that is, the Los Alamos National Lab (LANL), Icahn School of Medicine at Mount Sinai (ISMMS), J. Craig Venter Institute/Bacterial Viral Bioinformatic Resource Center (JCVI/BV-BRC), UC Riverside and Broad Institute teams) (Fig. 2b).

The functional impact of mutations

Cambridge prioritizes substitutions that are likely to cause immune escape by looking at both experimentally determined escape from polyclonal sera and the effect of mutations on spike protein structure. Substitutions are given higher priority if they appear to be emerging and if they are in a different Barnes class¹⁷ from previously observed substitutions, and lower priority if they have already been tested experimentally. The WRAIR team tracks the prevalence of substitutions at a set of sites selected based on the strength of the interaction with known SARS-CoV-2 antibodies (using complex structures in the Protein Data Bank; <https://www.rcsb.org>) as well as structural information or knowledge from deep mutational scanning or mutagenesis studies. Weight scores for ranking are also given for various characteristics, such as the fold increase in detection over time and geographical spread or population growth in the context of high vaccination coverage.

Prevalence and growth patterns

The ISMMS team has a similar approach, whereby variants are ranked on the basis of an aggregate score for sequence prevalence increase and genetic changes of concern in sites of importance associated with functional changes (such as ACE2 binding, antibody escape) but also assigns weight to mutations in the active sites of viral enzymes. Moreover, data from surveillance cohorts in the New York City metropolitan area are used to assess lineages associated with local outbreaks and breakthrough infections after vaccination. LANL identifies emergent mutational patterns within the spike, RBD and NTD supersite to determine global and regional sampling frequencies. Variant dynamics and global spread are tracked at multiple geographical levels using a suite of tools⁵ (<https://cov.lanl.gov/>). The JCVI/BV-BRC team uses an algorithm combining sequence prevalence dynamics with functional impact predictions to rank emerging variants. Each mutation is given a sequence-prevalence score, reflecting geographically localized prevalence changes, and a functional impact score, on the basis of the location of the mutation within important spike protein regions and whether studies have demonstrated significant changes in either antibody- or ACE2-receptor binding^{18,19,20,21}. UC Riverside uses relative growth in the prevalence of specific substitutions and deletions/insertions to identify the fastest growing variants and mutation combinations

(<https://coronavirus3d.org>). For the final variant and subvariant ranking, additional criteria are included, such as their potential impact on protein structure (by modelling) and the re-emergence of individual mutations in previously undescribed combinations in new variants. Finally, the team from the Broad Institute, similar to the UC Riverside team, examines the accelerated growth of a variant relative to its peers, across multiple geographical regions, but fits a binomial logistic regression to each lineage's proportion over time. Moreover, they fit hierarchical multinomial logistic regression models across geographical regions²².

Challenges for the early-detection and analysis group

The early-detection and analysis group has faced six main challenges in identifying emerging variants for functional testing: (1) the newest data are the most subject to bias and the least representative because of small numbers. The longer that one waits, the more accurate the data, but the greater the delay in identifying newly emergent variants for evaluation. (2) Disentanglement of epidemiological from evolutionary effects. A variant might show increased sequence prevalence within a geographical region due to founder effects, or increased incidence could be conferred by epidemiological factors rather than an evolutionary fitness advantage. An example of a founder effect is Delta AY.25, which is very common in North America but not increasing in frequency over time (Fig. 2a), versus AY.4.2, which was first sampled well after Delta was increasing in the UK and was constantly increasing in frequency in 27 countries where it was found and, furthermore, it never significantly decreased relative to other Delta variants once it emerged, suggesting positive selection. (3) Selective pressures on the virus are in flux, and mutations may be transient due to a balance with requirements for retention of fitness. Pressures are exerted by the host at the level of transmission, epidemiological interventions and immune evasion. (4) Under-representation of variant spread and evolution in countries with limited sampling and sequencing capacity. Although some parts of the world have an abundance of sequencing data(such as the UK and USA), others are under-represented (such as the African continent and China). There is an urgent need to increase sampling and sequencing capacity in resource-poor

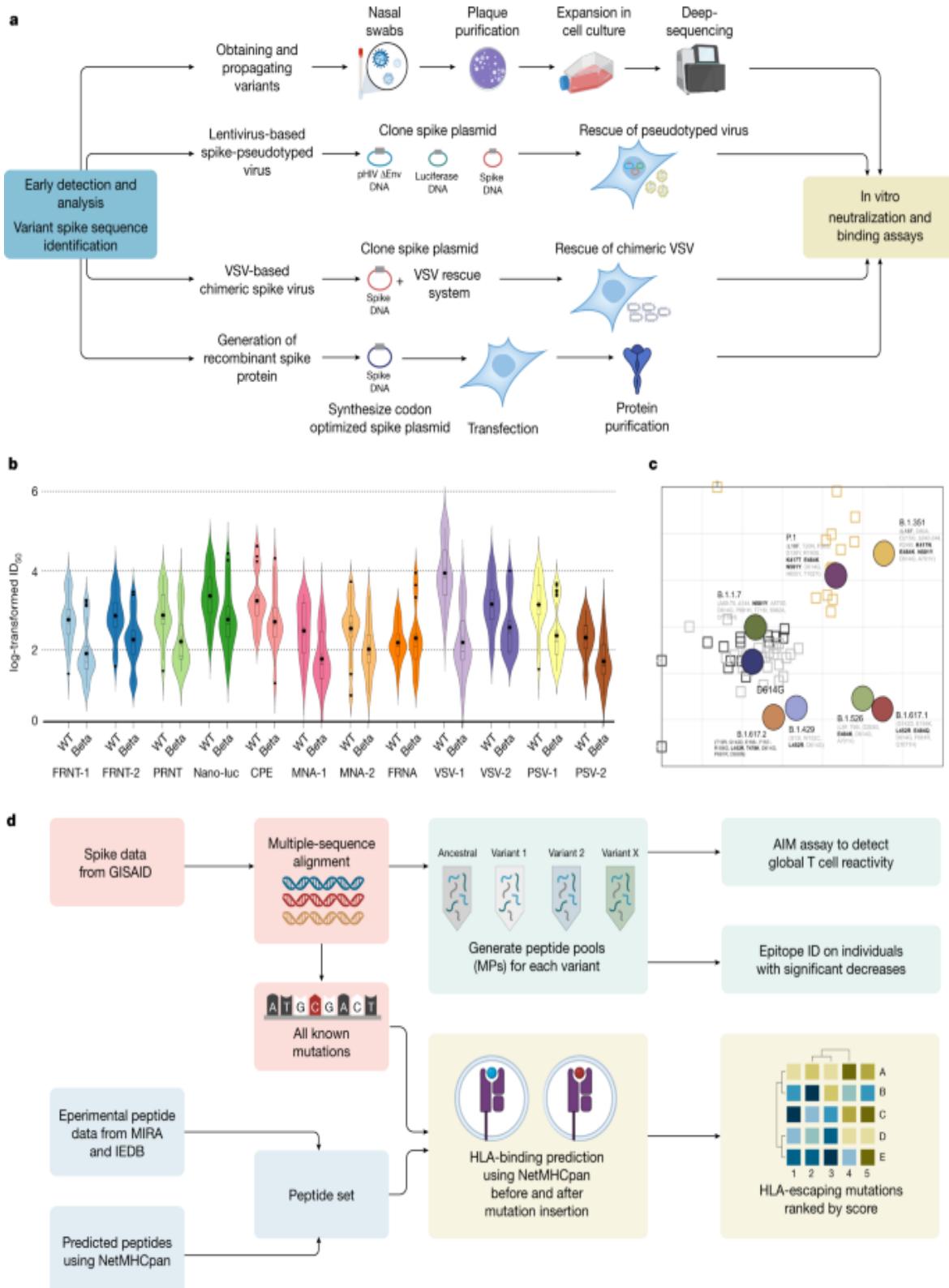
countries. (5) Variability in data quality. The submission of consensus assemblies without underlying raw read-level data means that quality cannot be independently evaluated. Erroneous genome sequences due to technical artifacts, low coverage or bioinformatic strategies that default to ancestral bases in regions without sequence coverage can affect the accuracy of variant amino acid calls²³. (6) The database curation quality-control steps can filter on the basis of criteria that do not apply uniformly across lineages. The B.1.621/Mu lineage had an unexpected stop codon in *ORF3a* that caused B.1.621 sequences to be flagged during automated uploads to the GISAID database, which initially led Mu to be undercounted. This can lead to a false understanding of the dynamics of a given variant lineage globally. Despite these challenges, our prioritization methods continue to evolve as more information becomes available. These efforts have allowed for the rapid generation of reagents for multiple variants before they have spread extensively in the USA and have been critical for guiding the in vitro and in vivo groups. A list of regularly updated prioritized variants is available online (https://docs.google.com/spreadsheets/d/167uJP9LfJN07410sWaMSKU1Se-4XX687j8IgVX4MV_w/edit?usp=sharing).

The in vitro group

The in vitro group performs antibody binding, neutralization, Fc effector and T cell stimulation assays to understand how SARS-CoV-2 variants affect vaccine- and infection-induced immunity. The in vitro group serves as a critical intermediary between the early detection and analysis and in vivo groups by providing valuable data to confirm variant lineage prioritization, and ranking viruses for prioritized in vivo challenge studies. The in vitro group was initially tasked with developing key reagents (for example, spike and RBD antigens, and plasmids for generating PSVs) and procuring biospecimens (such as authentic viruses and sera/plasma from infected and vaccinated individuals). At the beginning of 2021, reagents for generating data—including variant virus isolates, recombinant infectious clones, recombinant variant spike proteins for antibody binding assays, variant-specific expression plasmids for PSV particle entry inhibition assays and variant-specific sera—were not widely available (Fig. 3a). A key lesson from this process is that the streamlining of administrative procedures for

reagent sharing facilitates data generation that directly informs urgent policy- and decision-making. A substantial and ongoing challenge requiring numerous administrative steps is to obtain authentic virus isolates from domestic and international sources. To expedite this process, we developed a pipeline between SAVE investigators to isolate, propagate and sequence emerging viruses. This effort led to cataloguing and isolating hundreds of SARS-CoV-2 variants representing over 40 lineages. For more difficult to obtain SARS-CoV-2, additional efforts have been made to generate infectious clones^{[24,25,26](#)}. Furthermore, the early-detection group prioritized viral variants and curated sequences to accelerate the production of recombinant variant spike proteins and expression plasmids.

Fig. 3: In vitro group.



a, Live-virus nasal swabs in viral transport medium or seed stocks are obtained followed by plaque purification and deep-sequencing. Pseudotyped

virus plasmids encoding the variant spike sequence are synthesized to generate pseudotyped lentivirus stocks. The vesicular stomatitis virus (VSV) chimeric virus glycoprotein gene (*G*) is replaced with the spike protein of SARS-CoV-2 (VSV-eGFP-SARS-CoV-2) and a GFP reporter gene³³. **b**, The in vitro group conducted performance testing between 12 neutralization assays involving live authentic virus consisting of focus-reduction neutralization test (FRNT-1), recombinant SARS-CoV-2 reporter virus FRNT (FRNT-2), plaque-reduction neutralization test (PRNT), recombinant SARS-CoV-2 expressing nano-luc (Nano-luc), cytopathic effect assay (CPE), microneutralization assay (MNA), focus-reduction neutralization assay (FRNA), lentivirus and VSV pseudotyped neutralization assays, and VSV chimeric assays. An example comparison between the wild-type (WT) and Beta virus is presented. **c**, Antigenic cartography; 50% infectious dose (ID_{50}) neutralization titres in a lentivirus-based PSV assay were determined against a panel of SARS-CoV-2 variants and serum from individuals who were vaccinated with the Moderna vaccine, or individuals with SARS-CoV-2 infection. The distance between serum to an antigen corresponds to the titre of that serum for the antigen³⁷. The grid lines represent twofold dilution of antiserum. The *y* and *x* axes represent antigenic distance. Circles, antigens; squares, sera. **d**, T cell responses to SARS-CoV-2 variants. Sequencing data are curated for coding mutations (pink boxes). Curated mutations are tested on convalescent T cell responses using functional assays (activation-induced marker (AIM) assays; green boxes). Immune Epitope Database (IEDB) and the immunocode multiplex identification of T cell receptor antigen specificity dataset (MIRA) are analysed to generate curated peptide sets of immunodominant epitopes (blue boxes). Data are integrated to produce a ranked score list of variant epitope changes weighted by their likelihood to disrupt epitope binding and the relative size of the affected population (grey boxes). MPs, megapools. Partially created using BioRender.

At the start of the pandemic, the correlates of immune protection were unknown for COVID-19. Multiple teams within the in vitro group conducted assessments of vaccine-induced serum neutralization using parallel but independent methods across laboratories. Studies with clinical samples show neutralizing antibody titres are a strong predictor of protection against severe disease²⁷. As such, a major undertaking of the in vitro group has been to use

neutralization assays to assess the effect of spike mutations on the inhibitory activity of clinically approved monoclonal antibodies and serum/plasma from vaccinated or infected individuals. One of the strengths of the in vitro group is the use of orthogonal SARS-CoV-2 neutralization assays based on authentic live viruses, PSVs and chimeric viruses. An initial task of this group was to compare neutralization assay platforms across 12 independent laboratories using a defined serum panel from individuals vaccinated with the Pfizer and Moderna vaccines. Using either the ancestral wild-type virus (Wuhan-1) or more recent variants (for example, the Beta variant shown in Fig. 3b), team members performed neutralization assays that varied on the basis of live virus assay readouts (foci, plaques, cytopathic effect, luciferase and fluorescence), target cells, and expression of ACE2 and/or TMPRSS2 on target cells^{28,29,30,31,32,33,34,35}. This type of performance testing has highlighted differences between assay platforms, cell targets and readouts that can impact neutralization potency. Nonetheless, in most cases, there was considerable congruence across platforms. Another area of emphasis is using variant-infected serum/plasma samples to visualize the antigenic evolution of spike through a process called antigenic cartography^{36,37} (Fig. 3c). This two-dimensional map provides a landscape of how spike mutations drive loss in neutralizing activity.

For many viruses, the affinity and magnitude of antibody binding to viral glycoproteins associates with virus-neutralizing activity, and a strong correlation has been shown for SARS-CoV-2^{38,39,40,41}. Investigating the correlation between the neutralizing and binding activity of vaccine-induced antibodies showed that spike mutations alter this slope, and virus neutralization is often more affected than antibody binding^{42,43}. This has been confirmed through different platforms measuring changes in binding to either native spike proteins or the RBD, including ELISA⁴⁴ and multiplexed spike antigen detection platforms⁷. One potential explanation for this is that many more binding than neutralizing epitopes exist on the spike protein. Some antibodies that have neutralizing activity against the wild-type virus may lose activity to variants, yet overall binding is still maintained—a phenomenon observed for other viruses (such as influenza virus⁴⁵).

Binding antibodies can still have a considerable protective effect, irrespective of neutralizing activity due to Fc effector functions, as seen with

influenza virus or Ebola virus^{46,47,48}. The humoral immune response restricts microbes through the coordinated effort of the Fab (antigen-binding) and Fc (constant) domains⁴⁹. After infection or vaccination, polyclonal antibodies are induced that target pathogens at multiple sites through their Fab domains. Fab domains that directly or indirectly hinder virus entry are neutralizing; however, the remaining ‘non-neutralizing’ antibodies can bind to and opsonize the pathogen to form immune complexes, or bind to spike proteins on the surface of infected cells. Once complexed, the Fc domains act as molecular beacons that draw in immune cells through Fc-gamma receptors (FcγRs), providing instructions on how the immune system should destroy the antibody-opsonized material. Fc-effector functions of antibodies are linked to natural resolution of COVID-19^{50,51,52,53}, correlate with vaccine-mediated protection from infection in animal models^{54,55,56} and are associated with protection after the transfer of passive convalescent serum or monoclonal antibodies^{57,58,59,60}. Although emerging variants of SARS-CoV-2 can escape neutralizing antibodies, their substitutions alter a limited fraction of the overall humoral immune response to the SARS-CoV-2 spike^{56,61}. Thus, Fc-effector functions have more resilience in the face of variation across spike, for both mRNA and the adenoviral 26 (Ad26) vaccines, offering mechanisms through which antibodies may continue to confer protection despite escaping neutralization.

Growing evidence from animal models and human studies indicates that CD4⁺ and CD8⁺ T cells have protective roles in preventing severe disease and death from SARS-CoV-2 infection^{6,62,63,64}. T cells are an attractive target for intervention as they are less susceptible to viral escape than antibodies^{6,65}. This is largely for two reasons: (1) in convalescent individuals, T cells can target peptides derived from the entire proteome, not just surface-exposed epitopes; and (2) HLA-restriction and diversity creates interpersonal variation in the repertoire of targets, limiting the immunological pressure on any one epitope. Given the presumed role of T cells in limiting severe disease and their potential for sustaining protection against variant mutation, the SAVE in vitro group included assessment of T cell responses. The goal was to determine empirical drift from vaccination and infection-induced immunity, and to develop tools to predict the impact of variant-associated mutations on immunodominant T cell responses.

The T cell investigations follow two parallel approaches to assess the impact of variant mutations on T cell reactivity and a broad range of different variants (Fig. 3d). The first involves measuring the overall reactivity against the entire spike protein (in the case of vaccination) or the entire proteome (in the case of infection) and expressing the results as the fold difference relative to the ancestral sequences. A parallel approach characterizes the mutational impact on specific single epitopes, and monitors whether individuals with decreased T cell reactivity have responses that selectively recognize certain epitopes in the context of particular HLA types. Regarding the first approach, at the general population level, the results to date have detected a limited impact of mutations within spike after natural infection or mRNA vaccination⁶ against the most concerning variants at the time the study was performed (B.1.1.7, B.1.351, P.1 and B.1.427/429). These findings were corroborated⁶⁶ and expanded to adenoviral-vector-based vaccination⁶⁷. However, in a minority of individuals, two- to threefold decreases in the CD8⁺ T cell responses against the B.1.351/Beta and B.1.427/429/Epsilon variants were noted⁶. These findings suggest that a more in-depth characterization at the single-epitope level is required to understand the mechanisms behind the reduced CD8⁺ T cell response in specific individuals. Moreover, it is critical to monitor and predict the effect of emerging circulating variants on T cell reactivity, particularly regarding the most concerning (to date) B.1.617.2/Delta variant (including the AY.* sublineages) and B.1.1.529/Omicron variant. The experimental data will be used to confirm and improve the bioinformatics analysis and infer the impact of current and upcoming variants on SARS-CoV-2 specific T cell responses.

Advanced computational tools for assessing SARS-CoV-2 genome mutations on HLA binding have enabled prediction of the effect of mutations within a VOC on T cell reactivity. Owing to the broad diversity of HLA genotypes, T cell escape at the population level is not likely, as demonstrated for multiple VOCs⁶. However, previous work on HIV and influenza virus has identified associations between specific HLA class I alleles, disease severity^{68,69} and vaccine efficacy⁷⁰. We anticipate that, as SARS-CoV-2 continues to spread globally, T cell immunity will eventually drive viral evolution. In these situations, specific HLA alleles may become associated with a reduced ability to mount responses against dominant T cell

epitopes, which may affect clinical outcomes. The T cell subgroup has developed a computational pipeline to assess the effects of specific mutations on HLA binding by also ranking all individual mutations on a T cell escape score, based on experimentally verified and predicted T cell responses (Fig. 3d). This ranking will provide early identification of specific mutations associated with T cell escape, particularly CD8⁺ T cells, and testable hypotheses for T cell experiments. In our preliminary analyses of VOCs, the B.1.617.2 variant was identified as the first in which mutations were associated with reduced HLA binding at the population level. These data suggest that T cell cross-reactivity to B.1.617.2 may be reduced in some individuals. Owing to the extensive number of SARS-CoV-2 viral genomes, and large-scale clinical cohorts that are being studied, the T cell SAVE group plans to assemble a database linking HLA genotypes with clinical outcome and viral genomes, which may provide a unique opportunity to study HLA associations with clinical disease and viral evolution at a resolution that has not previously been attempted.

Challenges for the in vitro group

Work by the in vitro group has focused mostly on characterizing neutralizing antibody responses to the spike protein with some analysis of the impact of variants on T cell responses as well. With the recent increase in Omicron infections in vaccinated and unvaccinated individuals, a challenge moving forward will be to disentangle vaccine- and infection-induced immunity, breakthrough infections, waning immunity and other covariates associated with increased risk of symptomatic infection (immunocompromised, age, obesity, diabetes). Many other key aspects of SARS-CoV-2 and its variants remain uninvestigated. Although neutralizing antibodies correlate with protection from SARS-CoV-2, neutralization is not the only function of antibodies. In fact, non-neutralizing antibodies can afford substantial protection against influenza virus^{46,71,72,73} and similar mechanisms remain to be examined for SARS-CoV-2. Furthermore, differences between wild-type and variant viruses in ACE2 binding, fusion, impact of mutations on spike processing by proteases, and potentially fusion at the cell membrane and cell-to-cell fusion, remain poorly understood¹⁵. Furthermore, the spike protein is just one of many SARS-CoV-2 proteins. The effect of mutations in non-spike proteins on immunity and viral fitness, including transmission,

virus–host interaction and polymerase fidelity has not yet been assessed. The use of reverse genetics systems and PSVs can be leveraged to understand the contribution of individual mutations to viral fitness and evasion of antibody responses^{24,25,26,74}. We acknowledge that differences between the ancestral and variant viruses may also impact neutralization assays in different cell lines. Although we have seen some of these cell-line-specific effects in in vitro neutralization, we do not yet understand their underlying mechanisms. Furthermore, we need to increase the use of reference standards in binding and neutralization assays, such as the WHO International Standard and International Reference Panel for anti-SARS-CoV-2 immunoglobulin⁷⁵, to calibrate assays and provide a means to compare serological findings. Rare and volume-limited variant-specific sera/plasma are difficult to obtain and share across borders, and/or between academic institutions, and the process is often slowed by administrative hurdles. On occasion, SARS-CoV-2 variant sample sharing has not been possible within the needed timeframe, impeding the research response to this public health emergency. Although access to virus isolates outside the USA and variant-specific human sera remains limited, the in vitro group has created an extensive network of collaborations to overcome these hurdles. Finally, much remains to be examined for both antibody and T cell responses about emerging variants such as B.1.617.2, B.1.617.2 subvariants, B.1.1.529 and other new variants of interest and concern.

The *in vivo* group

SARS-CoV-2 animal models have been critical for the development and testing of vaccines and antiviral therapeutics^{76,77,78,79,80,81,82,83,84}. Initial countermeasures targeted the spike protein from the SARS-CoV-2 strain circulating during the early phase of the pandemic in 2020 and focused on efficacy testing against homologous strains. However, the emergence of variants and their possible effects on transmission, pathogenesis, and infection- or vaccine-mediated immunity required rapid adaptation of animal models to confirm vaccine efficacy against VOCs. The *in vivo* group was assembled to develop animal models, standardize reagents and assays, and examine the effect of SARS-CoV-2 variants on protection elicited by vaccine- or infection-induced immunity and transmission. The variants identified for investigation by the early-detection and analysis group that are

validated in the in vitro group are then forwarded to the in vivo group. The in vivo group studies protection against SARS-CoV-2 variants using an array of animal models, including mice, hamsters and non-human primates (NHPs). This has led to a collaborative process in which transmission, pathogenesis and protection data are shared to develop a consensus on the effect of emerging variants on protective immunity.

Development of animal models

One of the first tasks of the in vivo group was to standardize viral challenge stocks, routes and doses of infection, and vaccination strategies across each of the animal models. To minimize variability and adventitious mutations associated with virus propagation in different cell types, Vero-TMPRSS2 cells were used to generate challenge stocks and distributed among team members⁸⁵. As new variants emerge, viruses are tested in small-animal models to determine infectivity and pathogenicity. For vaccination studies, the group focused on evaluating both approved (including Emergency Use Authorization (EUA)) vaccines and those undergoing advanced clinical testing in humans, including mRNA vaccines (Pfizer BNT162b and Moderna mRNA-1273), protein-based vaccines (Novavax NVX-CoV2373) and virus-vectored vaccines (J&J Ad26.COV2.S). For these studies, vaccine doses were optimized to model magnitude and durability of vaccine-induced immunity across the animal models. For each vaccination experiment, many parameters are studied, including neutralizing antibody potency and kinetics, pathogenesis of ancestral and variant viruses, as well as the levels of virus in various respiratory tract tissues.

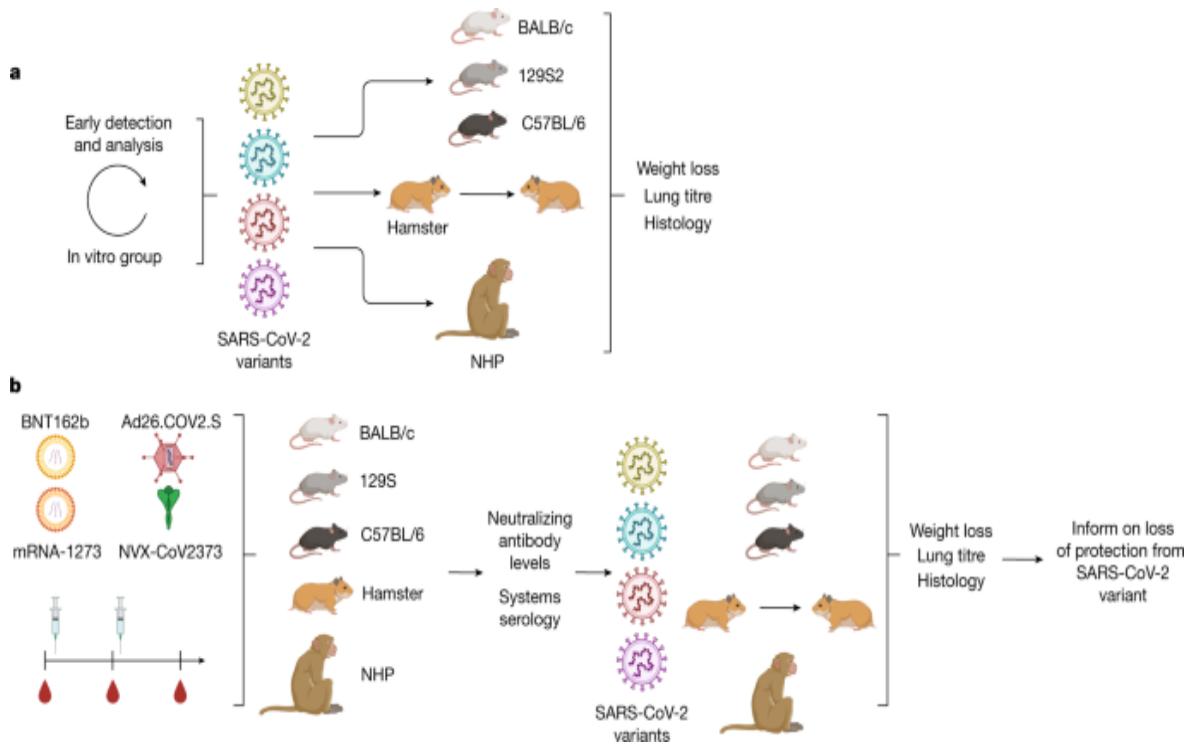
Mouse models

The ancestral SARS-CoV-2 strain does not replicate in conventional laboratory mice as the spike protein inefficiently binds to mouse ACE2^{86,87}. To overcome this obstacle, several mouse models were developed, including human ACE2 (hACE2) transgenic mice (for example, K18-hACE2⁸⁸, originally developed for studies of SARS-CoV⁸⁸) and mice that express hACE2 transiently after transduction with viral vectors (such as adenovirus)^{89,90}. The K18-hACE2 transgenic mice are highly permissive for

most SARS-CoV-2 strains and variants, and infection typically results in weight loss, nasal turbinete and lung infection, pneumonia and death^{60,91,92}. Lungs from SARS-CoV-2-infected mice show denuding bronchiolitis, mixed inflammatory infiltrate, alveolar oedema and alveolitis^{92,93,94}. Some mice, especially young K18-hACE2 mice, develop infection in the brain and encephalitis, which may confound interpretation of clinical disease⁹⁵. The spike mutation at position N501Y, which is found in mouse-adapted strains and several emerging variants (Alpha, Beta, Gamma, Mu and Omicron), increases the affinity of the SARS-CoV-2 spike protein for the murine ACE2 receptor and enables direct infection of inbred mice⁸⁶. Thus, in addition to the K18-hACE2 mouse model, challenge of conventional laboratory mice with mouse-adapted virus or SARS-CoV-2 variants containing an N501Y substitution within the spike protein cause pneumonia in BALB/c, 129S2 and C57BL/6 mice in an age-dependent manner^{89,96,97,98}. The initial characterization of each variant in a variety of mouse strains at different doses allowed for vaccination studies to be conducted with validated stocks and consistent phenotypes. The in vivo group has used both inbred mouse strains (such as BALB/c, C57BL/6, 129S1, 129S2) and K18-hACE2 transgenic mice for iterative infection and vaccination studies.

As part of the experimental design, the in vivo team uses a high and low vaccine dose strategy with Ad26.COV2.S, NVX-CoV2373, BNT162b and mRNA-1273 to study the effect of variant mutations on protection. Mice inoculated with a high vaccine dose are useful for evaluating antibody-mediated protection of the upper and lower respiratory tract after challenge. Mice inoculated with the lower vaccine dose serve as a model for suboptimal immunity (as might be seen in the elderly or immunocompromised) and breakthrough infection. Through this effort, the group has demonstrated that low doses of mRNA vaccines (BNT162b and mRNA-1273) or protein vaccine (NVX-CoV2373) show reduced protection compared with high vaccine doses (Fig. 4) against the B.1.351 and B.1.617.2 variants⁹⁹. These experiments are being extended to study additional variants and a range of vaccination doses.

Fig. 4: In vivo group.



a, Animal model development. After selecting variants for analysis by the early-detection and in vitro analysis group, isolates are grown, validated using next-generation sequencing, and analysed in each animal model at different doses to determine the pathogenicity, viral kinetics and transmission (in hamsters). Weight loss, lung titre and lung pathology was assessed to generate benchmarks for vaccine studies. **b**, Vaccine challenge. Each animal model is immunized with selected vaccines. Animal serum is examined after vaccination for neutralizing antibody levels and across a systems serology analysis before viral challenge with the chosen variants. Protection against infection and disease in each model is analysed to determine the protective ability of each vaccine and variant. Data on the protection of each animal model with each vaccine platform and challenged with variant viruses are shared with the SAVE consortium and SIG. Created using BioRender.

Another goal of the SAVE team is to assess the ability of previous infection to protect against secondary challenge with SARS-CoV-2. To model this, middle-aged C57BL/6 mice are inoculated with SARS-CoV-2 (B.1.1.7/Alpha, B.1.351/Beta or P.1/Gamma variants) followed by a homologous or heterologous challenge up to 120 days later with either variants or mouse-adapted SARS-CoV-2. Mice infected with any of the three

variants remain asymptomatic but develop a neutralizing antibody response that is measurable at 21 days after infection. However, neutralizing antibodies cannot be detected in most mice at 3 months after infection, yet mice are still partly protected from variant virus challenge or from a dose of mouse-adapted SARS-CoV-2 that is lethal to mice without pre-existing immunity. Part of the goal of this project is to measure T and B cell memory responses to understand why the mice remain protected.

Hamster models

Syrian golden hamsters are highly susceptible to SARS-CoV-2 infection and disease without any species-specific adaptation of the virus, and have a disease phenotype that resembles mild disease observed in human COVID-19 cases. Loss of 10–20% of the initial body weight is seen at 6–7 days after infection depending on the age and sex of the animal and the variant and dose of the virus^{[100,101,102](#)}. Virus replication is confined to the upper and lower respiratory tract, which peaks at 3 days after infection and then wanes to undetectable amounts by 10 days after infection. Imaging of the lungs of infected hamsters shows abnormalities during the course of infection that do not directly resolve even after virus clearance^{[100](#)}.

Pathological changes in the lungs of the hamsters are comparable to those in some humans and are characterized by widespread, moderate to severe bronchointerstitial pneumonia^{[100,103](#)}. Lung lesions comprise focal extensive areas of pulmonary oedema and consolidation with evidence of interstitial pneumonia. Histopathological lesions include fibrin deposits and oedema in alveolar spaces, influx of neutrophils and macrophages into alveolar spaces, the presence of syncytial cells and prominent type II pneumocyte hyperplasia. Secondary bacterial infections are often detected in the lungs. Despite a robust infection, SARS-CoV-2 is not lethal in healthy hamsters and infected animals recover.

As the use of the hamster model for SARS-CoV-2 infection was established using an early isolate of SARS-CoV-2, infection studies have been performed with several variants. Hamsters are largely agnostic to the variants, demonstrating little differences in viral replication and shedding kinetics between different variants including D614G, B.1.1.7 and

B.1.351¹⁰⁴. Although competition infections are more sensitive to revealing small effects of SARS-CoV-2 mutations on fitness for airway infection and transmission^{25,105}, recent studies indicate that B.1.1.529/Omicron is attenuated in hamsters with less infection in the lungs¹⁰⁶.

Given their general susceptibility to SARS-CoV-2, hamsters are an excellent model to study vaccine-induced immunity against variants. Although immunological reagents are less widely available for hamsters, antibody responses induced by vaccination can be measured in neutralization assays and ELISAs by using hamster-specific IgG, IgA and IgM secondary antibodies. Cohorts of animals receive two immunizations with either the Pfizer BNT162b2 and Moderna mRNA-1273 vaccine given three and four weeks apart, respectively. The vaccine dose in these studies has generally been one-third of the dose given to humans (that is, 10 µg of the Pfizer vaccine or 35 µg of the Moderna vaccine), and an additional freeze–thaw of the Pfizer and Janssen vaccines does not decrease immunogenicity. IgG antibody titres against the SARS-CoV-2 spike can be detected after the first vaccination but are more robust three weeks after the second vaccination, similar to data in human studies. As long-term vaccine immunity is a key question in SARS-CoV-2 research, vaccinated animals are held for extended periods of time before challenge with new emerging variants.

Transmission studies are well established in hamsters, as SARS-CoV-2 can transmit efficiently through the aerosol route from an infected to naive hamster. Increases in the transmission potential of the D614G and the B.1.1.7 variants have been observed compared with other isolates of SARS-CoV-2. Using a direct-contact transmission model, intramuscular or intranasal vaccinated ChadOx1/AZD1222 hamsters were protected from disease but not upper respiratory tract infection^{107,108}. This suggests the hamster transmission model is a useful tool to study vaccine efficacy in the context of natural exposure.

NHP model

Vaccinated NHPs are an important experimental model for demonstrating immunogenicity and protective efficacy against SARS-CoV-2^{30,109,110,111,112}. NHPs have several advantages for clinical translation.

First, NHPs are outbred, and their innate immune responses and B and T cell repertoires have greater similarity to humans than those of rodents. Owing to the diversity in class I and II MHC, NHPs also support the study of the breadth of T cell responses induced by vaccines. Second, NHPs enable the use of clinically relevant vaccine doses and are an excellent model to study the durability of immune responses. Third, after intranasal and intratracheal administration, viral replication occurs rapidly in the upper and lower airways, respectively, with similar kinetics to humans^{[109,110](#)}. For most SARS-CoV-2 strains, infection is cleared by 7–10 days after challenge. Inflammation and pathology in the lungs are consistent with mild infection as described in humans. The NHP model has been used to show immunogenicity and protection after vaccination with mRNA^{[77](#)}, ChAdOx-1^{[78](#)}, Ad26-spike^{[84](#)}, protein/adjuvant^{[79](#)} or inactivated whole virus^{[113,114](#)}, which have all been approved for use in humans. The NHP model has also been used to understand immune correlates and mechanisms of protection. The SAVE investigators are currently studying homologous and heterologous prime–boost vaccinations using EUA or approved vaccines for their ability to induce humoral and cellular immunity, longevity of immune responses, and the mechanisms associated with induction of long-lived immunity and protection in the upper and lower airway.

Challenges for the in vivo group

SARS-CoV-2 animal models provide an opportunity to understand mechanisms of infection, inflammation, pathogenesis and transmission across species and against different vaccine platforms. Similar to the concerns raised with the in vitro group, an initial challenge for the in vivo group was obtaining authentic viruses without cell-culture-adaptive mutations for challenge studies. To overcome this hurdle, a parallel pipeline for propagating and sequencing challenge stocks was developed to ensure not only the presence of lineage-defining mutations but also the absence of mutations associated with propagation in tissue culture, and that included standardizing virus dose and routes of inoculation and distributing the same stocks to all of the team members. SARS-CoV-2 strains are constantly evolving, challenging the decision as to which strains are most relevant for in vivo study. Each animal model has unique opportunities and limitations that are considered when evaluating protective immunity against a variant. Mice are a tractable system with an array of immunological tools, assays and genetic knockout strains that allow for experimental rigour and mechanistic analysis. However, inbred mouse strains are limited to SARS-CoV-2 variants that possess an N501Y mutation in the spike protein, and the genetic background (such as BALB/c, 129S2, C57BL/6) can impact viral replication and pulmonary pathology. Most variants appear to infect hamsters at similar levels with comparable lung inflammation and pathology, with the apparent exception of B.1.1.529 (ref. [106](#)). However, infection in hamsters causes a mild to moderate disease and there are limited immunological reagents to probe the response to infection and vaccination. The transmission models are exquisitely sensitive and further studies are needed to understand experimental parameters (such as airflow, contact time, relative humidity, temperature) that modulate the transmission efficiency and standardize experimental systems. The NHP model has been used in conjunction with the small-animal models using the same viral stocks to provide a more comprehensive analysis across species for how the vaccines are mediating protection. As NHPs are more limited in their availability, the rodent models can inform the best use of this model. As the SARS-CoV-2 pandemic continues, animal models will need to be adapted

to reflect the immune status of the population (for example, natural infection, vaccination and booster shots). These large- and small-animal models will be essential for further testing of next-generation vaccines, boosters formulated with variant spikes and immunological imprinting.

Summary

Collaborative science and open sharing of results in near real time have defined the SAVE programme. This cross-fertilization has enabled the efficient and rapid analysis of the effect of emerging variants on infection- and vaccine-induced immunity. The emergence of the B.1.1.529/Omicron variant, which contains more than 30 mutations in the spike protein, threatens clinically approved monoclonal antibodies and infection- and vaccine-induced immunity. The SAVE group rapidly responded by generating plasmids and spike protein, isolating, propagating and distributing authentic Omicron viral stocks, submitting reagents to public repositories, performing binding and neutralization assays and evaluating virus infection across different animal models^{106,115,116,117,118,119,120,121,122}. The data from these studies were rapidly shared with government agencies and submitted as manuscripts on preprint servers.

Over the past two decades, we have witnessed the emergence and re-emergence of several RNA viruses, including West Nile virus, H1N1 influenza virus, chikungunya virus, Zika virus, SARS-CoV-1, MERS-CoV and Ebola virus that have threatened global public health. Developing collaborative programmes between academic, industry and commercial partners is essential to respond to rapidly evolving viruses. This progressive approach combined with open communication and coordination by NIAID/NIH has facilitated rapid prioritization, reagent development, testing and assessment of SARS-CoV-2 variants. The mutual relationship between the SAVE group and the SIG has provided feedforward and feedback loops to aid with key decision matters involving risk assessment, SARS-CoV-2 countermeasures, diagnostics and public health policy. In addition to the SAVE group, other national and international networks have been developed for assessing the risk of SARS-CoV-2 mutations on immunity. This includes the Genotype to Phenotype–UK National Virology Consortium, the Genotype to Phenotype–Japan Consortium, the NIH

Accelerating COVID-19 therapeutic interventions and vaccines (ACTIV) tracking resistance and coronavirus evolution (Trace) consortium, the WHO R&D Blueprint and the National Cancer Institute Serological Sciences Network (SeroNet). These partnerships must continue to increasingly include scientists across the world to ensure that variants are rapidly identified and characterized to determine their impact on transmission, infection, replication and immune evasion. This SAVE programme is a template to develop reagents, models, assays and diagnostics, and test therapeutics and vaccines in preclinical models against rapidly evolving pathogens.

References

1. Krammer, F. SARS-CoV-2 vaccines in development. *Nature* **586**, 516–527 (2020).
2. Plante, J. A. et al. The variant gambit: COVID-19’s next move. *Cell Host Microbe* **29**, 508–515 (2021).
3. Walensky, R. P., Walke, H. T. & Fauci, A. S. SARS-CoV-2 variants of concern in the United States—challenges and opportunities. *JAMA* **325**, 1037–1038 (2021).
4. Fischer, W. et al. HIV-1 and SARS-CoV-2: patterns in the evolution of two pandemic pathogens. *Cell Host Microbe* **29**, 1093–1110 (2021).
5. 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 (2020). **This was the first study showing that a newly emerging mutation in the spike protein could come to rapidly dominate across the globe, indicating a fitness advantage.**
6. Tarke, A. et al. Impact of SARS-CoV-2 variants on the total CD4⁺ and CD8⁺ T cell reactivity in infected or vaccinated individuals. *Cell Rep. Med.* **2**, 100355 (2021).

7. Pegu, A. et al. Durability of mRNA-1273 vaccine-induced antibodies against SARS-CoV-2 variants. *Science* **373**, 1372–1377 (2021). **This study describes the impact of several SARS-CoV-2 variants on antibody binding and neutralization on mRNA-1273 vaccine sera between 1 and 6 months after the second dose.**
8. Carreno, J. M. et al. Evidence for retained spike-binding and neutralizing activity against emerging SARS-CoV-2 variants in serum of COVID-19 mRNA vaccine recipients. *EBioMedicine* **73**, 103626 (2021).
9. Jansen, J. E. & Retpen, J. A. Acute spontaneous streptococcal myositis. Case report. *Acta Chir. Scand.* **154**, 323–324 (1988).
10. Lusvarghi, S. et al. Key substitutions in the spike protein of SARS-CoV-2 variants can predict resistance to monoclonal antibodies, but other substitutions can modify the effects. *J. Virol.* **96**, JVI0111021 (2021).
11. Neerukonda, S. N. et al. Establishment of a well-characterized SARS-CoV-2 lentiviral pseudovirus neutralization assay using 293T cells with stable expression of ACE2 and TMPRSS2. *PLoS ONE* **16**, e0248348 (2021).
12. Pino, M. et al. A yeast expressed RBD-based SARS-CoV-2 vaccine formulated with 3M-052-alum adjuvant promotes protective efficacy in non-human primates. *Sci. Immunol.* **6**, eabh3634 (2021).
13. Mullen, J. L. et al. *Outbreak.info* <https://outbreak.info/> (2020).
14. *SARS-CoV-2 Variant Classifications and Definitions* (CDC, 2021).
15. Escalera, A. et al. Mutations in SARS-CoV-2 variants of concern link to increased spike cleavage and virus transmission. *Cell Host Microbe* **30**, 373–387 (2022).
16. Liu, Y. et al. Delta spike P681R mutation enhances SARS-CoV-2 fitness over Alpha variant. *Cell Reports*

<https://doi.org/10.1016/j.celrep.2022.110829> (2022).

17. Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).
18. Starr, T. N., Greaney, A. J., Dingens, A. S. & Bloom, J. D. Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell Rep. Med.* **2**, 100255 (2021).
19. Greaney, A. J. et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **29**, 463–476 (2021).
20. Starr, T. N. et al. Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850–854 (2021).
21. Starr, T. N. et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182**, 1295–1310 (2020).
22. Obermeyer, F. et al. Analysis of 6.4 million SARS-CoV-2 genomes identifies mutations associated with transmissibility. Preprint at *medRxiv* <https://doi.org/10.1101/2021.09.07.21263228> (2022).
23. Davis, J. J. et al. Analysis of the ARTIC version 3 and version 4 SARS-CoV-2 primers and their impact on the detection of the G142D amino acid substitution in the spike protein. *Microbiol. Spectr.* **9**, e0180321 (2021).
24. Xie, X. et al. An infectious cDNA clone of SARS-CoV-2. *Cell Host Microbe* **27**, 841–848 (2020).
25. Liu, Y. et al. The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature* **602**, 294–299 (2021). **This study tested all of the spike mutations identified in the first variant of**

concern, B.1.1.7/Alpha, and identified N501Y as being important for enhanced transmission and an adaptive mutation of concern.

26. Xie, X. et al. Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. *Nat. Med.* **27**, 620–621 (2021).
27. Khoury, D. S. et al. Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 1205–1211 (2021).
28. Vanderheiden, A. et al. Development of a rapid focus reduction neutralization test assay for measuring SARS-CoV-2 neutralizing antibodies. *Curr. Protoc. Immunol.* **131**, e116 (2020).
29. Wang, P. et al. Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. *Cell Host Microbe* **29**, 747–751 (2021).
30. Corbett, K. S. et al. Protection against SARS-CoV-2 beta variant in mRNA-1273 vaccine-boosted nonhuman primates. *Science* **374**, 1343–1353 (2021).
31. Chen, R. E. et al. Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* **27**, 717–726 (2021).
32. Hou, Y. J. et al. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **182**, 429–446 (2020).
33. Case, J. B. et al. Neutralizing antibody and soluble ACE2 inhibition of a replication-competent VSV-SARS-CoV-2 and a clinical isolate of SARS-CoV-2. *Cell Host Microbe* **28**, 475–485 (2020).
34. Saadat, S. et al. Binding and neutralization antibody titers after a single vaccine dose in health care workers previously infected with SARS-CoV-2. *JAMA* **325**, 1467–1469 (2021).

35. Muruato, A. E. et al. A high-throughput neutralizing antibody assay for COVID-19 diagnosis and vaccine evaluation. *Nat. Commun.* **11**, 4059 (2020).
36. Fonville, J. M. et al. Antibody landscapes after influenza virus infection or vaccination. *Science* **346**, 996–1000 (2014).
37. Smith, D. J. et al. Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376 (2004).
38. Wajnberg, A. et al. Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science* **370**, 1227–1230 (2020).
39. Amanat, F. et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020).
40. Suthar, M. S. et al. Rapid generation of neutralizing antibody responses in COVID-19 patients. *Cell Rep. Med.* **1**, 100040 (2020).
41. Cohen, K. W. et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep. Med.* **2**, 100354 (2021). **This study describes the durability of humoral and T-cell-mediated immunity after SARS-CoV-2 infection.**
42. Carreño, J. M. et al. Evidence for retained spike-binding and neutralizing activity against emerging SARS-CoV-2 variants in serum of COVID-19 mRNA vaccine recipients. *eBioMedicine* **73**, 103626 (2021).
43. Amanat, F. et al. SARS-CoV-2 mRNA vaccination induces functionally diverse antibodies to NTD, RBD, and S2. *Cell* **184**, 3936–3948 (2021).
44. Stadlbauer, D. et al. SARS-CoV-2 seroconversion in humans: a detailed protocol for a serological assay, antigen production, and test setup. *Curr. Protoc. Microbiol.* **57**, e100 (2020).

45. Roubidoux, E. K. et al. Mutations in the hemagglutinin stalk domain do not permit escape from a protective, stalk-based vaccine-induced immune response in the mouse model. *mBio* **12**, e03617–20 (2021).
46. Asthagiri Arunkumar, G. et al. Broadly cross-reactive, nonneutralizing antibodies against influenza b virus hemagglutinin demonstrate effector function-dependent protection against lethal viral challenge in mice. *J. Virol.* **93**, e01696–18 (2019).
47. Saphire, E. O. et al. Systematic analysis of monoclonal antibodies against ebola virus GP defines features that contribute to protection. *Cell* **174**, 938–952 (2018).
48. DiLillo, D. J., Tan, G. S., Palese, P. & Ravetch, J. V. Broadly neutralizing hemagglutinin stalk-specific antibodies require Fc γ R interactions for protection against influenza virus *in vivo*. *Nat. Med.* **20**, 143–151 (2014).
49. Lu, L. L., Suscovich, T. J., Fortune, S. M. & Alter, G. Beyond binding: antibody effector functions in infectious diseases. *Nat. Rev. Immunol.* **18**, 46–61 (2018).
50. Zohar, T. et al. Compromised humoral functional evolution tracks with SARS-CoV-2 mortality. *Cell* **183**, 1508–1519 (2020).
51. Atyeo, C. et al. Distinct early serological signatures track with SARS-CoV-2 survival. *Immunity* **53**, 524–532 (2020).
52. Anand, S. P. et al. Longitudinal analysis of humoral immunity against SARS-CoV-2 spike in convalescent individuals up to 8 months post-symptom onset. *Cell Rep. Med.* **2**, 100290 (2021).
53. Tso, F. Y. et al. Presence of antibody-dependent cellular cytotoxicity (ADCC) against SARS-CoV-2 in COVID-19 plasma. *PLoS ONE* **16**, e0247640 (2021).
54. McMahan, K. et al. Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature* **590**, 630–634 (2021).

55. Mercado, N. B. et al. Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* **586**, 583–588 (2020).
56. Gorman, M. J. et al. Fab and Fc contribute to maximal protection against SARS-CoV-2 following NVX-CoV2373 subunit vaccine with Matrix-M vaccination. *Cell Rep. Med.* **2**, 100405 (2021).
57. Begin, P. et al. Convalescent plasma for hospitalized patients with COVID-19: an open-label, randomized controlled trial. *Nat. Med.* **27**, 2012–2024 (2021).
58. Winkler, E. S. et al. Human neutralizing antibodies against SARS-CoV-2 require intact Fc effector functions for optimal therapeutic protection. *Cell* **184**, 1804–1820 (2021).
59. Schafer, A. et al. Antibody potency, effector function, and combinations in protection and therapy for SARS-CoV-2 infection in vivo. *J. Exp. Med.* **218**, e20201993 (2021).
60. Yamin, R. et al. Fc-engineered antibody therapeutics with improved anti-SARS-CoV-2 efficacy. *Nature* **599**, 465–470 (2021).
61. Kaplonek, P. et al. mRNA-1273 and BNT162b2 COVID-19 vaccines elicit antibodies with differences in Fc-mediated effector functions. *Sci Transl Med* <https://doi.org/10.1126/scitranslmed.abm2311> (2022).
62. Mateus, J. et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* **370**, 89–94 (2020).
63. Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501 (2020). **This study described the SARS-CoV-2-specific CD4⁺ and CD8⁺ T cell responses in patients with COVID-19 and in unexposed individuals.**
64. Dan, J. M. et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021). **This**

study assessed the durability of immune responses against SARS-CoV-2 variants in individuals who received the mRNA vaccine.

65. Goel, R. R. et al. mRNA vaccines induce durable immune memory to SARS-CoV-2 and variants of concern. *Science* **374**, eabm0829 (2021).
66. Geers, D. et al. SARS-CoV-2 variants of concern partially escape humoral but not T-cell responses in COVID-19 convalescent donors and vaccinees. *Sci. Immunol.* **6**, eabj1750 (2021).
67. Alter, G. et al. Immunogenicity of Ad26.COV2.S vaccine against SARS-CoV-2 variants in humans. *Nature* **596**, 268–272 (2021).
68. Moore, C. B. et al. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* **296**, 1439–1443 (2002).
69. Hertz, T. et al. HLA targeting efficiency correlates with human T-cell response magnitude and with mortality from influenza A infection. *Proc. Natl Acad. Sci. USA* **110**, 13492–13497 (2013).
70. Gartland, A. J. et al. Analysis of HLA A*02 association with vaccine efficacy in the RV144 HIV-1 vaccine trial. *J. Virol.* **88**, 8242–8255 (2014).
71. Henry Dunand, C. J. et al. Both neutralizing and non-neutralizing human H7N9 influenza vaccine-induced monoclonal antibodies confer protection. *Cell Host Microbe* **19**, 800–813 (2016).
72. Tan, G. S. et al. Broadly-reactive neutralizing and non-neutralizing antibodies directed against the H7 influenza virus hemagglutinin reveal divergent mechanisms of protection. *PLoS Pathog.* **12**, e1005578 (2016).
73. Jegaskanda, S., Weinfurter, J. T., Friedrich, T. C. & Kent, S. J. Antibody-dependent cellular cytotoxicity is associated with control of pandemic H1N1 influenza virus infection of macaques. *J. Virol.* **87**, 5512–5522 (2013).

74. Plante, J. A. et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592**, 116–121 (2020).
75. Knezevic, I. et al. WHO international standard for evaluation of the antibody response to COVID-19 vaccines: call for urgent action by the scientific community. *Lancet Microbe* **3**, e235–e240 (2021).
76. Corbett, K. S. et al. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature* **586**, 567–571 (2020).
77. Vogel, A. B. et al. BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature* **592**, 283–289 (2021).
78. van Doremalen, N. et al. ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature* **586**, 578–582 (2020).
79. Guebre-Xabier, M. et al. NVX-CoV2373 vaccine protects cynomolgus macaque upper and lower airways against SARS-CoV-2 challenge. *Vaccine* **38**, 7892–7896 (2020).
80. Tian, J. H. et al. SARS-CoV-2 spike glycoprotein vaccine candidate NVX-CoV2373 immunogenicity in baboons and protection in mice. *Nat. Commun.* **12**, 372 (2021).
81. Pruijssers, A. J. et al. Remdesivir inhibits SARS-CoV-2 in human lung cells and chimeric SARS-CoV expressing the SARS-CoV-2 RNA polymerase in mice. *Cell Rep.* **32**, 107940 (2020).
82. Hansen, J. et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* **369**, 1010–1014 (2020).
83. Boras, B. et al. Preclinical characterization of an intravenous coronavirus 3CL protease inhibitor for the potential treatment of COVID19. *Nat. Commun.* **12**, 6055 (2021).
84. He, X. et al. Low-dose Ad26.COV2.S protection against SARS-CoV-2 challenge in rhesus macaques. *Cell* **184**, 3467–3473 (2021).

85. Hirose, R. et al. A cytopathic effect-based tissue culture method for HCoV-OC43 titration using TMPRSS2-expressing VeroE6 cells. *mSphere* **6**, e00159–21 (2021).
86. Munoz-Fontela, C. et al. Animal models for COVID-19. *Nature* **586**, 509–515 (2020).
87. Ozono, S. et al. SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity. *Nat. Commun.* **12**, 848 (2021).
88. McCray, P. B. Jr et al. Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. *J. Virol.* **81**, 813–821 (2007).
89. Sun, J. et al. Generation of a broadly useful model for COVID-19 pathogenesis, vaccination, and treatment. *Cell* **182**, 734–743 (2020).
90. Sun, C. P. et al. Rapid generation of mouse model for emerging infectious disease with the case of severe COVID-19. *PLoS Pathog.* **17**, e1009758 (2021).
91. Oladunni, F. S. et al. Lethality of SARS-CoV-2 infection in K18 human angiotensin-converting enzyme 2 transgenic mice. *Nat. Commun.* **11**, 6122 (2020).
92. Winkler, E. S. et al. SARS-CoV-2 infection of human ACE2-transgenic mice causes severe lung inflammation and impaired function. *Nat. Immunol.* **21**, 1327–1335 (2020).
93. Golden, J. W. et al. Human angiotensin-converting enzyme 2 transgenic mice infected with SARS-CoV-2 develop severe and fatal respiratory disease. *JCI Insight* **5**, e142032 (2020).
94. Moreau, G. B. et al. Evaluation of K18-hACE2 mice as a model of SARS-CoV-2 infection. *Am. J. Trop. Med. Hyg.* **103**, 1215–1219 (2020).

95. Zheng, J. et al. COVID-19 treatments and pathogenesis including anosmia in K18-hACE2 mice. *Nature* **589**, 603–607 (2021).
96. Vanderheiden, A. et al. CCR2 signaling restricts SARS-CoV-2 infection. *mBio* **12**, e0274921 (2021).
97. Muruato, A. et al. Mouse-adapted SARS-CoV-2 protects animals from lethal SARS-CoV challenge. *PLoS Biol.* **19**, e3001284 (2021).
98. Rathnasinghe, R. et al. The N501Y mutation in SARS-CoV-2 spike leads to morbidity in obese and aged mice and is neutralized by convalescent and post-vaccination human sera. Preprint at *medRxiv* <https://doi.org/10.1101/2021.01.19.21249592> (2021).
99. Ying, B. et al. Protective activity of mRNA vaccines against ancestral and variant SARS-CoV-2 strains. *Sci. Transl. Med.* **14**, eabm3302 (2021).
100. Imai, M. et al. Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. *Proc. Natl Acad. Sci. USA* **117**, 16587–16595 (2020).
101. Bricker, T. L. et al. A single intranasal or intramuscular immunization with chimpanzee adenovirus-vectored SARS-CoV-2 vaccine protects against pneumonia in hamsters. *Cell Rep.* **36**, 109400 (2021).
102. Sia, S. F. et al. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* **583**, 834–838 (2020).
103. Imai, M. et al. Characterization of a new SARS-CoV-2 variant that emerged in Brazil. *Proc. Natl Acad. Sci. USA* **118**, e2106535118 (2021).
104. Schmitz, A. J. et al. A vaccine-induced public antibody protects against SARS-CoV-2 and emerging variants. *Immunity* **54**, 2159–2166 (2021).

105. Plante, J. A. et al. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592**, 116–121 (2021).
106. Halfmann, P. J. et al. SARS-CoV-2 Omicron virus causes attenuated disease in mice and hamsters. *Nature* **603**, 687–692 (2022).
107. van Doremalen, N. et al. Intranasal ChAdOx1 nCoV-19/AZD1222 vaccination reduces viral shedding after SARS-CoV-2 D614G challenge in preclinical models. *Sci. Transl. Med.* **13**, eabh0755 (2021).
108. Fischer, R. J. et al. ChAdOx1 nCoV-19 (AZD1222) protects Syrian hamsters against SARS-CoV-2 B.1.351 and B.1.1.7. *Nat. Commun.* **12**, 5868 (2021).
109. Francica, J. R. et al. Protective antibodies elicited by SARS-CoV-2 spike protein vaccination are boosted in the lung after challenge in nonhuman primates. *Sci. Transl. Med.* **13**, eabi4547 (2021).
110. Corbett, K. S. et al. Immune correlates of protection by mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. *Science* **373**, eabj0299 (2021).
111. Corbett, K. S. et al. mRNA-1273 protects against SARS-CoV-2 beta infection in nonhuman primates. *Nat. Immunol.* **22**, 1306–1315 (2021).
112. Gagne, M. et al. Protection from SARS-CoV-2 Delta one year after mRNA-1273 vaccination in rhesus macaques coincides with anamnestic antibody response in the lung. *Cell* **185**, 113–130 (2022).
113. Bewley, K. R. et al. Immunological and pathological outcomes of SARS-CoV-2 challenge following formalin-inactivated vaccine in ferrets and rhesus macaques. *Sci. Adv.* **7**, eabg7996 (2021).
114. Zhou, J. et al. Immunogenicity, safety, and protective efficacy of an inactivated SARS-associated coronavirus vaccine in rhesus monkeys. *Vaccine* **23**, 3202–3209 (2005).

115. Cameroni, E. et al. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **602**, 664–670 (2022).
116. Pajon, R. et al. SARS-CoV-2 Omicron variant neutralization after mRNA-1273 booster vaccination. *N. Engl. J. Med.* **386**, 1088–1091 (2022).
117. Edara, V.-V. et al. mRNA-1273 and BNT162b2 mRNA vaccines have reduced neutralizing activity against the SARS-CoV-2 Omicron variant. *Cell Rep. Med.* **3**, 100529 (2022).
118. Liu, L. et al. Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. *Nature* **602**, 676–681 (2022).
119. Tarke, A. et al. SARS-CoV-2 vaccination induces immunological T cell memory able to cross-recognize variants from Alpha to Omicron. *Cell* **185**, 847–859 (2022).
120. Valanparambil, R. M. et al. Antibody response to SARS-CoV-2 mRNA vaccine in lung cancer patients: reactivity to vaccine antigen and variants of concern. Preprint at *medRxiv* <https://doi.org/10.1101/2022.01.03.22268599> (2022).
121. Bartsch, Y. et al. Omicron variant Spike-specific antibody binding and Fc activity are preserved in recipients of mRNA or inactivated COVID-19 vaccines. *Sci Transl Med* <https://doi.org/10.1126/scitranslmed.abn9243> (2022).
122. Carreno, J. M. et al. Activity of convalescent and vaccine serum against SARS-CoV-2 Omicron. *Nature* **602**, 682–688 (2022).

Acknowledgements

A.A. is funded by HHSN272201700060C; G.A. is funded by the Ragon Institute of MGH, MIT and Harvard, the Massachusetts Consortium on Pathogen Readiness (MassCPR), the National Institutes of Health (NIH) (3R37AI080289-11S1, R01AI146785, U19AI42790-01, U19AI135995-02,

U19AI42790-01, 1U01CA260476-01), the Gates Foundation Global Health Vaccine Accelerator Platform funding (OPP1146996 and INV-001650) and the Musk Foundation; D.C.M. and X.S. are supported in part by the National Institute of Allergy and Infectious Diseases (NIAID)/NIH Collaborative Influenza Vaccine Innovation Centers (CIVIC) under contract 75N93019C00050, Duke University; A.E., A. Gordon, J.M.C., F.K. and V.S. are funded in part by NIH/NIAID CIVIC under contract 75N93019C00051, Icahn School of Medicine at Mount Sinai; G.A., B.L., V.A.M., S.S.-C. and P.G.T. are funded in part by the NIH/NIAID Center for Influenza Vaccine Research for High-Risk Populations (CIVR-HRP) CIVIC under contract 75N93019C0052, University of Georgia; M.S.D., V.S., L.B.T., H.v.B., L.A.V., R.A.M.F., A.G.-S., M.B.F., R.M.J., J.P.L., S.M.W., B.L.H., D.D.H., Y.H., Y.L., M.S.N., P.W., M.W., D.J.S., A.N., E.B.L., S.L.J., S. Türeli, S.M.W., S.A.T., J.M.C., F.K., B.M.W., B.R. and B.Y. are supported by the NIH/NIAID Centers of Excellence for Influenza Research and Response (CEIRR) under contract 75N93021C00014, Icahn School of Medicine at Mount Sinai; T.P.F., G.F., J.F., T.J., L.K., B.L., V.A.M., S.S.-C., A.S., S.T., P.G.T., E.K., L.C.-L., T.H., S. Sacharen and R.J.W. are funded in part by the NIH/NIAID CEIRR under contract 75N9302100016, St Jude Children's Research Hospital; R.R.A., V.-V.E., M.E.D.-G. and M.S.S. are funded in part by the NIH/NIAID CEIRR under contract 75N93021C00017, Emory University; A.E., T.P.F., J.F., T.J., L.K., B.L., V.A.M., S.S.-C., S.T. S.P.J.W., Z.L., L.-M.B., P.W.R., M.C.P., S. Stumpf, A.G., A.S. and R.J.W. are funded in part by the NIH/NIAID Centers of Excellence for Influenza Research and Surveillance (CEIRS) under contract HHSN272201400006C, St Jude Children's Research Hospital; A.P. is supported in part by the NIH/NIAID CEIRS under contract HHSN272201400007C, Johns Hopkins University; A.E., M.B.F., R.M.J., A.G.-S., J.P.L., S.M.W., P.J.H., Y.K., J.M.C., F.K., M. Schotsaert and V.S. are funded in part by the NIH/NIAID CEIRS under contract HHSN272201400008C, Icahn School of Medicine at Mount Sinai; R.R.A. is funded by 3R01AI148378-01S1; J.D.B. is supported in part by R01AI141707 and is an Investigator of the Howard Hughes Medical Institute; A.C.M.B. is funded by U01AI151810; E.A.B. is supported by an Intramural NIH Research Program project ZIA AI005156-02; D.L.B. is supported by NIH/NIAID under contract 75N93019D0025 to CAMRIS International; E.G. and V.J.M. are supported by the Intramural Research

Program of the NIAID, NIH; B.L.D. and M.R. are supported by a cooperative agreement between The Henry M. Jackson Foundation for the Advancement of Military Medicine and the US Department of the Army (W81XWH-18-2-0040); M.S.D. is supported in part by R01AI157155; M.D. is supported by 1R01AI143639; N.A.D.-R., D.C.D., A.B.M., N.J.S., R.A.K., S.O., S.D.S. and R.A.S. are supported by the Intramural Research Program of the Vaccine Research Center, NIAID and NIH; C.D., B.A.M. and T.C.J. are funded in part by the German Ministry of research under project codes DZIF, MolTrax and PREPARED; A.E. is supported by NIAID grants U01AI141990 and 1U01AI150747; V.-V.E., M.E.D.-G. and M.S.S. are supported in part by the NIH/NIAID (P51OD011132 and 3U19AI057266-17S1), Emory Executive Vice President for Health Affairs Synergy Fund award, and Woodruff Health Sciences Center 2020 COVID-19 CURE Award, Agod; L.J. and M. Sedova are supported by HHSN272201700060C; A. Gordon is supported by R01AI20997; M.R.H. is supported by NIAID contract HHSN272201800013C; M.P.G.K. is supported by VEO, versatile emerging infectious disease observatory: forecasting, nowcasting and tracking in a changing world. VEO has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 874735, COVID-19 vaccination in kidney patients (RECOVAC), ZonMw 10430072010002 monitoring the evolution, spread and transmission of SARS-CoV-2 through whole-genome sequencing to enable fast genotype to phenotype prediction' funded by ZonMw (project 10150062010005, APW 202645827); to support the structured evaluation of SARS-CoV-2 evolution funded by the WHO (APW 202605269); to gain in-depth understanding of the evolution, spread and transmission of SARS-CoV-2 during the coming phase of the pandemic funded by the WHO. B.K., J.T. and H.Y. are funded by the Los Alamos National Laboratory and the Gates Foundation OPP1169339; R.A.K. is funded by Intramural NIH funding; M.J.M. is supported by UM1AI068618; V.D.M. is supported by R01AI153602; S.P. is supported by P01AI060699 and R01AI129269; P.-Y.S. is supported by NIH grants AI134907 and awards from the Sealy Smith Foundation, the Kleberg Foundation, the John S. Dunn Foundation, the Amon G. Carter Foundation, the Gillson Longenbaugh Foundation and the Summerfield Robert Foundation; V.S. is supported by the NIH/NCI: Serological Sciences Network (SeroNet) 75N91019D00024; P.G.T. is supported by U01AI144616, U01AI150747,

R01AI136514, R01AI154470; Z.S.W., R.H.S., A.M.N. and M. Shukla are supported by NIH, NIAID Contract 75N93019C00076; W.W. and C.D.W. are supported by US Food and Drug Administration institutional research funds; and S.C.W. is supported by NIH grant R24 AI120942 and by the Sealy and Smith Foundation. The findings and conclusions in this Review are those of the authors and do not necessarily represent the official position of the US Centers for Disease Control and Prevention. Use of trade names is for identification only and does not imply endorsement by the US Centers for Disease Control and Prevention or the US Department of Health and Human Services.

Author information

Author notes

1. These authors contributed equally: Marciela M. DeGrace, Elodie Ghedin, Matthew B. Frieman, Florian Krammer, Alba Grifoni

Authors and Affiliations

1. National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD, USA

Marciela M. DeGrace, Elodie Ghedin, Eli A. Boritz, Debbie L. Bratt, Liliana Brown, Nicole A. Doria-Rose, Daniel C. Douek, William C. Florence, Richard A. Koup, Adrian B. McDermott, Sijy O'Dell, Stephen D. Schmidt, Robert A. Seder, Reed S. Shabman, Nancy J. Sullivan, Leah R. Vincent & Diane J. Post

2. Division of Microbiology and Infectious Diseases, National Institutes of Health, Rockville, MD, USA

Marciela M. DeGrace, Eli A. Boritz, Debbie L. Bratt, Liliana Brown, William C. Florence, Reed S. Shabman, Leah R. Vincent & Diane J. Post

3. Systems Genomics Section, Laboratory of Parasitic Diseases, National Institutes of Health, Rockville, MD, USA

Elodie Ghedin

4. Center for Pathogen Research, Department of Microbiology and Immunology, The University of Maryland School of Medicine, Baltimore, MD, USA

Matthew B. Frieman, Robert M. Johnson, James P. Logue & Stuart M. Weston

5. Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Florian Krammer, Juan Manuel Carreño, Adolfo García-Sastre, Michael Schotsaert & Viviana Simon

6. Department of Pathology, Molecular and Cell Based Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Florian Krammer, Adolfo García-Sastre & Viviana Simon

7. Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology, La Jolla, CA, USA

Alba Grifoni & Alessandro Sette

8. University of California Riverside School of Medicine, Riverside, CA, USA

Arghavan Alisoltani, Adam Godzik, Lukasz Jaroszewski & Mayya Sedova

9. Ragon Institute of MGH, MIT, and Harvard, Boston, MA, USA

Galit Alter

10. Department of Microbiology and Immunology, Emory Vaccine Center, Division of Microbiology and Immunology, Yerkes National Primate Research Center, Emory University School of Medicine, Atlanta, GA, USA

Rama R. Amara

11. Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Ralph S. Baric & Jenny E. Munt

12. Department of Microbiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Ralph S. Baric

13. Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, MA, USA

Dan H. Barouch

14. Fred Hutch Cancer Center, Howard Hughes Medical Institute, Seattle, WA, USA

Jesse D. Bloom

15. Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO, USA

Louis-Marie Bloyet, Zhuoming Liu, Marjorie C. Pontelli, Paul W. Rothlauf, Spencer Stumpf & Sean P. J. Whelan

16. CDC COVID-19 Emergency Response, Centers for Disease Control and Prevention, Atlanta, GA, USA

Gaston Bonenfant, Han Di, Li Wang, David E. Wentworth & Bin Zhou

17. Department of Medicine, Washington University in St Louis, St Louis, MO, USA

Adrianus C. M. Boon, Traci L. Bricker, Tamarand L. Darling, Astha Joshi, Larissa B. Thackray, Laura A. VanBlargan, Bradley M. Whitener, Baoling Ying & Michael S. Diamond

18. Vaccine Research Center, Bethesda, MD, USA

Eli A. Boritz, Nicole A. Doria-Rose, Daniel C. Douek, Richard A. Koup, Adrian B. McDermott, Sijy O'Dell, Stephen D. Schmidt, Robert A. Seder & Nancy J. Sullivan

19. CAMRIS, Contractor for NIAID, Bethesda, MD, USA

Debbie L. Bratt

20. High Throughput Screening Center, Washington University School of Medicine, St Louis, MO, USA

William J. Buchser, Mariel J. Liebeskind & Maria A. Vakaki

21. National Institute for Biotechnology in the Negev, Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Be'er-Sheva, Israel

Liel Cohen-Lavi, Eilay Koren & Sinai Sacharen

22. Center for Childhood Infections and Vaccines of Children's Healthcare of Atlanta, Department of Pediatrics, Emory Vaccine Center, Emory University School of Medicine, Atlanta, GA, USA

Meredith E. Davis-Gardner, Venkata-Viswanadh Edara & Mehul S. Suthar

23. US Military HIV Research Program, Henry M. Jackson Foundation for the Advancement of Military Medicine, Walter Reed Army Institute of Research, Silver Spring, MD, USA

Bethany L. Dearlove & Morgane Rolland

24. Microbiology Department, New York University Grossman School of Medicine, New York, NY, USA

Meike Dittmann

25. Institute of Virology, Charité–Universitätsmedizin and German Center for Infection Research (DZIF), Berlin, Germany

Christian Drosten, Terry C. Jones & Barbara Mühlemann

26. Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO, USA

Ali Ellebedy

27. Department of Infectious Diseases, St Jude Children's Research Hospital, Memphis, TN, USA

Thomas P. Fabrizio, John Franks, Trushar Jeevan, Lisa Kercher, Brandi Livingston, Victoria A. Meliopoulos, Stacey Schultz-Cherry, Sanja Trifkovic & Richard J. Webby

28. Department of Surgery, Duke University Medical Center, Durham, NC, USA

Guido Ferrari & Xiaoying Shen

29. Department Viroscience, Erasmus MC, Rotterdam, The Netherlands

Ron A. M. Fouchier, Bart L. Haagmans, Marion P. G. Koopmans & Barry Rockx

30. Department of Medicine, Division of Infectious Diseases, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre & Viviana Simon

31. The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre

32. Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Adolfo García-Sastre, Michael Schotsaert & Viviana Simon

33. Department of Genetics and Genomic Sciences, Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Ana Silvia Gonzalez-Reiche & Harm van Bakel

34. Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA

Aubree Gordon

35. Influenza Research Institute, Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin, Madison, WI, USA

Peter J. Halfmann & Yoshihiro Kawaoka

36. Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA

David D. Ho, Yaxing Huang, Yang Luo, Manoj S. Nair, Maple Wang & Pengfei Wang

37. National Institute of Allergy and Infectious Diseases Integrated Research Facility, Frederick, MD, USA

Michael R. Holbrook

38. Center for Pathogen Evolution, Department of Zoology, University of Cambridge, Cambridge, UK

Sarah L. James, Terry C. Jones, Eric B. LeGresley, Barbara Mühlemann, Antonia Netzl, Sina Türeli, Samuel A. Turner, Samuel H. Wilks & Derek J. Smith

39. Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Yoshihiro Kawaoka

40. Disease Control and Prevention Center, National Center for Global Health and Medicine Hospital, Tokyo, Japan

Yoshihiro Kawaoka

41. Los Alamos National Laboratory, New Mexico Consortium, Los Alamos, NM, USA

Will M. Fischer, Bette Korber, James Theiler & Hyejin Yoon

42. The Shraga Segal Department of Microbiology and Immunology, Ben-Gurion University of the Negev, Be'er Sheva, Israel

Eilay Koren & Sinai Sacharen

43. Broad Institute of MIT and Harvard, Boston, MA, USA

Jacob E. Lemieux

44. Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Margaret J. McElrath

45. Department of Microbiology and Immunology, Institute for Human Infection and Immunity, World Reference Center for Emerging Viruses and Arboviruses, University of Texas Medical Branch, Galveston, TX, USA

Vineet D. Menachery & Scott C. Weaver

46. Duke University Medical Center, Durham, NC, USA

David C. Montefiori

47. Laboratory of Virology, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT, USA

Vincent J. Munster

48. Department of Informatics, J. Craig Venter Institute, La Jolla, CA, USA

Anna M. Niewiadomska, Richard H. Scheuermann & Zachary S. Wallace

49. Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Andrew Pekosz

50. Department of Microbiology and Immunology, University of Iowa, Iowa City, IA, USA

Stanley Perlman

51. Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA, USA

Alessandro Sette

52. Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA

Pei-Yong Shi & Xuping Xie

53. University of Chicago Consortium for Advanced Science and Engineering, University of Chicago, Chicago, IL, USA

Maulik Shukla

54. Data Science and Learning Division, Argonne National Laboratory, Argonne, IL, USA

Maulik Shukla

55. Department of Immunology, St Jude Children's Research Hospital, Memphis, TN, USA

Paul G. Thomas

56. Department of Computer Science and Engineering, University of California, San Diego, CA, USA

Zachary S. Wallace

57. Center for Biologics Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

Wei Wang & Carol D. Weiss

58. Department of Microbiology, Immunology and Genetics Faculty of Health Sciences Ben-Gurion University of the Negev, Be'er Sheva, Israel

Tomer Hertz

59. Department of Pathology & Immunology, Washington University School of Medicine, St Louis, MO, USA

Michael S. Diamond

60. Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO, USA

Michael S. Diamond

61. The Andrew M. and Jane M. Bursky Center for Human Immunology and Immunotherapy Programs, Washington University School of Medicine, St Louis, MO, USA

Michael S. Diamond

Contributions

M.M.D., E.G., M.B.F., F.K., A.G., N.J.S., A.S., P.G.T., D.J.S., M.S.D., D.J.P. and M.S.S. wrote the initial draft. A.A., G.A., R.R.A., R.S.B., D.H.B., J.D.B., L.B., G.B., A.C.M.B., E.A.B., D.L.B., T.L.B., L.-M.B., W.J.B., J.M.C., L.C.-L., T.L.D., M.E.D.-G., B.L.D., H.D., M.D., N.A.D.-R., D.C.D., C.D., V.-V.E., A.E., T.P.F., G.F., W.C.F., R.A.M.F., J.F., A.G.-S., A. Godzik, A.S.G.-R., A. Gordon., B.L.H., P.J.H., D.D.H., M.R.H., Y.H., S.L.J., L.J., T.J., R.M.J., T.C.J., A.J., Y.K., L.K., M.P.G.K., B.K., E.K., R.A.K., E.B.L., J.E.L., M.J.L., Z.L., B.L., J.P.L., Y.L., A.B.M., M.J.M., V.A.M., V.D.M., D.C.M., B.M., V.J.M., J.E.M., M.S.N., A.N., A.M.N., S.O., A.P., S.P., M.C.P., B.R., M.R., P.W.R., S. Sacharen, R.H.S., S.D.S., M. Schotsaert, S.S.-C., R.A.S., M. Sedova, A.S., R.S.S., X.S., P.-Y.S., M. Shukla., V.S., S. Stumpf, L.B.T., J.T., S. Trifkovic, S. Tureli, S.A.T., M.A.V., H.v.B., L.A.V., L.R.V., Z.S.W., L.W., M.W., P.W., W.W., S.C.W., R.J.W., C.D.W., D.E.W., S.M.W., S.P.J.W., B.M.W., S.H.W., X.X., B.Y., H.Y., B.Z., T.H., M.M.D., E.G., M.B.F., F.K., A.G., N.J.S., P.G.T., D.J.S., M.S.D., D.J.P. and M.S.S. provided input, concepts, editing and revisions.

Corresponding authors

Correspondence to [Florian Krammer](#), [Tomer Hertz](#), [Derek J. Smith](#), [Michael S. Diamond](#), [Diane J. Post](#) or [Mehul S. Suthar](#).

Ethics declarations

Competing interests

D.H.B. is listed as a co-inventor on provisional vaccine patents (63/121,482; 63/133,969; 63/135,182). J.D.B. consults for Moderna and Flagship Labs 77 on topics related to viral evolution, and is an inventor on Fred Hutch licensed patents related to viral deep mutational scanning (62/692,398; PCT/US2019/039952; 17/281,540; 19824586.2; 62/935,954; 17/097,853; 62/812,804; PCT/US2020/020429). The Boon laboratory has received unrelated funding support in sponsored research agreements from AI Therapeutics, GreenLight Biosciences, and Nano targeting & Therapy Biopharma. The Boon laboratory has received funding support from AbbVie, for the commercial development of SARS-CoV-2 monoclonal antibodies. A.C.M.B. was a recipient of a licensing agreement with Abbvie for the commercial development of SARS-CoV-2 monoclonal antibodies. M.S.D. is a consultant for Inbios, Vir Biotechnology, Senda Biosciences and Carnival Corporation, and on the scientific advisory boards of Moderna and Immunome. The Diamond laboratory has received unrelated funding support in sponsored research agreements from Vir Biotechnology, Moderna and Emergent BioSolutions. The Ellebedy laboratory received funding under sponsored research agreements that are unrelated to the data presented in the current study from Emergent BioSolutions and from AbbVie. A.E. has received consulting fees from InBios International, Fimbrion Therapeutics, Mubadala Investment Company and Goldman Sachs and is the founder of ImmuneBio Consulting. M.B.F. has funding from Novavax, which is outside the scope of this research. They had no role in the funded research from the SAVES consortium. The Garcia-Sastre laboratory has received research support from Pfizer, Senhwa Biosciences, Kenall Manufacturing, Avimex, Johnson & Johnson, Dynavax, 7Hills Pharma, Pharmamar, ImmunityBio, Accurius, Nanocomposix, Hexamer, N-fold, Model Medicines and Merck, outside of the reported work. A.G.-S. has consulting agreements for the following companies involving cash and/or stock: Vivaldi Biosciences, Contrafект, 7Hills Pharma, Avimex, Vaxalto, Pagoda, Accurius, Esperovax, Farmak, Applied Biological Laboratories, Pharmamar and Pfizer, outside of the reported work. A.G.-S. is listed as an inventor on patents and patent applications on the use of antivirals and vaccines for the treatment and prevention of virus infections and cancer, owned by the Icahn School of Medicine at Mount Sinai, New York, outside of the reported work (5,820,871; 5,854,037; 6,001,634; 6,146,642; 6,451,323; 6,468,544;

6,544,785; 6,573,079; 6,635,416; 6,649,372; 6,669,943; 6,740,519; 6,852,522; 6,866,853; 6,884,414; 6,887,699; 7,060,430; 7,384,774; 7,442,379; 7,494,808; 7,588,768; 7,833,774; 8,012,490; 8,057,803; 8,124,101; 8,137,676; 8,591,881; 8,629,283; 8,673,314; 8,709,442; 8,709,730; 8,765,139; 8,828,406; 8,999,352; 9,051,359; 9,096,585; 9,175,069; 9,217,136; 9,217,157; 9,238,851; 9,352,033; 9,371,366; 9,387,240; 9,387,242; . 9,549,975; 9,701,723; 9,708,373; 9,849,172; 9,908,930; 9,968,670; 10,035,984; .10,098,945; 10,131,695; 10,137,189; 10,179,806; 10,251,922; 10,308,913; 10,543,268; 10,544,207; 10,583,188; 10,736,956; 11,254,733; 11,266,734). A. Gordon serves on a scientific advisory board for Janssen, Erasmus MC has a proprietary IP on MERS. B.K. is part of provisional patent applications for strategies for next-generation SARS-CoV-2 vaccines that address diversity (63/256,848, 17/234,590; S133955). The Icahn School of Medicine at Mount Sinai has filed patent applications relating to SARS-CoV-2 serological assays that list V.S. as a co-inventor (62/994,252; 63/018,457; 63/020,503; 63/024,436) and NDV-based SARS-CoV-2 vaccines that list F.K. as a co-inventor (63/251,020). Mount Sinai has spun out a company, Kantaro, to market serological tests for SARS-CoV-2. F.K. has consulted for Merck and Pfizer (before 2020), and is currently consulting for Pfizer, Seqirus, 3rd Rock Ventures and Avimex. The Krammer laboratory is also collaborating with Pfizer on animal models of SARS-CoV-2. V.D.M. has filed a patent on the reverse genetic system and reporter SARS-CoV-2 (63/000,713; 63/041,667). D.C.M. receives funding from Moderna to perform blinded assessments of vaccine-elicited neutralizing antibody responses in clinical studies of their COVID-19 vaccines. A.S. is a consultant for Gritstone Bio, Flow Pharma, Arcturus Therapeutics, ImmunoScape, CellCarta, Avalia, Moderna, Fortress and Repertoire. P.-Y.S. laboratory has received funding support in sponsored research agreements from GSK, Pfizer, Gilead, Novartis, Merck, IGM Biosciences and Atea Pharmaceuticals. P.-Y.S. is a member of the scientific advisory boards of AbImmune and is Founder of FlaviTech. M.S.S. serves on the advisory board for Moderna and Ocugen. P.G.T. serves on the scientific advisory board for Immunoscape and Cytoagents and has consulted for Johnson and Johnson. P.G.T. has received travel support and honoraria from Illumina and 10x Genomics. P.G.T. has patents related to viral infection treatment and T cell receptor biology (11,083,725; 2021/0299118; US-2019-0040381; 17/616,279;

WO2021/214637). S.P.J.W. has filed a patent with Washington University for VSV- SARS-CoV-2 mutants to characterize antibody panels (PCT/US2021/027275). S.P.J.W. has received unrelated funding support in sponsored research agreements with Vir Biotechnology, AbbVie, and sAB therapeutics.

Peer review

Peer review information

Nature thanks Ravindra Gupta, Volker Thiel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

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

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

DeGrace, M.M., Ghedin, E., Frieman, M.B. *et al.* Defining the risk of SARS-CoV-2 variants on immune protection. *Nature* **605**, 640–652 (2022). <https://doi.org/10.1038/s41586-022-04690-5>

- Received: 25 November 2021
- Accepted: 24 March 2022
- Published: 31 March 2022

- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04690-5>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04690-5>

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

[Download PDF](#)

- Article
- Open Access
- [Published: 01 April 2022](#)

In situ recording of Mars soundscape

- [S. Maurice](#) [ORCID: orcid.org/0000-0001-5702-8002¹](#),
- [B. Chide](#) [ORCID: orcid.org/0000-0002-5124-9119²](#),
- [N. Murdoch](#) [ORCID: orcid.org/0000-0002-9701-4075³](#),
- [R. D. Lorenz⁴](#),
- [D. Mimoun](#) [ORCID: orcid.org/0000-0002-3427-2974³](#),
- [R. C. Wiens](#) [ORCID: orcid.org/0000-0002-3409-7344²](#) nAff⁵,
- [A. Stott³](#),
- [X. Jacob](#) [ORCID: orcid.org/0000-0001-6718-4981⁶](#),
- [T. Bertrand](#) [ORCID: orcid.org/0000-0002-2302-9776⁷](#),
- [F. Montmessin](#) [ORCID: orcid.org/0000-0002-4187-1457⁸](#),
- [N. L. Lanza](#) [ORCID: orcid.org/0000-0003-4445-7996²](#),
- [C. Alvarez-Llamas](#) [ORCID: orcid.org/0000-0001-7793-7000⁹](#),
- [S. M. Angel¹⁰](#),
- [M. Aung¹¹](#),
- [J. Balaram¹¹](#),
- [O. Beyssac¹²](#),
- [A. Cousin¹](#),
- [G. Delory¹³](#),
- [O. Forni](#) [ORCID: orcid.org/0000-0001-6772-9689¹](#),
- [T. Fouchet](#) [ORCID: orcid.org/0000-0001-9040-8285⁷](#),
- [O. Gasnault](#) [ORCID: orcid.org/0000-0002-6979-9012¹](#),
- [H. Grip¹¹](#),
- [M. Hecht](#) [ORCID: orcid.org/0000-0002-4114-4583¹⁴](#),

- [J. Hoffman¹⁵](#),
- [J. Laserna⁹](#),
- [J. Lasue ORCID: orcid.org/0000-0001-9082-4457¹](#),
- [J. Maki¹¹](#),
- [J. McClean¹⁴](#),
- [P.-Y. Meslin ORCID: orcid.org/0000-0002-0703-3951¹](#),
- [S. Le Mouélic ORCID: orcid.org/0000-0001-5260-1367¹⁶](#),
- [A. Munguira ORCID: orcid.org/0000-0002-1677-6327¹⁷](#),
- [C. E. Newman ORCID: orcid.org/0000-0001-9990-8817¹⁸](#),
- [J. A. Rodríguez Manfredi ORCID: orcid.org/0000-0003-0461-9815¹⁹](#),
- [J. Moros⁹](#),
- [A. Ollila²](#),
- [P. Pilleri¹](#),
- [S. Schröder ORCID: orcid.org/0000-0003-1870-3663²⁰](#),
- [M. de la Torre Juárez ORCID: orcid.org/0000-0003-1393-5297¹¹](#),
- [T. Tzanetos¹¹](#),
- [K. M. Stack¹¹](#),
- [K. Farley¹¹](#),
- [K. Williford ORCID: orcid.org/0000-0003-0633-408X^{11,21}](#) &
- [the SuperCam team](#)

Nature volume **605**, pages 653–658 (2022)

- 5501 Accesses
- 972 Altmetric
- [Metrics details](#)

Subjects

- [Atmospheric dynamics](#)
- [Characterization and analytical techniques](#)

Abstract

Before the Perseverance rover landing, the acoustic environment of Mars was unknown. Models predicted that: (1) atmospheric turbulence changes at centimetre scales or smaller at the point where molecular viscosity converts kinetic energy into heat¹, (2) the speed of sound varies at the surface with frequency^{2,3} and (3) high-frequency waves are strongly attenuated with distance in CO₂ (refs. ^{2,3,4}). However, theoretical models were uncertain because of a lack of experimental data at low pressure and the difficulty to characterize turbulence or attenuation in a closed environment. Here, using Perseverance microphone recordings, we present the first characterization of the acoustic environment on Mars and pressure fluctuations in the audible range and beyond, from 20 Hz to 50 kHz. We find that atmospheric sounds extend measurements of pressure variations down to 1,000 times smaller scales than ever observed before, showing a dissipative regime extending over five orders of magnitude in energy. Using point sources of sound (Ingenuity rotorcraft, laser-induced sparks), we highlight two distinct values for the speed of sound that are about 10 m s⁻¹ apart below and above 240 Hz, a unique characteristic of low-pressure CO₂-dominated atmosphere. We also provide the acoustic attenuation with distance above 2 kHz, allowing us to explain the large contribution of the CO₂ vibrational relaxation in the audible range. These results establish a ground truth for the modelling of acoustic processes, which is critical for studies in atmospheres such as those of Mars and Venus.

Main

Before the landing of Perseverance (18 February 2021), no pressure fluctuations had ever been monitored on Mars at a frequency >20 Hz, namely, in the acoustic domain. The recording of sounds offers the unique opportunity to study the atmosphere as the main natural source of sound and as the propagation medium for acoustic waves. From the knowledge of Mars atmospheric pressure (about 0.6 kPa) and the physical properties of CO₂, one can predict (see [Methods](#)) that: the acoustic impedance results in approximately 20 dB weaker sounds on Mars than on Earth if produced by the same source, the speed of sound should be around 240 m s⁻¹ near the surface and acoustic waves are heavily damped in CO₂ at these atmospheric pressures and temperatures. A few studies^{2,3} proposed very detailed models

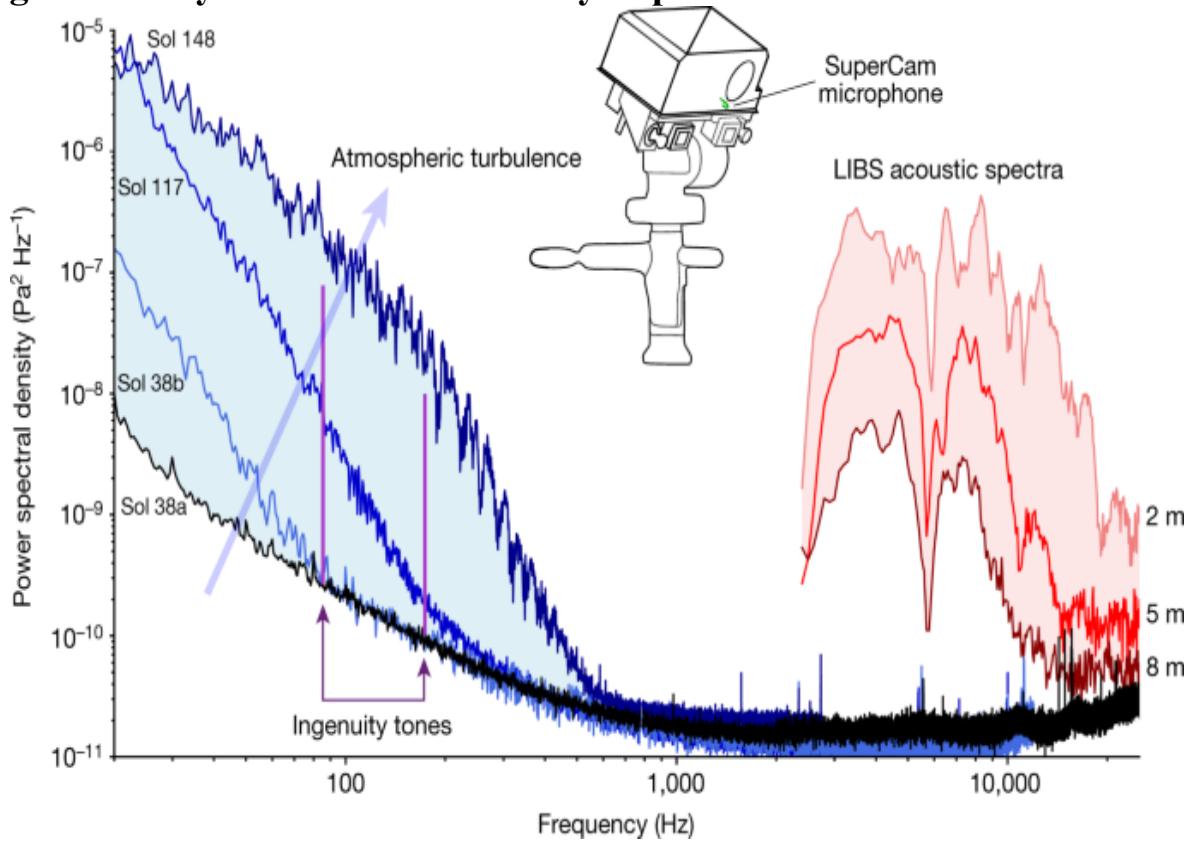
of acoustic propagation on Mars but with large discrepancies between their results because of a lack of experimental data at low pressure and appropriate temperatures, and the difficulty of characterizing attenuation in a closed environment. Acoustic data are also sensitive to wind speed and direction and, to a lesser extent, other environmental parameters^{5,6}. As such, owing to the high sampling frequency of microphones (up to 100 kHz), the acoustic data allow us to explore the atmospheric behaviour on a microscale that has never been accessible before on Mars.

The SuperCam instrument suite^{7,8} on Perseverance carries an electret microphone, similar to that carried by the Mars Polar Lander⁹, lost during atmospheric entry, and the Phoenix spacecraft¹⁰, on which technical issues prevented the device from being operated. SuperCam's microphone is able to record air pressure fluctuations from 20 Hz to 12.5 kHz or 50 kHz, at sampling rates of 25 kHz or 100 kHz, respectively. After landing (Martian solar day 'Sol' 0; one Sol = 88,775 s), the microphone was turned on for the first time on Sol 1 while the mast was still stowed. Since deployment on Sol 2, the microphone is approximately 2.1 m above the ground; it has performed nominally up to the time of writing. SuperCam also consists of a laser-induced breakdown spectroscopy (LIBS) capability to analyse the chemistry of Mars at stand-off distances from 1.5 to 7 m (refs. ^{7,8}). When the laser pulse interacts with the target, a luminous plasma emits characteristic optical emission lines of the elements present in the target¹¹. Plasma expansion generates a shock wave that decouples from the plasma within the first microsecond after laser interaction¹² and results in a clearly detectable acoustic signal^{13,14}. Moreover, Perseverance carries a second microphone as part of the Entry, Descent, and Landing Camera (EDLCAM¹⁵), which has a frequency response from 20 Hz to 20 kHz at a sampling rate of 48 kHz. The EDL microphone is mounted on the port side of the rover, 1 m above the ground. It was activated on Sol 2.

Figure 1 provides an overview of sounds acquired by SuperCam's microphone (see [Methods](#)). Sol 38a is the quietest recording in our dataset. Later on that same day (Sol 38b), the power spectral density (PSD) increases above the quiet state at frequencies below 100 Hz. On Sol 117, we associate this increase of power to an increase in the turbulent activity, which extends up to 300 Hz; this is the situation we observe most often. The recording of

Sol 148 is the most active one shown, with the same shape starting towards higher frequencies but with a slope break near 200 Hz; turbulence is detected up to 600 Hz. All non-saturated atmospheric recordings from Sol 0 to Sol 216 fit between the boundaries given by the Sol 38a and Sol 148 spectra. The laser-excited plasma generates a short, roughly 300- μ s acoustic pulse (see [Methods](#)), with 95% of its energy between 3 and 15 kHz. Various spectral notches are caused by acoustic interferences owing to echoes from the base of the microphone itself (6 kHz and 12 kHz) or from nearby rocks. The total intensity varies as a function of target distance, as shown for recordings at 2 m, 5 m and 8 m. During laser-induced spark recording sessions, the atmospheric signal below 1 kHz is masked by electromagnetic interference⁸. The Ingenuity rotorcraft tones (see [Methods](#)) are also shown.

Fig. 1: Variety of sounds recorded by SuperCam.



Atmospheric spectra spread over the light blue area; turbulence increases in the direction of the arrow. LIBS acoustic spectra spread over the light red area. Ingenuity tones are recorded at 84 Hz and 168 Hz (purple). The black

spectrum is the quietest recording so far below 1 kHz. SuperCam's microphone is located on the rover mast (green).

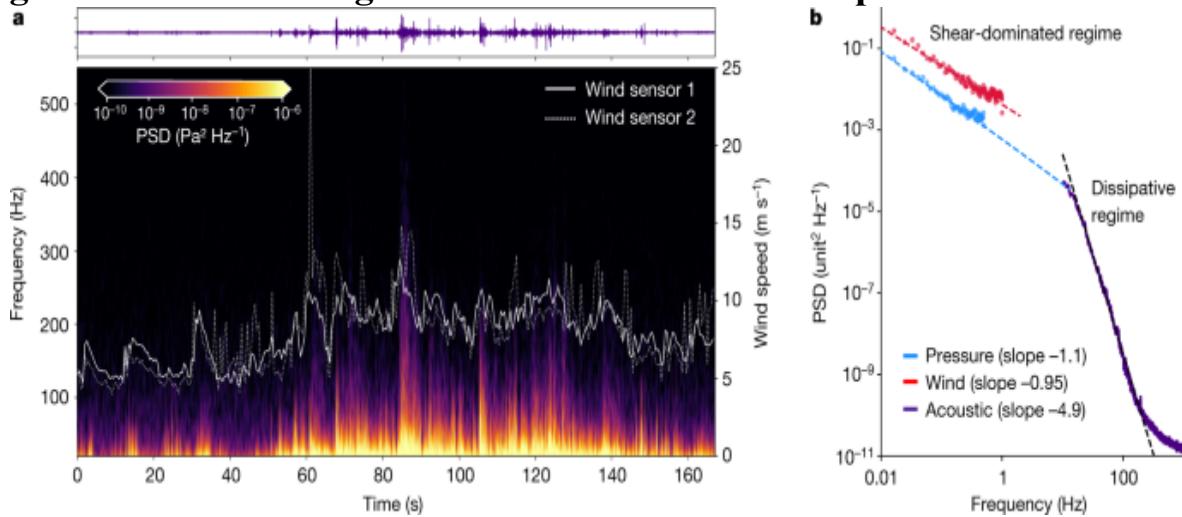
Atmospheric turbulence

The Martian planetary boundary layer (PBL) is the part of the atmosphere in contact with the surface¹⁶, extending to several km. It is prone to convective turbulence and vertical mixing during daytime, owing to the thin atmosphere and low surface thermal inertia that induce strong and unstable near-surface temperature gradients^{17,18,19}. This turbulence translates into high-frequency variations in atmospheric pressure, wind speed and temperature that can be measured by in situ instruments. Conversely, during night-time, the strong radiative cooling of the atmosphere induces highly stable conditions, which efficiently inhibit most convection and turbulence¹⁶. Analysing the PBL at the surface is therefore important to understand how the Martian atmosphere transports and mixes heat, momentum, aerosols and chemical species²⁰. The Mars Environmental Dynamics Analyzer (MEDA²¹) instrument on Perseverance and the meteorological suites of previously landed missions^{20,22} typically measure pressure, temperature and wind fluctuations with sampling frequencies of 0.1 Hz to 10 Hz. These instruments study the turbulence variability^{23,24} and the Martian turbulent energy cascade^{1,17,25}.

Specifically, we report here the observation of the dissipative turbulence regime in the PBL, in which the InSight mission could see a hint of a regime change at the limits of the instrument capability¹. This regime, in which molecular viscosity dissipates the turbulent kinetic energy into heat, is now fully characterized by a rapid decrease of the power spectrum with increasing frequency (Fig. 1, 2b) over roughly five orders of magnitude. The scale at which the viscous dissipation becomes notable is characterized by the Kolmogorov length scale²⁶, $\eta = (v^3/\varepsilon)^{0.25}$, in which v is the kinematic viscosity and ε is the turbulence energy dissipation rate per unit mass, typically around $0.001 \text{ m}^2 \text{ s}^{-1}$ and $0.005 \text{ m}^2 \text{ s}^{-1}$ on Mars, respectively¹⁷. Thus η is about 0.02 m and the timescale of these small eddies, $t_\eta = (v/\varepsilon)^{0.5}$, is about 0.45 s. Hence the dissipation regime should be observable at frequencies above 2 Hz on Mars, at centimetre or smaller scales only (on Earth, this transition occurs at millimetre scales or smaller¹⁷). This

theoretical prediction is confirmed by the acoustic data; the threshold moves with frequency, depending on the dissipation rate^{25,27}. The balance between energy production and molecular dissipation controls the total amount of turbulent kinetic energy in the boundary layer and, as such, the dissipation mechanism is intrinsically linked to the PBL dynamics; a larger dissipation leads to a faster turbulence decay, in turn suppressing small-scale wind gustiness, and vice versa.

Fig. 2: Sound recordings and correlation with atmospheric data.



Recording of Sol 38b. **a**, On top, the y axis of the time series ranges from -0.2 to 0.2 Pa. The spectrogram (bottom) shows bursts that extend to 300 Hz. Overlaid, with the y axis on the right, are wind speeds from MEDA booms. **b**, The PSD calculated for SuperCam's microphone (in $\text{Pa}^2 \text{Hz}^{-1}$ for 167 s) and for MEDA pressure (in $\text{Pa}^2 \text{Hz}^{-1}$ for 51 min around the microphone acquisition time) and MEDA wind data (in $(\text{m s}^{-1})^2 \text{Hz}^{-1}$). The wind PSD is artificially offset by 10^{-2} in the y axis.

The microphone records rapid deviations from ambient pressure (>20 Hz) that are correlated to variations in the wind flow, as shown by Fig. 2a, in which a spectrogram of Sol 38b microphone data (see [Methods](#)) is overlaid with the wind speed as measured by the MEDA (see [Methods](#)). As expected^{6,13}, there is a clear correlation between the intensity of acoustic data and the wind speed. This can be owing to the flow-induced turbulence from the rover/mast itself but also to the direct sensing of the incoming flow fluctuations, seen to be the dominant factor for outdoor microphones in other

studies⁶. Moreover, the daytime local turbulence is known to increase for larger ambient wind speeds²⁴. The high microphone sampling rate provides an opportunity to observe very intense but short wind gusts, on a timescale of 10 s. In Fig. 2b, the same acoustic data are plotted in the frequency domain and combined with low-frequency measurements of pressure and wind from the MEDA, for a 51-min time period of continuous data around the microphone acquisition. The large difference in slope between the MEDA and microphone data is indicative of regime change. The transition from the probable shear-dominated regime²⁸ to the dissipation regime occurs in this case between 1 and 20 Hz.

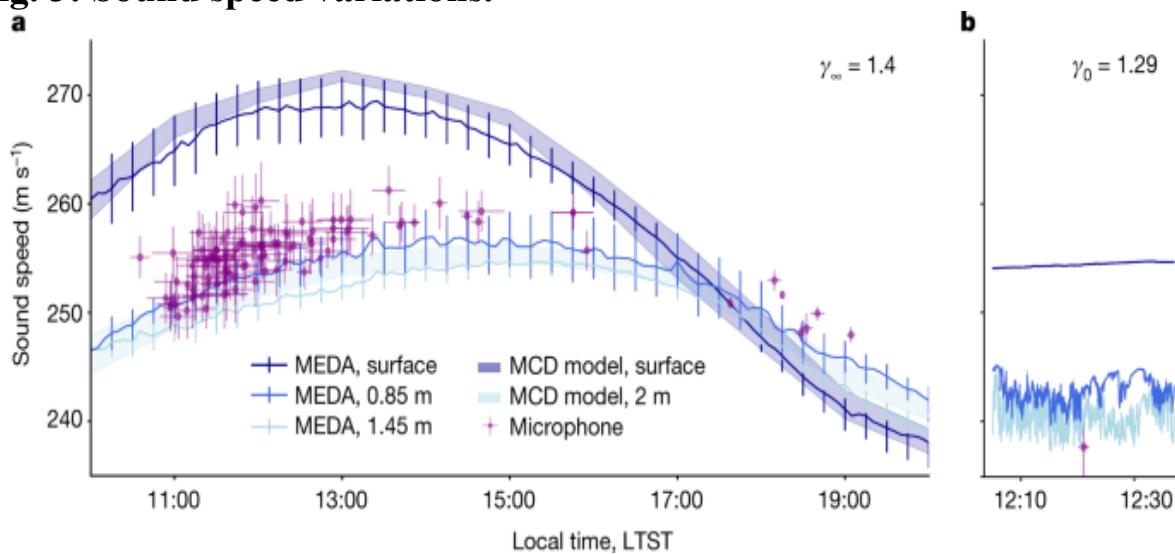
Speed of sound on Mars

In a cold CO₂ atmosphere, the speed of sound is expected to be lower than on Earth. Furthermore, owing to the low pressure and the physical properties of CO₂, we also expect a dispersion of this speed with frequency^{2,3}. On Earth, the adiabatic ratio γ is constant up to a few MHz at ambient pressure²⁹ and sound speed does not vary with frequency near the surface. At low pressure on Mars, still within the framework of small Knudsen numbers³⁰ (10^{-6} at 100 Hz to 2.10^{-4} at 20 kHz), the continuum theory still holds, but energy exchanges at molecular scales are modified. Part of the energy associated with the translational motions of molecules, which constitute the acoustic waves, is spent on the excitation of inner degrees of freedom (vibrational modes and rotational motions). The relaxation of the rotational motion is almost instantaneous, whereas relaxation of the vibrational modes occurs over a much longer timescale, a property of small and rigid polyatomic molecules such as CO₂. If the frequency f is smaller than $f_R = 1/\tau_R$, in which τ_R is the relaxation time, all modes are equally excited and then relaxed. The seven degrees of freedom that result from three translational modes, two rotational modes and one doubly-degenerate vibrational mode (v_2 , bending) lead to an adiabatic index $\gamma_0 = 9/7 = 1.2857$. Conversely, if $f > f_R$, there is no time to relax the vibrational mode; in that case, there are only five active degrees of freedom and $\gamma_\infty = 7/5 = 1.4$. In CO₂ at Earth-ambient pressure, f_R is about 40 kHz (ref. ³¹). This frequency depends on the rate at

which molecules can collide, hence f_R is proportional to the pressure. As a result, at 0.6 kPa, the relaxation frequency is about 240 Hz on Mars.

The recording of pulsed waves generated in LIBS mode provides a unique opportunity to measure directly and repetitively the local speed of sound for acoustic waves above 2 kHz, that is, for $f > f_R$ (see [Methods](#)). From the daytime measurements, sound speeds between 246 m s^{-1} and 257 m s^{-1} are obtained (Fig. [3a](#)), with maximum values between 11:00 to 14:00 Local True Solar Time (LTST) and minimum values around 18:00. The 1σ -dispersion of the sound speed during the approximately 20 min of a target analysis with LIBS is at its maximum at noon (1.5%) and is reduced to 0.5% at 18:00, which highlights the vanishing of the atmospheric turbulence at dusk. These measurements are compared with temperature-derived speeds of sound obtained from: (1) the MEDA temperature datasets at the surface, at the heights of 0.85 m and 1.45 m, and (2) the temperature at the surface and at a height of 2 m given by the Mars Climate Database (MCD^{[32](#)}) (see [Methods](#)), using $\gamma_\infty = 1.4$ (because $f > f_R$). The agreement between the MEDA and MCD predictions is excellent. SuperCam sound speeds are comparable with temperature-derived values at the height of the MEDA's 0.85-m temperature sensor or higher. This is consistent with the fact that the speed is integrated between a height of 2.1 m and the surface, possibly biased towards the surface when the temperature gradient is larger.

Fig. 3: Sound speed variations.



a, Sound speeds as a function of local time from LIBS time-of-flight data in purple. Other sound speeds are calculated at the three heights from the MEDA temperatures and at the surface and at 2-m altitude from MCD simulations; for these conversions, the adiabatic index above f_R is used. Error bars for microphone data: standard deviation of the sound speeds during each laser burst (vertical); total duration of the burst (horizontal). Error bars for the MEDA data: standard deviation of 1-h bins between Sols 37 and 216. **b**, Sound speeds are calculated at three heights from the MEDA temperatures during Ingenuity’s fourth flight; the adiabatic index below f_R is used. The sound speed estimated from the Ingenuity Doppler effect is in purple. Error bars: 95% confidence interval of the Doppler shift fit.

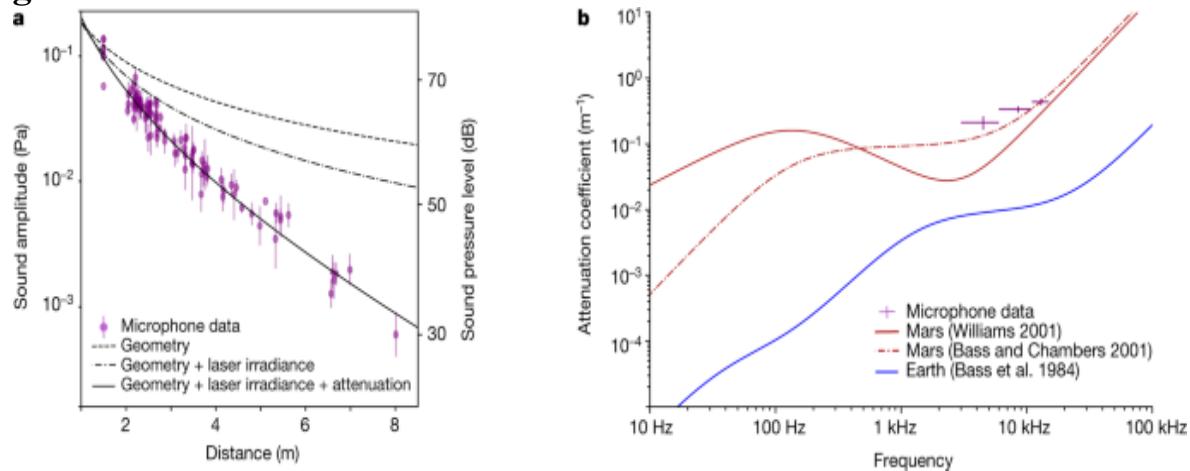
Ingenuity’s blade passage frequency (BPF)³³ is close to a harmonic source centred around 84 Hz, and — in that case — for $f < f_R$ (see [Methods](#)). This signal recorded by SuperCam’s microphone is modulated by the variations of the distance range between the microphone and the helicopter. An emitted frequency at 84.43 Hz and a speed of sound $c = 237.7 \pm 3 \text{ m s}^{-1}$ are estimated on the basis of a fit of the Doppler effect for Ingenuity’s fourth flight (see [Methods](#)). Accounting for the presence of a wind of about 2.5 m s^{-1} along the microphone-to-helicopter line of sight towards the helicopter (MEDA data), the true sound speed is about 240 m s^{-1} at this frequency. At the time of the flight, the atmospheric temperatures ranged between 232 K and 240 K at a height of 1.45 m. Using $\gamma_0 = 1.2857$ (the BPF is below f_R), the temperature-derived speed of sound ranges from 238.8 m s^{-1} to 242.9 m s^{-1} , which is consistent with the speed directly derived from Ingenuity’s flight plus wind (Fig. [3b](#)). As a summary, SuperCam’s microphone highlights a sound speed dispersion of about 10 m s^{-1} in the audible range at the surface of Mars.

Sound attenuation

The most remarkable property of sound propagation on Mars is the magnitude of the attenuation at all frequencies, especially above 1 kHz. The decrease of the LIBS acoustic signal with distance is an opportunity to verify the theory in situ and to test two different attenuation models^{3,4} that suffer from a lack of field data under Mars conditions.

As the spherical LIBS acoustic wave propagates, sound pressure decreases as $1/r$, in which r is the distance between the target and the microphone. This decrease is scaled by a factor $r^{-0.698}$ to account for the variation of laser irradiance⁸, multiplied by $e^{-\alpha r}$, in which $\alpha = \alpha(f)$ is the atmospheric attenuation coefficient as a function of frequency. The frequency spectrum of the LIBS acoustic signal is divided into three bands, which account for the three main lobes observed in Fig. 1: from 3 kHz to 6 kHz, from 6 kHz to 11 kHz and from 11 kHz to 15 kHz. The evolution of the sound amplitude with distance for the second frequency band is shown in Fig. 4a. Over the three bands, we find $\alpha = 0.21 \pm 0.04 \text{ m}^{-1}$ (95% confidence interval of the fit), $\alpha = 0.34 \pm 0.05 \text{ m}^{-1}$ and $\alpha = 0.43 \pm 0.05 \text{ m}^{-1}$ respectively. As expected, high-pitched sounds are strongly attenuated. Compared with a signal emitted at 1 m, attenuation of an 8-kHz wave ranges from -9 dB at 2 m to -40 dB at 8 m. At 5 m, the atmospheric absorption takes precedence over the geometrical attenuation. On Earth, for which $\alpha = 0.01 \text{ m}^{-1}$ for the same frequency³⁴, the attenuation ranges from -6 dB at 2 m to -20 dB at 8 m, and is almost exclusively resulting from the wavefront spreading. To reach an attenuation of -40 dB on Earth, the source would need to be at 65 m.

Fig. 4: Sound attenuation with distance.



a, Sound amplitude as a function of target distance r from LIBS acoustic data between 6 kHz and 11 kHz. The second vertical axis on the right is for sound pressure level in dB. Signal intensities are in dB relative to 20 μPa . Error bars: standard deviation of the acoustic amplitudes during each laser burst. **b**, Comparison of the attenuation models for Mars^{3,4} (computed at 240 K and 740 Pa) and Earth³⁴ (293 K and 30% relative humidity). The

experimental points correspond to this study. Error bars: 95% confidence interval of the fit performed in Fig. 4a (vertical) and width of each frequency range (horizontal).

Such attenuation coefficients are compared with theoretical³ and semi-empirical⁴ attenuation models in Fig. 4b. In situ data tend towards the behaviour described by Bass and Chambers³, with a plateau at frequencies <6 kHz and then an increase for higher frequencies. Conversely, the data do not show an attenuation gap as suggested by the model of Williams⁴. This result confirms the large contribution of CO₂ vibrational relaxation in this frequency range, the same process that explains the two values for the speed of sound (above). However, the attenuation coefficient for the 2–6-kHz band is still higher than that predicted by Bass and Chambers³. It may highlight a different relaxation strength than the one forecasted by the model (see [Methods](#)). However, these measurements do not reach frequencies low enough to constrain the large discrepancies observed between models below 1 kHz.

Mars soundscape

Sound is a new, rich source of information on Mars. Thanks to sensors measuring only a few millimetres in diameter, turbulence-induced noise and artificial sources have been recorded. Acoustic waves are governed by the macroscopic thermodynamic properties of fluids (molar mass, heat capacity and temperature, or — alternatively — compressibility and density).

However, given the small displacements and timescales that come into play, we confirm that energy exchanges at the molecular scale also need to be considered to accurately model the sound propagation parameter variations (speed, attenuation) with frequency. More sound speed measurements at different local times and seasons will allow the study of atmospheric fluctuations at a scale of a few metres on Mars^{35,36,37}. The first in situ retrieval of the acoustic attenuation coefficient already provides new constraints on theoretical models, which are key parameters for geophysical studies in CO₂-dominated atmospheres^{38,39}. Wind and turbulence, driven by heat fluxes, are natural sources of pressure fluctuations on Mars. We show that acoustic data yield new insights into the boundary layer turbulence with

10 to 1,000 times higher temporal resolution than before, highlighting for the first time the dissipative regime and a transition to this regime above a few Hz. Characterizing this regime in more detail, and the associated transition, is necessary to settle the assumptions used in the numerical modelling of the PBL (including large-eddy simulations), telling us what the fraction of missing energy is in the unresolved scales of the models^{40,41}. In the future, this will lead to a measurement of the dissipation rate, related to the diffusion of heat in the atmosphere, which is not well known for Mars at present^{17,42}. Finally, beyond the rumble of the wind, the acoustic signatures of our robotic presence on Mars are rich in information on the health of the rover subsystems.

Methods

This section starts with a reminder on acoustics adapted to Mars. It provides details of the different datasets: acoustic data from the microphones of SuperCam and EDL; artificial sounds from the Mars Oxygen In-Situ Resource Utilization Experiment (MOXIE) and Ingenuity, in addition to the recording of LIBS shock waves and rover noises; wind speed, temperature and pressure data from the MEDA and temperatures extracted from the MCD. Processing methods are presented: computation of PSDs; analyses of LIBS shock wave time series; extraction of the Doppler effect from Ingenuity. Finally, supporting explanations on the attenuation with distance are given.

Acoustics reminder

We justify the three main assertions in the introduction to the text. First, the acoustic impedance describes the strength of a medium to sustain acoustic waves. It is given by $Z = \rho c$ in the far field of a source, in which ρ is the density and c is the speed of sound. Typically, with $\rho = 0.02 \text{ kg m}^{-3}$ and $c = 238 \text{ m s}^{-1}$ (see below), we obtain $Z = 4.76 \text{ kg m}^{-2} \text{ s}^{-1}$ at the surface of Mars, whereas $\rho = 1.217 \text{ kg m}^{-3}$ and $c = 340 \text{ m s}^{-1}$ yield $Z = 413 \text{ kg m}^{-2} \text{ s}^{-1}$ on Earth. This difference of two orders of magnitude translates into signals on Mars being roughly 20 dB weaker than on Earth when produced by the same source. Second, at Mars pressure, the approximately 95% CO₂ atmosphere

can be efficiently modelled as an ideal gas. In the microphone frequency range, and given the small amplitude of the acoustic pressure, sound waves are considered as adiabatic disturbances. It follows that the temperature-derived speed of sound is given by $c^2 = \gamma RT/M$, with R the molar gas constant ($8.314 \text{ J mol}^{-1} \text{ K}^{-1}$), M the molar mass of the atmosphere (43.34 g mol^{-1}), T the temperature in kelvins and γ the adiabatic index. Using $\gamma = 9/7$, the standard value for CO_2 — this value is discussed in the main text — we find $c = 238 \text{ m s}^{-1}$ at 230 K . Third, in a rarefied atmosphere, absorption is intrinsically larger as classical (thermal and viscous) and rotational absorption are inversely proportional to the pressure³. Moreover, at low frequencies, vibrational absorption dominates over the classical types of absorption and rotation. It turns out that the vibrational specific heat of CO_2 is 20 times greater than for N_2 (ref. ³) for a comparison with Earth. Hence the doubly degenerate mode of CO_2 vibration attenuates sounds at low frequencies, whereas the viscosity strongly attenuates frequencies higher than a few kHz. Both effects, in terms of attenuation coefficient per metre, are one or two orders of magnitude stronger than on Earth at the same frequency.

Acoustic dataset

The SuperCam microphone dataset used in this study extends from Sol 1 to Sol 216, when the first solar conjunction of the mission occurred. At this date, a total of 4 h and 40 min of Martian sounds have been recorded, including atmospheric turbulence (46% of the total duration), the accompanying pressure waves of LIBS sparks (12%) and mechanical noises (for example, MOXIE⁴³, Ingenuity³³, mast rotation of Perseverance, Mastcam-Z mechanisms, 42%). In the same period, the EDL¹⁵ microphone has recorded a total of 56 min of Martian sounds, mainly during rover operations (for example, rover drive, arm motion). Extended Data Table 1 lists all acoustic files used or mentioned in this study, except for those shown in Figs. [3a](#), [4a](#), which are too numerous to cite individually (see below).

SuperCam's microphone records air pressure fluctuations from 20 Hz to 12.5 kHz at a 25-kHz sampling frequency, and up to 50 kHz when the 100-kHz sampling mode is used. The analogue signal from the microphone,

ranging from 0 to 5 V, is digitized (12-bit depth) using one of four electronic gains to boost the sensitivity from 0.6 to 21 V Pa^{-1} and the resolution from 2 to 0.06 mPa. Gain 0 is used to record the accompanying sounds from LIBS sparks on calibration targets, gains 1, 2 and 3 to record the accompanying sounds from LIBS sparks at various distances and gain 3 is used for atmospheric recordings. Recordings of the atmosphere and mechanical noises are generally 167 s long. The EDL microphone can record 10 mn and longer time series, with a fixed gain amplifier followed by a 24-bit/44-kHz digitizer (the key characteristics are summarized in Extended Data Table 2).

Typical LIBS sequences consist of 30 laser shots at the same position fired at 3 Hz (for technical reasons, only 29 shots are recorded). For a given target, such a sequence is usually repeated 5–10 times, on fresh sampling points separated by a few millimetres. The laser-induced acoustic signal is monitored at a 100-kHz sampling frequency for 60 ms around each laser pulse. The start of the recording window is precisely timed on the laser trigger, so that the propagation time of the sound wave can be measured with an uncertainty $<10 \mu\text{s}$. Up to Sol 216, SuperCam’s microphone has recorded sound sequences for 123 Martian targets located at distances ranging from 2.05 m (target Garde) to 8.01 m (target Pepin) from the microphone. On seven occasions, it has also recorded the acoustic signal related to LIBS measurements of the titanium (Ti) calibration target⁴⁴ located on the rover deck 1.51 m from the microphone.

For the derivation of the sound speed (Fig. 3a), targets farther than 6 m are excluded because of a small signal-to-noise ratio that prevents a good time-of-flight measurement. Recordings of the LIBS acoustic signals from Ti are also excluded, as their sound propagates above the rover and are biased by extra heating and turbulence induced by the warm body of the rover. In total, 109 targets between Sol 1 and Sol 216 are considered for Fig. 3. For the attenuation study (Fig. 4), regolith or loose material targets, which generally lead to a lower sound amplitude, are excluded, as well as out-of-focus points for the same reason. In total, 96 targets are used. The measurements from Ti are included and provide a useful constraint on the attenuation at short distance. As the laser energy used on Ti is lower than the laser energy used on Mars targets (110 A pumping current on Ti compared with 155 A on Mars targets), the PSD amplitude from the Ti measurements are normalized by a factor of 155/110, as the pumping current is proportional to the laser

energy, which is proportional to the laser irradiance, as the spot size and the pulse duration remains the same.

Dataset of artificial sounds

Both EDL and SuperCam microphones are also used to record sounds produced by the rover. They help to inform operators about equipment health (for example, rover driving, MOXIE) and provide sources of sound that are well localized in space and time (for example, during Ingenuity flights or LIBS sparks). All recordings also pick up some perturbations, such as intense single-frequency emissions at 195 Hz, 198.75 Hz and the following harmonics at 780 Hz and 795 Hz that result from the rover's internal heat pump used for thermal management. Sounds of the rover's instruments or pumps propagate through both structural vibrations (microphonics) and acoustic propagation in the atmosphere.

The EDL microphone¹⁵ was used to record the rover drive on Sol 16 (Extended Data Fig. 1a). Broad, quasi-continuous 'screech' signals in the 520–700 Hz, 1.2–1.4 kHz and 1.6–1.9 kHz bandwidths are assumed to arise directly from frictional interaction of the metal wheel tread with surface rocks. Sonorous transients or 'clanks' are seen at 13 s with several narrow-frequency components but with a lower total sound intensity than the aforementioned phenomenon. It is suggested that these are structural resonances of mobility system elements (for example, suspension) excited by near-impulsive changes in loading, for example, when a wheel slips off the edge of a rock.

The Ingenuity rotorcraft³³ provides a localized but moving source of sound on Mars. On Sol 69 during the fourth flight, SuperCam's microphone recorded the entire 116-s duration of the flight. A prominent acoustic signal, up to $2 \times 10^{-7} \text{ Pa}^2 \text{ Hz}^{-1}$ (1.5 mPa sound pressure level), associated with the BPF at 84 Hz and its first overtone 168 Hz was detected by SuperCam's microphone (Extended Data Fig. 1b). All phases of the flight are visible but the take-off occurred during a gust (at the rover location) as high as 20 mPa. The BPF clearly stands out but its overtone is much fainter, owing to greater atmospheric absorption at higher frequency. After landing, the microphone captured the blades spinning down.

The MOXIE instrument⁴³ operates every 1–2 months to produce a few grams of gaseous O₂. The primary objective of these repeated operations is to look for possible degradation of the O₂ production efficiency associated with the harsh environment of Mars. MOXIE uses SuperCam’s microphone recordings for independent diagnosis of compressor performance, including precise measurements of the motor rotation rates, as indicated by the fundamental frequency of the observed comb of harmonics. Distinct transitions in Extended Data Fig. 1c, recorded during a night-time run (Sol 81), correspond to commanded changes in motor speed from 50 to 58.3 Hz. The loudest harmonics are near 500 Hz, at which several more frequencies are also excited. This range corresponds to resonant frequencies of the MOXIE instrument, as observed during dynamics testing. Even recorded during one of the quietest times of the day, the amplitude of the signal only reaches 1.5 mPa.

Recording LIBS sparks was the main rationale to develop SuperCam’s microphone to infer physical properties of rock targets, such as their hardness¹³. Typical LIBS sequences consist of 30 laser shots at the same position per observation, fired at 3 Hz (Extended Data Fig. 1d). LIBS operations are monitored by the microphone at a 100-kHz sampling rate for 60 ms around each laser pulse. The mean amplitude of the signal is 0.25 ± 0.08 Pa (1σ) for this shot sequence.

Wind speed, temperature and pressure data

The wind, temperature and pressure data are recorded from the MEDA²¹ instrument. The wind data are acquired up to 2 Hz, pressure at 1 Hz and temperature at 1 or 2 Hz. Wind speed and direction are independently acquired from two individual booms separated by 120° (termed boom 1 and boom 2), for which one is preferred for a given wind direction (see accuracies and resolutions in Extended Data Table 2).

The MCD³² provides climate predictions derived from 3D simulations of Mars atmosphere performed with the Mars global climate model developed at the Laboratoire de Météorologie Dynamique (<http://www-mars.lmd.jussieu.fr>). The Laboratoire de Météorologie Dynamique Mars global climate model is described in ref. ³² but — since then — it has

adopted more sophisticated and realistic modelling for the CO₂, dust and water cycles, photochemistry, radiative transfer and so on. In this work, we use the climatology scenario⁴⁵ from the MCD Version 5.3, in which: (1) the simulated spatial and vertical dust distributions are reconstructed from observations during Martian years 24 to 31 without global dust storms (thus representative of standard climate conditions) and (2) average solar extreme ultraviolet conditions are assumed. In this study, the MCD outputs (surface temperature and atmospheric temperature at 2 m above the surface) are provided for daytime local times in increments of 1 h and between Ls = 5.2° and Ls = 104.7° in increments of 10° to capture the seasonal variations in temperature.

Spectral analysis of acoustic data

The microphone data from SuperCam are converted from volts to pascals using the instrument sensitivity for each gain (0.6, 1.3, 5.3 and 21.6 V Pa⁻¹, corresponding to amplification factors of 29 to 972). The microphone's electronic response function for each gain (bandpass filter between 100 Hz and 10 kHz) is used to correct raw spectra below 100 Hz and above 10 kHz. EDL microphone data are not converted into physical units. PSDs represented in Fig. 1, 2b were computed from a Fourier transform, using a Welch's estimator. Spectrograms represented in Fig. 2a and Extended Data Fig. 1b, c are computed with a Hanning window of 2 s. Extended Data Figure 1a is computed with a window of 1 s and Extended Data Fig. 1d with a window of 5 ms.

Time series (laser-induced spark recordings)

The creation of the laser-induced plasma is accompanied by a shock wave, which can be described as an N-wave acoustic pulse⁴⁶ primarily, a short, approximately 300-μs-long compression/rarefaction acoustic signal. This signal is followed by echoes on nearby rocks and the rover structure, plus diffraction. The whole acoustic signal typically lasts less than 5 ms. A bandpass filter is applied to remove electromagnetic interferences, atmospheric signal below 2 kHz and to reduce noises above 20 kHz. There are residuals of the laser warm-ups but they do not affect the determination of the sound speed (Extended Data Fig. 2).

Time series data are used to calculate the local speed of sound. The distance to each target, which is returned by the instrument's autofocus, is known to an accuracy of $\pm 0.5\%$ (ref. [8](#)). The laser trigger time is known to a few microseconds and the shock wave becomes sonic after $1 \mu\text{s}$ (ref. [12](#)), which is less than 0.1% of the propagation time. The arrival of the pressure wave is considered to be detected when the signal increases 3σ above the background.

Doppler effect (Ingenuity recording)

The fourth (Sol 69), fifth (Sol 76), sixth (Sol 91) and eighth (Sol 120) flights of Ingenuity were recorded by SuperCam at a 25-kHz sampling rate. We use data from the fourth flight, as this flight came closer to Perseverance than any other flight SuperCam could record. During this flight, Ingenuity climbed to an altitude of 5 m, accelerated to 3.5 m s^{-1} , travelled 130 m at constant height, decelerated, turned around and returned to its base by the same route. Taking off at a distance of 76 m from Perseverance, it came as close as 69 m and moved as far as 123 m from Perseverance (Extended Data Fig. [3b](#), bottom).

On the PSD obtained during the whole recording, the BPF (two times the rotation rate for a two-bladed rotor) at 84 Hz and its first harmonic at 168 Hz are clearly visible above the background, which itself is higher than that of Sol 38a, a very quiet recording on Mars (Extended Data Fig. [3a](#)). There is no other tone above the background. There is a period of atmospheric turbulence up to 56 s into the recording that explains why the spectrum at low frequency is above that of Sol 38a. Discontinuities in the amplitude of the tones at 84 Hz and 168 Hz are visible during the cruise phase of the flight. Such a modulation beat results from the interferences of two signals with slightly different frequencies (about 50 mHz apart), each originating from the two blades that are frequency shifted. The study of this phase shift is outside the scope of this paper.

Each tone is fitted by a Gaussian function every 0.5 s. In the main text, we report on the study of the BPF at 84 Hz. The received frequency varies along Ingenuity's flight (Extended Data Fig. [3b](#), top) as a function of the variation of the distance range between the rover and the helicopter. The received

frequency, the classical Doppler effect, varies by $\pm 1.5\%$. The fit of this tone, when the atmosphere is quiet ($t > 60$ s), as a function of the range rate, yields $f = 84.44$ Hz for the BPF at the source and $c = 237.7 \pm 3$ m s $^{-1}$. A similar fit to the first harmonic yields $f = 168.90$ Hz and $c = 236.9 \pm 4$ m s $^{-1}$, which are coherent with values derived from the BPF.

Sound attenuation with distance

As the sound wave propagates through an atmosphere, part of the acoustic energy is transferred to the propagation medium as heat by an absorption mechanism called atmospheric (or intrinsic) attenuation. This process has been largely described and validated for the atmosphere on Earth^{34,47,48}. The atmospheric attenuation, linked to the motion of molecules, depends on the frequency of the wave. It can be attributed to two phenomena. First, the classical attenuation includes the heat loss caused by viscous friction and losses from non-adiabatic heat diffusion between the compression and rarefaction areas. This phenomenon is all the more important for short-period waves, when there is less time to establish an equilibrium. The classical attenuation is proportional to the square of the frequency. The second phenomenon is the molecular attenuation, owing to the excitation of the internal degrees of freedom of the polyatomic molecules (rotational and vibrational modes), each one taking some time, called ‘relaxation time’, to return to equilibrium. The shorter the period of the wave, compared with the relaxation time, the less time molecules have to relax their energy and hence the greater the absorption of the acoustic energy.

This theory has been applied to the atmosphere of Mars to compute empirical attenuation models^{3,4} (Fig. 4). Although Mars models are in good agreement above 10 kHz (classical attenuation is well constrained thanks to a good knowledge of dynamic viscosity and thermal conductivity for CO₂ as a function of temperature), they differ strongly in the infrasonic and part of the audible range, depending on the way molecular relaxation is modelled. The Williams model⁴ extrapolates experimental data for CO₂ that have been acquired at 1 bar and above 273 K. The model considers that molecular attenuation increases proportionally to the frequency, reaches a maximum value at the relaxation frequency, f_R (240 Hz), and then decreases as $1/f$. On

the other hand, the Bass and Chambers model³ discriminates molecular relaxation into rotational and vibrational relaxations. Rotational relaxation is modelled by f squared, just as in classical attenuation. For vibrational relaxation, the model considers that, below f_R , vibrational attenuation grows as f squared. Above f_R , vibration modes are not excited and the vibrational attenuation stays at a constant level.

As a supplementary note, CO₂ has three vibrational modes at 1,341 cm⁻¹ (ν_1 symmetric stretching), 667 cm⁻¹ (ν_2 degenerate bending) and 2,349 cm⁻¹ (ν_3 asymmetric stretching). Associated vibrational temperatures are 1,890 K, 960 K and 3,360 K, respectively. At 240 K, the respective contribution of each mode to the vibrational specific heat are 3.7% (ν_1), 96.3% (ν_2) and <0.1% (ν_3) at 240 K. This justifies why first-order models can only consider the contribution of the ν_2 bending mode to the vibrational specific heat.

Data availability

All acoustic data are publicly available at the Planetary Data System Geosciences Node: <https://doi.org/10.17189/1522646>.

References

1. Banfield, D. et al. The atmosphere of Mars as observed by InSight. *Nat. Geosci.* **13**, 190–198 (2020).
2. Petculescu, A. & Lueptow, R. M. Atmospheric acoustics of Titan, Mars, Venus, and Earth. *Icarus* **186**, 413–419 (2007).
3. Bass, H. E. & Chambers, J. P. Absorption of sound in the Martian atmosphere. *J. Acoust. Soc. Am.* **109**, 3069–3071 (2001).
4. Williams, J.-P. Acoustic environment of the Martian surface. *J. Geophys. Res.* **106**, 5033–5041 (2001).
5. Chide, B. et al. Experimental wind characterization with the SuperCam microphone under a simulated martian atmosphere. *Icarus* **354**, 114060

(2021).

6. Morgan, S. & Raspé, R. Investigation of the mechanisms of low-frequency wind noise generation outdoors. *J. Acoust. Soc. Am.* **92**, 1180–1183 (1992).
7. Wiens, R. C. et al. The SuperCam instrument suite on the NASA Mars 2020 rover: body unit and combined system tests. *Space Sci. Rev.* **217**, 4 (2021).
8. Maurice, S. et al. The SuperCam instrument suite on the Mars 2020 rover: science objectives and Mast-Unit description. *Space Sci. Rev.* **217**, 47 (2021).
9. Delory, G. T., Luhmann, J., Friedman, L. & Betts, B. Development of the first audio microphone for use on the surface of Mars. *J. Acoust. Soc. Am.* **121**, 3116–3116 (2007).
10. NASA. Mars Descent Imager (MARDI); www.nasa.gov/mission_pages/phoenix/spacecraft/mardi.html.
11. Mizolek, A. W., Palleschi, V. & Schechter, I. (eds) *Laser-Induced Breakdown Spectroscopy* (Cambridge Univ. Press, 2006).
12. Seel, F. *Laboratory Studies on Laser-Induced Shock Waves for LIBS Measurements on Mars*. Master's Thesis, Technische Universität Berlin (2021).
13. Chide, B. et al. Recording laser-induced sparks on Mars with the SuperCam microphone. *Spectrochim. Acta B At. Spectrosc.* **174**, 106000 (2020).
14. Murdoch, N. et al. Laser-induced breakdown spectroscopy acoustic testing of the Mars 2020 microphone. *Planet. Space Sci.* **165**, 260–271 (2019).
15. Maki, J. N. et al. The Mars 2020 engineering cameras and microphone on the Perseverance rover: a next-generation imaging system for Mars exploration. *Space Sci. Rev.* **216**, 137 (2020).

16. Savijarvi, H. A model study of the PBL structure on Mars and the Earth. *Contrib. Atmos. Phys.* **64**, 219–229 (1991).
17. Petrosyan, A. et al. The Martian atmospheric boundary layer. *Rev. Geophys.* **49**, RG3005 (2011).
18. Tillman, J. E., Landberg, L. & Larsen, S. E. The boundary layer of Mars: fluxes, stability, turbulent spectra, and growth of the mixed layer. *J. Atmos. Sci.* **51**, 1709–1727 (1994).
19. Larsen, S. E. et al. Aspects of the atmospheric surface layers on Mars and Earth. *Boundary Layer Meteorol.* **105**, 451–470 (2002).
20. Martínez, G. M. et al. The modern near-surface Martian climate: a review of in-situ meteorological data from Viking to Curiosity. *Space Sci. Rev.* **212**, 295–338 (2017).
21. Rodriguez-Manfredi, J. A. et al. The Mars Environmental Dynamics Analyzer, MEDA. A suite of environmental sensors for the Mars 2020 mission. *Space Sci. Rev.* **217**, 48 (2021).
22. Banfield, D. et al. InSight Auxiliary Payload Sensor Suite (APSS). *Space Sci. Rev.* **215**, 4 (2019).
23. Ullán, A. et al. Analysis of wind-induced dynamic pressure fluctuations during one and a half Martian years at Gale Crater. *Icarus* **288**, 78–87 (2017).
24. Chatain, A., Spiga, A., Banfield, D., Forget, F. & Murdoch, N. Seasonal variability of the daytime and nighttime atmospheric turbulence experienced by InSight on Mars. *Geophys. Res. Lett.* **48**, e2021GL095453 (2021).
25. Temel, O. et al. Large eddy simulations of the Martian convective boundary layer: Towards developing a new planetary boundary layer scheme. *Atmos. Res.* **250**, 105381 (2021).
26. Kolmogorov, A. N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Cr. Acad. Sci. URSS* **30**,

- 301–305 (1941).
27. Chen, W., Lovejoy, S. & Muller, J.-P. Mars' atmosphere: the sister planet, our statistical twin. *J. Geophys. Res. Atmos.* **121**, 11,968–11,988 (2016).
 28. Tchen, C. M. On the spectrum of energy in turbulent shear flow. *J. Res. Natl. Bur. Stand.* **50**, 51–62 (1953).
 29. Bass, H. E., Hetzer, C. H. & Raspet, R. On the speed of sound in the atmosphere as a function of altitude and frequency. *J. Geophys. Res.* **112**, D15110 (2007).
 30. Hanford, A. D. & Long, L. N. The direct simulation of acoustics on Earth, Mars, and Titan. *J. Acoust. Soc. Am.* **125**, 640–650 (2009).
 31. Zhang, X., Wang, S. & Zhu, M. Locating the inflection point of frequency-dependent velocity dispersion by acoustic relaxation to identify gas mixtures. *Meas. Sci. Technol.* **31**, 115001 (2020).
 32. Forget, F. et al. Improved general circulation models of the Martian atmosphere from the surface to above 80 km. *J. Geophys. Res.* **104**, 24155–24175 (1999).
 33. Balaram, J., Aung, M. & Golombek, M. P. The Ingenuity Helicopter on the Perseverance Rover. *Space Sci. Rev.* **217**, 56 (2021).
 34. Bass, H., Sutherland, L., Piercy, J. & Evans, L. in *Physical Acoustics: Principles and Methods* Vol. 17 (eds Mason, W. P. & Thurston, R. N.) 145–232 (Academic Press, 1984).
 35. Noble, J. M. & Auvermann, H. J. The effects of large and small scale turbulence on sound propagation in the atmosphere. Army Research Laboratory, ARL-TR-565 (1995).
 36. Berengier, M. C., Gauvreau, B., Blanc-Benon, P., Juve, D. & de Collongue, G. Outdoor sound propagation: a short review on analytical and numerical approaches. *Acta Acust. United Acust.* **89**, 980–991 (2003).

37. Ostashev, V. E. in *Sound-Flow Interactions* (eds Aurégan, Y., Pagneux, V., Pinton, J.-F. & Maurel, A.) 169–191 (Springer, 2002).
38. Garcia, R. F. et al. Finite-difference modeling of acoustic and gravity wave propagation in Mars atmosphere: application to infrasounds emitted by meteor impacts. *Space Sci. Rev.* **211**, 547–570 (2017).
39. Garcia, R. F., Lognonné, P. & Bonnin, X. Detecting atmospheric perturbations produced by Venus quakes. *Geophys. Res. Lett.* **32**, L16205 (2005).
40. Temel, O. et al. Large eddy simulations of the Martian convective boundary layer: towards developing a new planetary boundary layer scheme. *Atmos. Res.* **250**, 105381 (2021).
41. Spiga, A. et al. A study of daytime convective vortices and turbulence in the Martian planetary boundary layer based on half-a-year of InSight atmospheric measurements and large-eddy simulations. *J. Geophys. Res. Planets* **126**, e2020JE006511 (2021).
42. Colaïtis, A. et al. A thermal plume model for the Martian convective boundary layer. *J. Geophys. Res. Planets* **118**, 1468–1487 (2013).
43. Hecht, M. et al. Mars oxygen ISRU experiment (MOXIE). *Space Sci. Rev.* **217**, 9 (2021).
44. Manrique, J. A. et al. SuperCam calibration targets: design and development. *Space Sci. Rev.* **216**, 138 (2020).
45. Millour, E. et al. The Mars Climate Database (Version 5.3). Scientific Workshop: “From Mars Express to ExoMars”, 27–28 February 2018, ESAC Madrid, Spain.
46. Chiba, R., Ishikawa, Y., Hasegawa, J. & Horioka, K. Time evolution of laser-ablation plumes and induced shock waves in low-pressure gas. *Phys. Plasmas* **24**, 063520 (2017).
47. Qin, Q. & Attenborough, K. Characteristics and application of laser-generated acoustic shock waves in air. *Appl. Acoust.* **65**, 325–340

(2004).

48. Sutherland, L. C. & Bass, H. E. Atmospheric absorption in the atmosphere up to 160 km. *J. Acoust. Soc. Am.* **115**, 1012–1032 (2004).

Acknowledgements

Many people helped with this project in addition to the co-authors, including hardware and operation teams, and we are most grateful for their support. This project was supported in the USA by NASA's Mars Exploration Program and in France is conducted under the authority of CNES. The work of A. Munguira is supported by grant PID2019-109467GB-I00 funded by MCIN/AEI/10.13039/501100011033.

Author information

Author notes

1. R. C. Wiens

Present address: Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, USA

Authors and Affiliations

1. Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse 3 Paul Sabatier, CNRS, CNES, Toulouse, France

S. Maurice, A. Cousin, O. Forni, O. Gasnault, J. Lasue, P.-Y. Meslin, P. Pilleri, A. Fau, J.-A. Manrique, Y. Parot, P. Pinet, W. Rapin & M. Toplis

2. Space and Planetary Exploration Team, Los Alamos National Laboratory, Los Alamos, NM, USA

B. Chide, R. C. Wiens, N. L. Lanza, A. Ollila, R. C. Wiens, R. Beal, S. M. Clegg, D. Delapp, C. Legett IV, T. Nelson, R. T. Newell, A. Reyes-Newell, S. Robinson & D. Venhaus

3. Institut Supérieur de l'Aéronautique et de l'Espace (ISAE-SUPAERO),
Université de Toulouse, Toulouse, France

N. Murdoch, D. Mimoun, A. Stott, M. Bassas-Portus, A. Cadu & A.
Sournac

4. Space Exploration Sector, Johns Hopkins Applied Physics Laboratory,
Laurel, MD, USA

R. D. Lorenz & J. R. Johnson

5. Institut de Mécanique des Fluides, Université de Toulouse 3 Paul
Sabatier, INP, CNRS, Toulouse, France

X. Jacob

6. Laboratoire d'Etudes Spatiales et d'Instrumentation en Astrophysique,
Observatoire de Paris, CNRS, Sorbonne Université, Université Paris
Diderot, Meudon, France

T. Bertrand, T. Fouchet, P. Bernardi, A. Dorresoundiram, L.
Mandon & C. Royer

7. Laboratoire Atmosphères, Milieux, Observations Spatiales, CNRS,
Université Saint-Quentin-en-Yvelines, Sorbonne Université,
Guyancourt, France

F. Montmessin, E. W. Knutson & G. Lacombe

8. Universidad de Málaga, Málaga, Spain

C. Alvarez-Llamas, J. Laserna & J. Moros

9. Department of Chemistry and Biochemistry, University of South
Carolina, Columbia, SC, USA

S. M. Angel

10. Jet Propulsion Laboratory, California Institute of Technology,
Pasadena, CA, USA

M. Aung, J. Balaram, H. Grip, J. Maki, M. de la Torre Juárez, T.
Tzanatos, K. M. Stack, K. Farley, K. Williford, R. Francis, I.
Gontijo, S. Madsen, V. Shridar & P. Willis

11. Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie,
CNRS, Sorbonne Université, MNHN, Paris, France

O. Beyssac, K. Benzerara, S. Bernard & V. Sautter

12. Heliospace Corporation, Berkeley, CA, USA

G. Delory

13. Haystack Observatory, Massachusetts Institute of Technology,
Westford, MA, USA

M. Hecht & J. McClean

14. Department of Aeronautics and Astronautics, Massachusetts Institute of
Technology, Cambridge, MA, USA

J. Hoffman

15. Laboratoire de Planétologie et Géosciences, CNRS, Nantes Université,
Université Angers, Nantes, France

S. Le Mouélic, S. Le Mouélic & N. Mangold

16. Escuela de Ingeniería de Bilbao, Universidad del País Vasco
UPV/EHU, Bilbao, Spain

A. Munguira

17. Aeolis Corporation, Sierra Madre, CA, USA

C. E. Newman

18. Centro de Astrobiología (INTA-CSIC), Madrid, Spain
J. A. Rodríguez Manfredi
19. Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institute of Optical Sensor Systems, Berlin, Germany
S. Schröder & D. Vogt
20. Blue Marble Space Institute of Science, Seattle, WA, USA
K. Williford
21. University of Hawai‘i at Mānoa, Mānoa, HI, USA
T. Acosta-Maeda, E. Kelly & S. Sharma
22. U.S. Geological Survey, Flagstaff, AZ, USA
R. B. Anderson & T. Gabriel
23. University of Winnipeg, Winnipeg, Canada
D. M. Apmlin, E. Cloutis, S. Connell & N. Turenne
24. University of the Basque Country UPV/EHU, Leioa, Bilbao, Spain
G. Arana, K. Castro, J. M. Madariaga & I. Torre-Fdez
25. Institut de Planétologie et Astrophysique de Grenoble, CNRS, Université Grenoble Alpes, Grenoble, France
P. Beck & E. Lewin
26. Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA
T. Bosak

27. Centre Lasers Intenses et Applications, CNRS, CEA, Université de Bordeaux, Bordeaux, France
B. Bousquet & E. Clavé
28. Plancius Research, Severna Park, MD, USA
A. Brown
29. Laboratoire d'Astrophysique de Bordeaux, CNRS, Université de Bordeaux, Bordeaux, France
P. Caïs & B. Quertier
30. Centre National d'Études Spatiales, Toulouse, France
A. Debus, C. Donny, E. Lorigny, M. Mann, F. Meunier, V. Mousset, G. Pont & L. Rochas
31. Université de Lyon, UCBL, ENSL, UJM, CNRS, LGL-TPE, Villeurbanne, France
E. Dehouck, G. Dromart, G. Montagnac & C. Quantin-Nataf
32. Groupe d'Instrumentation Scientifique, Observatoire Midi-Pyrénées, Toulouse, France
B. Dubois
33. GeoRessources, CNRS, Université de Lorraine, Nancy, France
C. Fabre
34. California Institute of Technology, Pasadena, CA, USA
W. Fischer & H. Kalucha
35. University of Copenhagen, Copenhagen, Denmark
J. Frydenvang & M. Madsen

36. McGill University, Montreal, Canada
E. Gibbons & R. Leveille
37. University of Valladolid, Valladolid, Spain
G. Lopez-Reyes, J.-A. Manrique, F. Rull & M. Veneranda
38. Agencia Estatal Consejo Superior de Investigaciones Científicas,
Madrid, Spain
J. Martinez-Frias
39. Department of Geological Sciences, University of Colorado Boulder,
Boulder, CO, USA
L. E. Mayhew
40. University of Maryland, College Park, MD, USA
T. McConnochie
41. State University of New York, Stony Brook, NY, USA
S. M. McLennan
42. Department of Physics and Applied Physics, Kennedy College of
Sciences, University of Massachusetts Lowell, Lowell, MA, USA
N. Melikechi
43. Institut d'Astrophysique Spatiale, CNRS, Université Paris-Saclay,
Orsay, France
C. Pilorget & F. Poulet
44. Institut Universitaire de France, Paris, France
C. Pilorget

45. University of Nevada, Las Vegas, Las Vegas, NV, USA

A. Udry

Consortia

the SuperCam team

- R. C. Wiens
- , S. Maurice
- , T. Acosta-Maeda
- , C. Alvarez-Llamas
- , R. B. Anderson
- , S. M. Angel
- , D. M. Applin
- , G. Arana
- , M. Bassas-Portus
- , R. Beal
- , P. Beck
- , K. Benzerara
- , S. Bernard
- , P. Bernardi
- , T. Bertrand
- , O. Beyssac
- , T. Bosak
- , B. Bousquet
- , A. Brown
- , A. Cadu
- , P. Caïs
- , K. Castro
- , B. Chide
- , E. Clavé
- , S. M. Clegg
- , E. Cloutis
- , S. Connell
- , A. Cousin
- , A. Debus

- , E. Dehouck
- , D. Delapp
- , C. Donny
- , A. Dorresoundiram
- , G. Dromart
- , B. Dubois
- , C. Fabre
- , A. Fau
- , W. Fischer
- , O. Forni
- , T. Fouchet
- , R. Francis
- , J. Frydenvang
- , T. Gabriel
- , O. Gasnault
- , E. Gibbons
- , I. Gontijo
- , X. Jacob
- , J. R. Johnson
- , H. Kalucha
- , E. Kelly
- , E. W. Knutsen
- , G. Lacombe
- , N. L. Lanza
- , J. Laserna
- , J. Lasue
- , S. Le Mouélic
- , C. Leggett IV
- , R. Leveille
- , E. Lewin
- , G. Lopez-Reyes
- , R. D. Lorenz
- , E. Lorigny
- , J. M. Madariaga
- , M. Madsen
- , S. Madsen
- , L. Mandon

- , N. Mangold
- , M. Mann
- , J.-A. Manrique
- , J. Martinez-Frias
- , L. E. Mayhew
- , T. McConnochie
- , S. M. McLennan
- , N. Melikechi
- , P.-Y. Meslin
- , F. Meunier
- , D. Mimoun
- , G. Montagnac
- , F. Montmessin
- , J. Moros
- , V. Mousset
- , N. Murdoch
- , T. Nelson
- , R. T. Newell
- , A. Ollila
- , Y. Parot
- , P. Pilleri
- , C. Pilorget
- , P. Pinet
- , G. Pont
- , F. Poulet
- , C. Quantin-Nataf
- , B. Quertier
- , W. Rapin
- , A. Reyes-Newell
- , S. Robinson
- , L. Rochas
- , C. Royer
- , F. Rull
- , V. Sautter
- , S. Schröder
- , S. Sharma
- , V. Shridar

- , A. Sournac
- , A. Stott
- , M. Toplis
- , I. Torre-Fdez
- , N. Turenne
- , T. Tzanetos
- , A. Udry
- , M. Veneranda
- , D. Venhaus
- , D. Vogt
- & P. Willis

Contributions

S.M. and B.C. equally led the writing of the manuscript. S.M., B.C., N.M., R.D.L., A.S., X.J., T.B. and F.M. performed data processing and interpreted the data. D.M. is the lead of SuperCam’s microphone, R.C.W. the lead of the SuperCam investigation, and N.L.L. and B.C. the leads of the SuperCam Acoustics Working Group. Other investigations provided data and support to this study: J.M. is the lead of EDL’s microphone; M.A., J.B., H.G. and T.T. are leading the Ingenuity project; M.H., J.H. and J.McClean are leading the MOXIE investigation; J.A.R.M., M.dIT.J. and C.E.N. are leading the MEDA investigation and mission atmospheric working group. All other co-authors provided helpful comments and inputs to the manuscript. The SuperCam team built the instrument, helps daily to operate it and to process and interpret the data.

Corresponding authors

Correspondence to [S. Maurice](#) or [B. Chide](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Andi Petculescu, Peter Read and Roger Waxler for their contribution to the peer review of this work.

Additional information

Extended data

is available for this paper at <https://doi.org/10.1038/s41586-022-04679-0>.

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 Artificial sounds recorded by Perseverance.](#)

For each panel: time is along the *x* axis; on top time series data in mPa (see hereafter for the range when it exists); at the bottom PSD in Pa² Hz⁻¹ (a colour scale is shown next to panel **a**; ranges are different for each panel but not indicated). **a**, Unfiltered spectrogram of a segment (240–270 s) of the 16-min EDL microphone recording during the rover drive (Sol 16). Time series data are not calibrated in pressure. **b**, Spectrogram of Ingenuity's fourth flight (Sol 69). Time series data range from -10 mPa to +10 mPa. **c**, Unfiltered spectrogram of MOXIE oxygen production (Sol 81). Time series data range from -2.5 mPa to +2.5 mPa. **d**, Series of ten LIBS shots on target Hedgehog (Sol 37), labelled 1 to 10. Time series data range from -500 mPa to 500 mPa.

[Extended Data Fig. 2 Recording of laser-induced shock wave.](#)

Target Hedgehog (point #1, shot #1). The inset focuses on the time when the acoustic wave is recorded.

Extended Data Fig. 3 Recording of the BPF of Ingenuity's fourth flight.

a, PSD for the entire recording (purple). The rover's thermal pumps are shown at 195 Hz and 198.75 Hz. PSD for Sol 38a is the reference of a quiet day. **b**, Top, variations of the received frequency along the trajectory of the flight (diamond symbol). During the first 60 s, the recording is perturbed by high winds. In red, the Doppler effect is modelled with $f = 84.44$ Hz at the source and $c = 237.7 \text{ m s}^{-1}$. Bottom, range between the rover and Ingenuity (solid line, left y axis) and rotorcraft speed (dotted line, right y axis).

Extended Data Table 1 List of audio recordings used in this study

Extended Data Table 2 Origin and characteristics of the dataset used in this study

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[Reprints and Permissions](#)

About this article

Cite this article

Maurice, S., Chide, B., Murdoch, N. *et al.* In situ recording of Mars soundscape. *Nature* **605**, 653–658 (2022). <https://doi.org/10.1038/s41586-022-04679-0>

- Received: 07 December 2021
- Accepted: 23 March 2022
- Published: 01 April 2022
- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04679-0>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

[**Snippets of the Martian soundscape**](#)

Nature Research Briefing 04 Apr 2022

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04679-0>

[Download PDF](#)

- Article
- Open Access
- [Published: 25 May 2022](#)

Qubit teleportation between non-neighbouring nodes in a quantum network

- [S. L. N. Hermans](#) [ORCID: orcid.org/0000-0002-1060-5199^{1 na1}](#),
- [M. Pompili](#) [ORCID: orcid.org/0000-0002-8360-1957^{1 na1}](#),
- [H. K. C. Beukers](#) [ORCID: orcid.org/0000-0001-9934-1099¹](#),
- [S. Baier](#) [ORCID: orcid.org/0000-0002-2840-5590^{1 nAff2}](#),
- [J. Borregaard¹](#) &
- [R. Hanson](#) [ORCID: orcid.org/0000-0001-8938-2137¹](#)

[Nature](#) volume **605**, pages 663–668 (2022)

- 4685 Accesses
- 288 Altmetric
- [Metrics details](#)

Subjects

- [Quantum information](#)
- [Quantum optics](#)

Abstract

Future quantum internet applications will derive their power from the ability to share quantum information across the network^{1,2}. Quantum teleportation allows for the reliable transfer of quantum information between distant nodes, even in the presence of highly lossy network connections³. Although many experimental demonstrations have been performed on different quantum network platforms^{4,5,6,7,8,9,10}, moving beyond directly connected nodes has, so far, been hindered by the demanding requirements on the pre-shared remote entanglement, joint qubit readout and coherence times. Here we realize quantum teleportation between remote, non-neighbouring nodes in a quantum network. The network uses three optically connected nodes based on solid-state spin qubits. The teleporter is prepared by establishing remote entanglement on the two links, followed by entanglement swapping on the middle node and storage in a memory qubit. We demonstrate that, once successful preparation of the teleporter is heralded, arbitrary qubit states can be teleported with fidelity above the classical bound, even with unit efficiency. These results are enabled by key innovations in the qubit readout procedure, active memory qubit protection during entanglement generation and tailored heralding that reduces remote entanglement infidelities. Our work demonstrates a prime building block for future quantum networks and opens the door to exploring teleportation-based multi-node protocols and applications^{2,11,12,13}.

Main

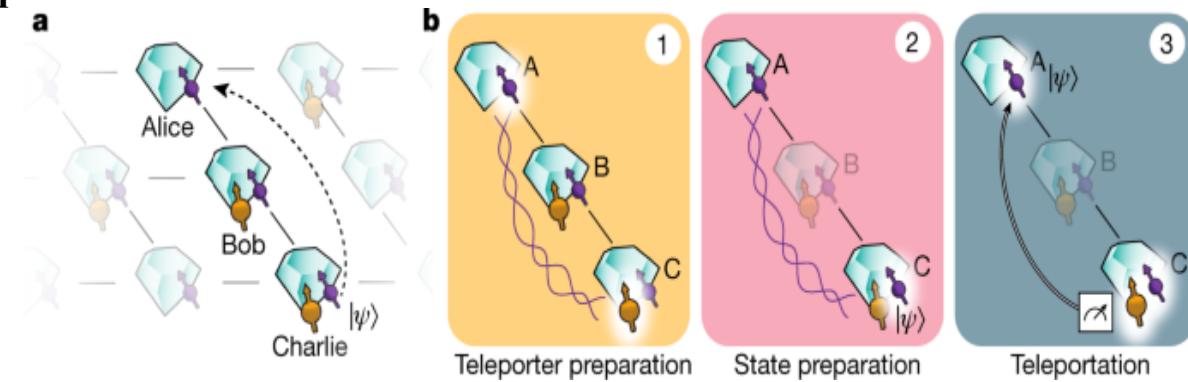
Quantum teleportation is the central routine for reliably sending qubits across lossy network links³, as well as a key primitive of quantum network protocols and applications^{2,11,12}. Using a teleporter in the form of a pre-shared entangled state, the quantum information is transferred by performing a joint Bell-state measurement (BSM) on the sender's part of the entangled state and the qubit state to be teleported. The state is recovered on the receiving node by a gate operation conditioned on the BSM outcome³. Because the quantum information is not transmitted by a physical carrier, the protocol is insensitive to loss in the connecting photonic channels and on intermediate nodes. A deterministic BSM combined with real-time feed-forward enables unconditional teleportation, in which state transfer is achieved each time a qubit state is inserted into the teleporter.

Pioneering explorations of quantum teleportation protocols were performed using photonic states^{4,5,6}. Following the development of quantum network nodes with stationary qubits, remote qubit teleportation was realized between trapped ions⁷, trapped atoms^{8,10}, diamond nitrogen-vacancy (NV) centres⁹ and memory nodes based on atomic ensembles¹⁴.

Although future quantum network applications will widely use teleportation between non-connected nodes in the network, the demanding set of requirements on the pre-shared entanglement, the BSM and the coherence times for enabling real-time feed-forward has, so far, prevented the realization of teleportation beyond directly connected stationary network nodes.

Here we overcome these challenges by a set of key innovations and achieve qubit teleportation between non-neighbouring network nodes (see Fig. 1a). Our quantum network consists of three nodes in a line configuration, Alice, Bob and Charlie. Each node contains a NV centre in diamond. Using the NV electronic spin as the communication qubit, we are able to generate remote entanglement between each pair of neighbouring nodes. In addition, Bob and Charlie each use a nearby ¹³C nuclear spin as a memory qubit. The steps of the teleportation protocol are shown in Fig. 1b. To prepare the teleporter, we use an entanglement swapping protocol mediated by Bob, similar to a quantum repeater protocol¹⁵, to establish entanglement between Alice and Charlie. Once successful preparation of the teleporter is heralded, the input qubit state is prepared on Charlie and finally teleported to Alice.

Fig. 1: Teleporting a qubit between non-neighbouring nodes of a quantum network.

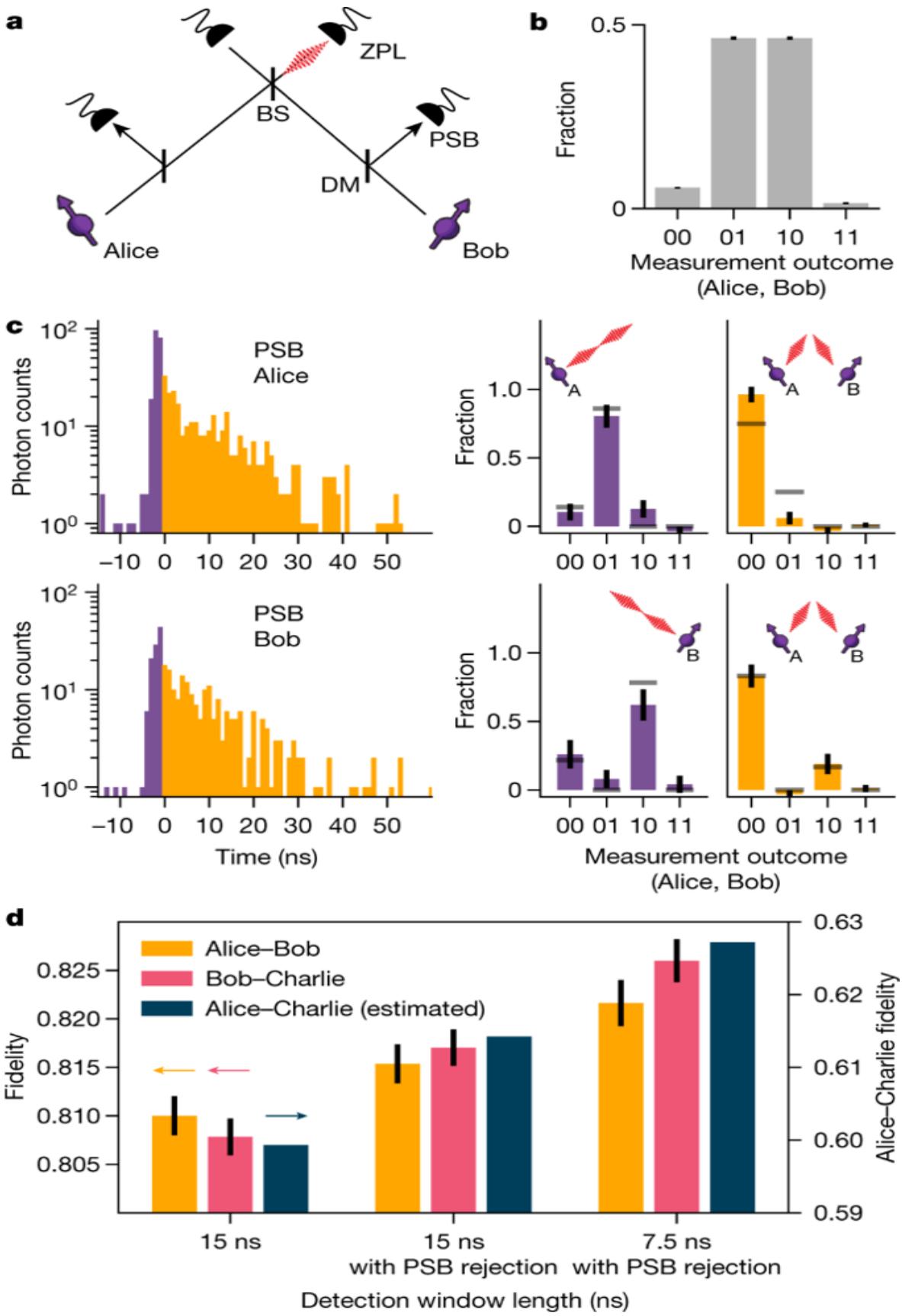


a, Three network nodes, Alice (A), Bob (B) and Charlie (C), are connected by means of optical fibre links (lines) in a line configuration. Each setup has a communication qubit (purple) that enables entanglement generation with its neighbouring node. Furthermore, Bob and Charlie contain a memory qubit (yellow). **b**, The steps of the teleportation protocol. (1) We prepare the teleporter by establishing entanglement between Alice and Charlie using an entanglement swapping protocol on Bob, followed by swapping the state at Charlie to the memory qubit. (2) The qubit state to be teleported is prepared on the communication qubit on Charlie. (3) A BSM is performed on Charlie's qubits and the outcome is communicated to Alice over a classical channel. Dependent on this outcome, Alice applies a quantum gate to obtain the teleported qubit state.

Entanglement fidelity of the network links

A key parameter for quantum teleportation is the fidelity of the pre-shared entangled state between Alice and Charlie. As we generate this state by entanglement swapping, its fidelity can be increased by mitigating errors on the individual links. Our network generates entanglement between neighbouring nodes using a single-photon protocol^{16,17} in an optical-phase-stabilized architecture¹⁸. The building block of this protocol is a qubit–photon entangled state created at each node. To generate this entangled state, we initialize the communication qubit in a superposition state $\langle|\psi\rangle = \sqrt{\alpha}|0\rangle + \sqrt{1-\alpha}|1\rangle$ and apply a state-selective optical pulse that transfers the population from $|0\rangle$ to an optically excited state. Following spontaneous emission, the qubit state is entangled with the photon number (0 or 1 photon). We perform this protocol on both nodes and interfere the resonant photonic states on a beam splitter (Fig. 2a). Detection of a single photon in one of the output ports ideally heralds the generation of an entangled state $\langle|\psi\rangle = (|01\rangle \pm |10\rangle)/\sqrt{2}$ in which the \pm phase is set by the detector that clicked. Figure 2b shows the joint outcomes of qubit measurements in the computational basis after entanglement is heralded, showing the expected correlations.

Fig. 2: High-fidelity entangled network links.



a, Simplified schematic of the optical link used for generating entanglement between neighbouring nodes. Photons emitted by the communication qubits are filtered by a dichroic mirror (DM) to separate the resonant (zero-phonon line, ZPL) photons (3% of emission) from the off-resonant (phonon-side band, PSB) photons (97% of emission). The resonant photons are sent to the beam splitter (BS); detection of a single photon at one of the ZPL detectors heralds successful generation of an entangled state between the two nodes. **b**, Measured correlations of the communication qubits in the computational basis, conditioned on a heralding event on the ZPL detectors. **c**, Left, histograms of the PSB photon detection times on Alice (top) or Bob (bottom), conditioned on a simultaneous ZPL detection in the same entanglement generation attempt. Grey lines show expected correlations on the basis of a quantum-optical model (see [Supplementary Information](#)). The correlations measured in the other measurement bases can be found in Extended Data Fig. [1](#). **d**, Measured fidelity of the network links, without PSB rejection (left), with PSB rejection (middle) and with PSB rejection plus shortened detection window (right). The dark blue bars indicate the corresponding expected fidelity on Alice–Charlie after entanglement swapping for each case (see Methods). All error bars represent one standard deviation.

The infidelity of the generated state has three main contributions: double $\langle|0\rangle\langle|0|\rangle$ state occupancy, double optical excitation and finite distinguishability of the photons^{[18,19](#)}. In the case of double $\langle|0\rangle\langle|0|\rangle$ state occupancy (which occurs with probability α), both communication qubits are in the $\langle|0\rangle\langle|0|\rangle$ state and have emitted a photon. Detection of one of these photons leads to false heralding of an entangled state. The second effect, double excitation, is due to the finite length of the optical pulse compared with the optical lifetime of the emitter. There is a finite chance that the communication qubit emits a photon during this pulse, is subsequently re-excited and then emits another photon, resulting in the qubit state being entangled with two photons. Detection or loss of the first photon destroys the coherence of the qubit–photon entangled state and detection of the second photon can then falsely herald the generation of an entangled state.

Crucially, false heralding events caused by double $\langle|0\rangle\langle|0|\rangle$ state occupancy and double excitation are both accompanied by an extra emitted

photon. Therefore, detection of this extra photon allows for unambiguous identification of such events and thus for real-time rejection of false heralding signals. We implement this rejection scheme by monitoring the off-resonant phonon-side band (PSB) detection path on both setups during and after the optical excitation (see Fig. 2a).

To investigate the effect of this scheme, we generate entanglement on the individual links and extract the entanglement heralding events for which the PSB monitoring flagged the presence of an extra photon. For these events, we analyse the corresponding qubit measurements in the computational basis (Fig. 2c).

We identify two separate regimes: one during the optical pulse (purple) and one after the optical pulse (yellow). When a photon is detected on Alice's (Bob's) PSB detector during the optical pulse, we see that the outcome 01 (10) is most probable (purple data in Fig. 2c), showing that only one setup was in the $|0\rangle$ state and thus that both detected photons originated from Alice (Bob). The detection of PSB photons during the optical pulse thus primarily flags double excitation errors. By contrast, when a photon is detected after the optical pulse in either Alice's or Bob's PSB detector, the outcome 00 is most probable (yellow data in Fig. 2c), indicating that both setups were in the $|0\rangle$ state and emitted a photon. PSB photon detection after the optical pulse thus flags the double $|0\rangle$ state occupancy error. We find similar results to Fig. 2c for the entangled states generated on the Bob–Charlie link; see Extended Data Fig. 2. The improvement in fidelity from rejecting these false heralding events in our experiment is set by the combined probability of occurrence ($\approx 9\%$; see [Supplementary Information](#)) multiplied by the probability to flag them (given here by the total PSB photon detection efficiency of $\approx 10\%$).

The third main source of infidelity, the finite distinguishability, can arise from frequency detunings between the emitted photons²⁰. Whereas most of these detunings are eliminated up front by the charge-resonance (CR) check before the start of the protocol (see [Supplementary Information](#)), the communication qubits may still be subject to a small amount of spectral diffusion. In our single-photon protocol, this leads to dephasing that is stronger for photons that are detected later relative to the optical pulse. By shortening our detection window, we can increase the fidelity of the

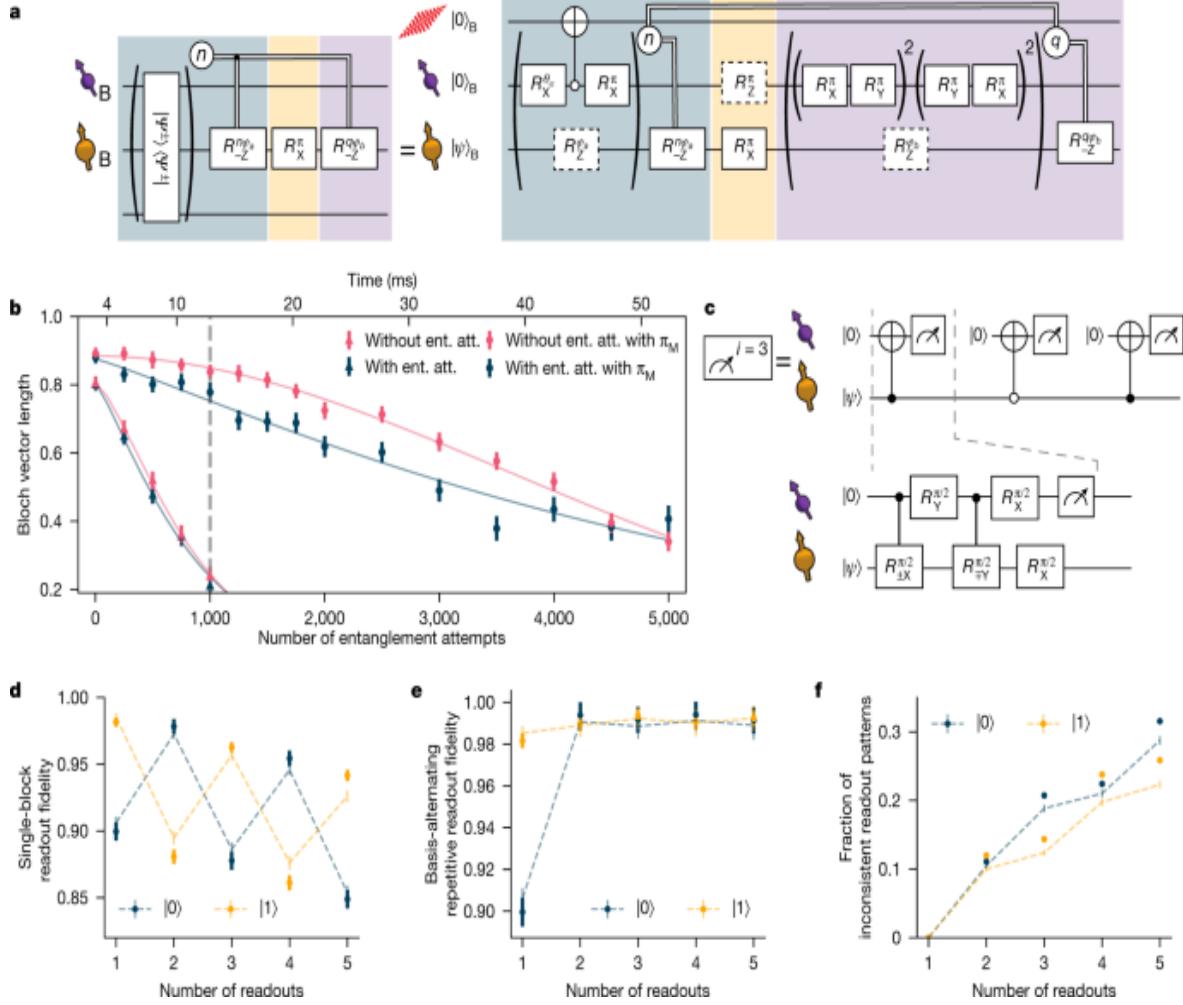
entangled state at the expense of a lower entangling rate. For the experiments below (unless mentioned otherwise), we use a detection window length of 15 ns. Figure 2d summarizes the measured improvements on the individual links and the estimated effect on the Alice–Charlie entangled state fidelity. The increase of $\approx 3\%$ is instrumental in pushing the teleportation fidelity above the classical bound.

Memory qubit coherence

In the preparation of the teleporter, we reliably preserve the Alice–Bob entangled link on the memory qubit, by aborting the sequence and starting over when the Bob–Charlie entangled state is not heralded within a fixed number of attempts, the timeout.

The ^{13}C memory qubits can be controlled with high fidelity by means of the communication qubit, although they can be efficiently decoupled when no interaction is desired. Recent work showed that, in a magnetic field of 189 mT, entanglement generation attempts with the communication qubit do not limit the memory dephasing time $\langle T_2^{\text{ast}} \rangle$ (ref. 18), opening the door to substantially extending the memory preservation time with active coherence protection from the spin bath²¹. We realize this protection by integrating a decoupling π -pulse on the memory qubit into the experimental sequence that follows a heralding event, while ensuring that all phases that are picked up owing to the probabilistic nature of the remote entangling process are compensated in real time (Fig. 3a).

Fig. 3: Memory qubit coherence and readout.



a, Gate sequence on Bob for entanglement generation with the communication qubit while preserving states stored on the memory qubit. Entanglement generation attempts are repeated until success or a predetermined timeout. On success in the n th attempt, a phase feed-forward is applied to maintain the correct reference frame of the memory qubit¹⁸, followed by a decoupling pulse on the memory qubit. The decoupling π_M pulse causes a Z rotation on the communication qubit. Afterwards, we rephase the memory qubit for the same amount of time as it took to herald entanglement (by applying q blocks of XY8 decoupling sequences on the communication qubit, in which q depends on the number of entanglement attempts needed n) and we end with another phase feed-forward on the memory qubit, to compensate for any phase picked up during this decoupling. **b**, Bloch vector length of a superposition state stored on the memory qubit for different number of entanglement attempts or a time-averaged value.

equivalent wait element. In the case of no decoupling (no π_M) on the memory qubit, the gates in the yellow shaded box in **a** are left out. The grey dashed line indicates the chosen timeout of 1,000 entanglement attempts. **c**, Gate sequence for the basis-alternating repetitive readout of the memory qubit. **d**, Readout fidelity for each readout repetition, for states $|0\rangle$ and $|1\rangle$. **e**, Readout fidelity of the basis-alternating repetitive readout scheme for different number of readout repetitions. **f**, Fraction of inconsistent readout patterns for different number of readout repetitions. In **d–f**, the dashed lines show a numerical model using measured parameters. All error bars represent one standard deviation.

In Fig. 3**b**, we check the performance of this sequence by storing a superposition state on the memory qubit and measuring the Bloch vector length. We observe that, without the decoupling pulse, the decay of the Bloch vector length is not altered by the entanglement attempts, in line with previous findings¹⁸. By contrast, when we apply the decoupling pulse, the decay is slowed down by more than a factor of 6, yielding a $N_{1/e}$ decay constant of $\approx 5,300$ entanglement attempts, the highest number reported so far for diamond devices. The difference in the shape of the decay indicates that intrinsic decoherence is no longer the only limiting factor. The improved memory coherence enables us to use a timeout of 1,000 entangling attempts, more than double that of ref. ¹⁸, which doubles the entanglement swapping rate.

Memory qubit readout

High-fidelity memory qubit readout is required both in the preparation of the teleporter (at Bob) and during the teleportation protocol itself (at Charlie). The memory qubit is read out by mapping its state onto the communication qubit using quantum logic followed by single-shot readout of the communication qubit using state-dependent optical excitation and detection²². Owing to limited photon collection efficiency ($\approx 10\%$) and finite cyclicity of the optical transition ($\approx 99\%$), the communication qubit readout fidelity is different for $|0\rangle$ and $|1\rangle$ and the probability that the correct state was assigned is much larger if one or more photons were detected (assigned outcome 0) than if no photons were detected (assigned

outcome 1)²³. In previous work, we circumvented this issue by conditioning on obtaining the outcome 0 (ref. ¹⁸). However, this approach scales unfavourably, as it forces the protocol to prematurely abort with probability >50% at each memory qubit readout.

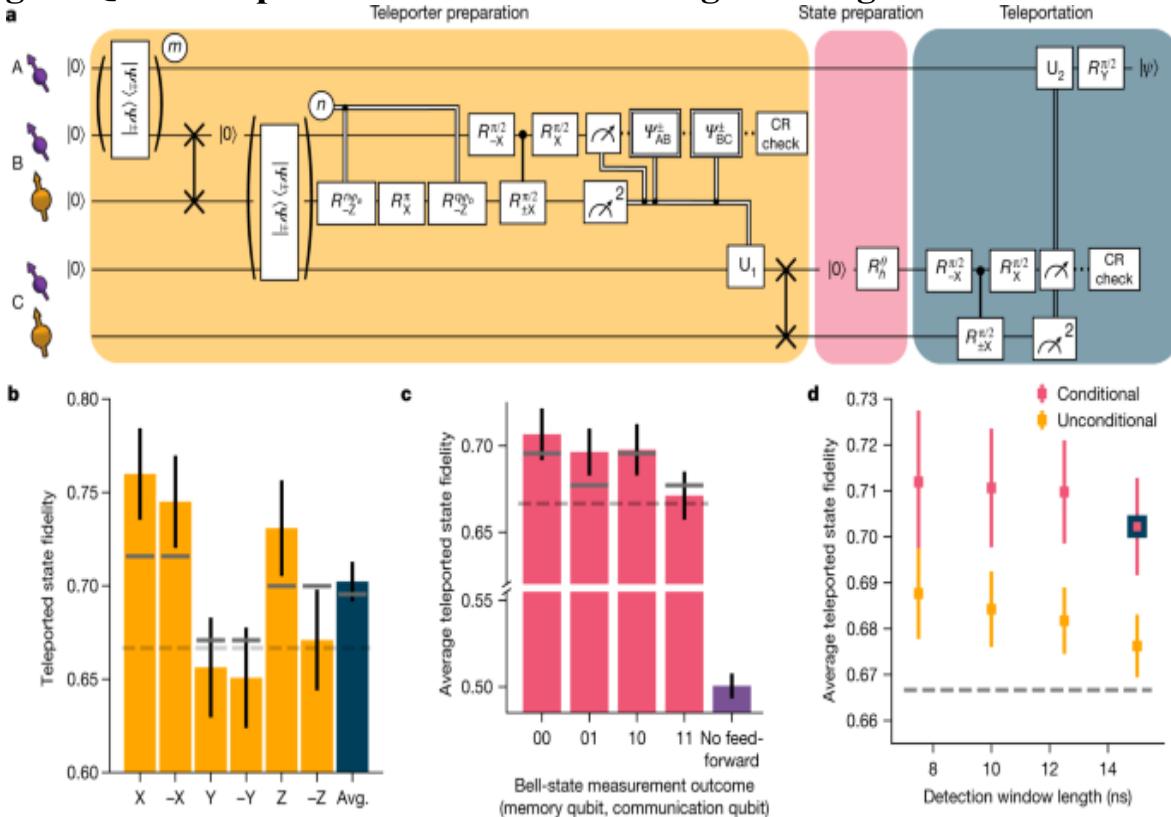
We resolve this challenge by introducing a basis-alternating repetitive readout for the memory qubit (see Fig. ^{3c}). The key point of this readout strategy is, in contrast to earlier work²⁴, to alternatingly map the computational basis states of the memory qubit to the communication qubit state $\langle 0 | \rangle$ Figure ^{3d} shows the readout fidelities of the n th readout repetition for the two initial states for the memory qubit on Bob (for Charlie, see Extended Data Fig. ³). We clearly observe the expected alternating pattern owing to the asymmetry of the communication qubit readout fidelities. Notably, the readout fidelity decays only by $\approx 1\%$ per readout, showing that the readout is mostly non-demolition and several readouts are possible without losing the state.

Next, we assign the state using the first readout and continue the sequence only when the consecutive readouts are consistent with the first readout. The subsequent readouts therefore add confidence to the assignment in the case of consistent outcomes, whereas cases of inconsistent outcomes (which have a higher chance of indicating an incorrect assignment) are filtered out. In Fig. ^{3e}, we plot the readout fidelity resulting from this strategy for up to five readouts, with the corresponding rejected fraction due to inconsistent outcomes plotted in Fig. ^{3f}. We observe that using two readouts already eliminates most of the asymmetry, reducing the average infidelity from $\approx 6\%$ to below 1%. At this point, the remaining observed infidelity mainly results from cases in which the memory qubit was flipped during the first readout block because of imperfect memory qubit gates. For the experiments reported below (unless mentioned otherwise), we use two readout repetitions to benefit from a high average readout fidelity (Bob: 99.2(4)%; Charlie: 98.1(4)%) and a high probability to continue the sequence (Bob and Charlie: $\approx 88\%$).

Teleporting qubit states from Charlie to Alice

With all innovations described above implemented, we perform the protocol as shown in Fig. 4a. First, we generate entanglement between Alice and Bob and store Bob's part of the entangled state on the memory qubit using a compiled SWAP operation. Second, we generate entanglement between Bob and Charlie, while preserving the first entangled state on the memory qubit with the pulse sequence as described in Fig. 3a. Next, we perform a BSM on Bob followed by a CR check. We continue the sequence if the communication qubit readout yields outcome 0, the memory qubit readout gives a consistent outcome pattern and the CR check is passed. At Charlie, we perform a quantum gate that depends on the outcome of the BSM and on which detectors clicked during the two-node entanglement generation. Next, we swap the entangled state to the memory qubit. At this point, the teleporter is ready and Alice and Charlie share an entangled state with an estimated fidelity of 0.61.

Fig. 4: Qubit teleportation between non-neighbouring network nodes.



a. Circuit diagram of the teleportation protocol using notation defined in Fig. 3. m (n) is the number of attempts needed to herald entanglement for the AB (BC) entangled link. See the [Supplementary Information](#) for the full circuit

diagram. **b**, Teleported state fidelities for the six cardinal states and their average (Avg.). The grey lines show the expected fidelities from simulations. The dashed lines in **b–d** represent the classical bound of 2/3. **c**, Average teleported state fidelity for the different outcomes of the BSM on Charlie. The right-most bar shows the resulting fidelity when no feed-forward operation on Alice would be applied. The numerical values of the bar plots shown in **b** and **c** can be found in Extended Data Tables [1](#) and [2](#). **d**, Average state fidelity for a conditional and an unconditional teleportation, for different detection window lengths of the two-node entanglement generation processes. The blue-bordered data point is the same point as shown in **b**. All error bars represent one standard deviation.

Subsequently, we generate the qubit state to be teleported, $\langle|\psi\rangle$ on Charlie’s communication qubit and run the teleportation protocol. First, a BSM is performed on the communication and memory qubits at Charlie. With the exception of unconditional teleportation (discussed below), we only continue the sequence when we obtain a 0 outcome on the communication qubit, when we have a consistent readout pattern on the memory qubit and when Charlie passes the CR check. The outcomes of the BSM are sent to Alice and, by applying the corresponding gate operation, we obtain $\langle|\psi\rangle$ on Alice’s side.

We teleport the six cardinal states $\{(\pm \text{rm}\{X\}, \pm \text{rm}\{Y\}, \pm \text{rm}\{Z\})\}$ which form an unbiased set^{[25](#)}, and measure the fidelity of the teleported states to the ideally prepared state (Fig. [4b](#)). We find an average teleported state fidelity of $F = 0.702(11)$ at an experimental rate of 1/(117 s). This value exceeds the classical bound of 2/3 by more than three standard deviations, thereby proving the quantum nature of the protocol. We note that this value provides a lower bound to the true teleportation fidelity, as the measured fidelity is decreased by errors in the preparation of the qubit states at Charlie (estimated to be 0.5%; see [Supplementary Information](#)).

The differences in fidelity between the teleported states arise from an interplay of errors in different parts of the protocol that either affect all three axes (depolarizing errors) or affect only two axes (dephasing errors). These differences are qualitatively reproduced by our model (grey bars in Fig. [4b](#)). In Fig. [4c](#), we plot the teleportation fidelity for each possible outcome of the BSM. Owing to the basis-alternating repetitive readout, the dependence on

the second bit (from the memory qubit readout) is small, whereas for the first bit (communication qubit readout), the best teleported state fidelity is achieved for outcome 0, due to the asymmetric readout fidelities. We also analyse the case in which no feed-forward is applied at Alice (see [Methods](#)); as expected, the average state fidelity reduces to a value consistent with a fully mixed state (fidelity $F = 0.501(7)$), emphasizing the critical role of the feed-forward in the teleportation protocol.

Finally, we demonstrate that the network can achieve unconditional teleportation between Alice and Charlie by using the BSM in a deterministic fashion. To this end, we revise the protocol at Charlie to accept both communication qubit outcomes, use all memory qubit readout patterns, including the inconsistent ones, and disregard the outcome of the CR check after the BSM. Using this fully deterministic BSM lowers the average teleportation fidelity by a few percent (Fig. [4d](#)). At the same time, shortening the detection windows of the two-node entanglement generation is expected to yield an improvement in the fidelity, as discussed above. We find that, indeed, the average unconditional teleportation fidelity increases with shorter window lengths, reaching $F = 0.688(10)$ for a length of 7.5 ns and a rate of 1/(100 s); see Extended Data Fig. [4](#). The current quantum network is thus able to perform teleportation beyond the classical bound, even under the strict condition that every state inserted into the teleporter be transferred.

Outlook

In this work, we have realized unconditional qubit teleportation between non-neighbouring nodes in a quantum network. The innovations introduced here on memory qubit readout and protection during entanglement generation, as well as the real-time rejection of false heralding signals, will be instrumental in exploring more complex protocols^{[2,11,12,13,26](#)}. Also, these methods can be readily transferred to other platforms, such as the group IV colour centres in diamond, the vacancy-related qubits in SiC and single rare-earth ions in solids^{[27,28,29,30,31,32,33](#)}.

The development of an improved optical interface for the communication qubit^{[34](#)} will increase both the teleportation protocol rate and fidelity.

Because of the improved memory qubit performance reported here, the network already operates close to the threshold at which nodes can reliably deliver a remote entangled state while preserving previously stored quantum states in their memory qubits. With further improvements, for instance, by integrating multi-pulse memory decoupling sequences²¹ into the entanglement generation, demonstration of deterministic qubit teleportation (with no pre-shared entangled state) may come within reach, which opens the door to exploring applications that call the teleportation routine several times. In addition, future work will focus on further improving the phase stabilization and extending the current schemes for use in deployed fibre³⁵.

Finally, by implementing a recently proposed link layer protocol³⁶, qubit teleportation and applications making use of the teleportation primitive may be executed and tested on the network through platform-independent control software, an important prerequisite for a large-scale future network.

Methods

Experimental setup

The basics of the experimental setup are described in ref. ¹⁸. In the current experiment, Charlie has access to a carbon-13 nuclear spin that acts as a memory qubit. The parameters used for the memory qubits of Bob and Charlie can be found in Extended Data Table ³. Furthermore, we have set up a classical communication channel between Charlie and Alice, such that Charlie can directly send the results of the BSM to Alice.

Temporal selection of heralding photons

To eliminate any reflected excitation light in the heralding detectors, we make use of a cross-polarization scheme and perform temporal selection of the detected photons as described in ref. ³⁷. We start the detection windows 4 ns (5 ns) after the highest intensity point of the excitation pulse, for the AB (BC) entangled link, to ensure sufficient suppression of excitation laser light in the detection window.

Memory qubit coherence Bob

We use the sequence described in Fig. 3a to preserve the state of the memory qubit during entanglement attempts. To characterize the decoupling sequence, we compare it to the sequence in which we do not apply the decoupling pulse on the memory qubit and/or the sequence in which we idle instead of performing entanglement attempts. We characterize the coherence of the memory qubit by storing the six cardinal states. We average the results for the eigenstates $\langle |0\rangle, |1\rangle \rangle$ and superposition states $\langle |\pm\text{rm}\{X\}\rangle, |\pm\text{rm}\{Y\}\rangle \rangle$. In Extended Data Fig. 5a, we plot the Bloch vector length $b = \sqrt{b_x^2 + b_y^2 + b_z^2}$, with b_i the Bloch vector component in direction i .

Over the measured range, the eigenstates show little decay. The decay of the superposition states is fitted with the function $f(x) = A \cdot e^{-((x/N)^{1/\text{rm}\{e\}})^n}$. The fitted parameters can be found in Extended Data Fig. 5b.

The use of the decoupling pulse π_M on the memory qubit increases the $N_{1/e}$ by more than a factor of 6. Moreover, the initial Bloch vector length A is higher with the π_M pulse. This is mainly explained by the second round of phase stabilization¹⁸ in between swapping the state onto the memory qubit and starting the entanglement generation process. The phase stabilization takes $\approx 350 \mu\text{s}$ and, during this time, the memory qubit is subject to intrinsic $\langle T \rangle^2 \text{last}$ dephasing, which can be efficiently decoupled using the π_M pulse.

Communication qubit coherence

In various parts of the protocol, we decouple the communication qubits from the spin bath environment to extend their coherence time. On Alice, we start the decoupling when the first entangled link is established and stop when the results of the BSM to teleport the state are sent by Charlie. On Bob, we decouple the communication qubit when the memory qubit is being rephased. On Charlie, the communication qubit is decoupled from the point that entanglement with Bob is heralded up to the point at which Bob has

finished the BSM, performed the CR check and has communicated the results. All these decoupling times are dependent on how many entanglement attempts are needed to generate the entangled link between Bob and Charlie.

We characterize the average state fidelities for different decoupling times; see Extended Data Fig. 6a. We investigate eigenstates and superposition states separately. We fit the fidelity with the function $f(t) = A \{ \{ \text{rm} \{ e \} \} \}^{-\{ (t/\{\tau\}_{\{ \{ \text{rm} \{ coh \} \} \}})\}^{\{ n \}} + 0.5.}$ The fitted parameters are summarized in Extended Data Fig. 6b. For each setup, the minimum and maximum decoupling times used are indicated by the shaded regions in Extended Data Fig. 6a. The left-most border is the decoupling time when the first entanglement attempt on Bob and Charlie would be successful and the right-most border is when the last attempt before the timeout of 1,000 attempts would herald the entangled state.

Model of the teleported state

A detailed model of the teleported state can be found at <https://doi.org/10.4121/16645969>. The model comprises elements from ref. 18 and is further extended for the teleportation protocol. We take the following noise sources into account:

- Imperfect Bell states between Alice and Bob, and between Bob and Charlie.
- Dephasing of the memory qubit of Bob during entanglement generation between Bob and Charlie.
- Depolarizing noise on the memory qubits of Bob and Charlie, owing to imperfect initialization and swap gates.
- Readout errors on the communication qubits of Bob and Charlie and readout errors on the memory qubits of Bob and Charlie when using the basis-alternating readout scheme, which result in incorrect feed-forward gate operations after the BSMs.
- Depolarizing noise on Alice during the decoupling sequence.

- Ionization probability on Alice.

An overview of the input parameters and the effect of the different error sources are given in Extended Data Table 4.

Calculation of teleported state fidelity without feed-forward operation

In Fig. 4c, we show the fidelity of the teleported state in case no feed-forward operations would have been applied on Alice. To extract this data, we follow the same method as in ref. 9. We perform classical bit flips on the measurement outcomes to counteract the effect of the feed-forward gate operations (as if the gate was not applied) for each BSM outcome. We do this for all six cardinal states and compute the average fidelity. We assume the errors of the gate in the feed-forward operations to be small.

Data availability

The datasets that support this manuscript and the software to analyse them are available at <https://doi.org/10.4121/16645969>.

References

1. Kimble, H. J. The quantum internet. *Nature* **453**, 1023–1030 (2008).
2. Wehner, S., Elkouss, D. & Hanson, R. Quantum internet: a vision for the road ahead. *Science* **362**, eaam9288 (2018).
3. Bennett, C. H. et al. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
4. Bouwmeester, D. et al. Experimental quantum teleportation. *Nature* **390**, 575–579 (1997).
5. Boschi, D., Branca, S., De Martini, F., Hardy, L. & Popescu, S. Experimental realization of teleporting an unknown pure quantum state

- via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.* **80**, 1121–1125 (1998).
6. Furusawa, A. et al. Unconditional quantum teleportation. *Science* **282**, 706–709 (1998).
 7. Olmschenk, S. et al. Quantum teleportation between distant matter qubits. *Science* **323**, 486–489 (2009).
 8. Nölleke, C. et al. Efficient teleportation between remote single-atom quantum memories. *Phys. Rev. Lett.* **110**, 140403 (2013).
 9. Pfaff, W. et al. Unconditional quantum teleportation between distant solid-state quantum bits. *Science* **345**, 532–535 (2014).
 10. Langenfeld, S. et al. Quantum teleportation between remote qubit memories with only a single photon as a resource. *Phys. Rev. Lett.* **126**, 130502 (2021).
 11. Ben-Or, M., Crépeau, C., Gottesman, D., Hassidim, A. & Smith, A. Secure multiparty quantum computation with (only) a strict honest majority. In *Proc. 2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)* 249–258 (IEEE, 2006).
 12. Arora, A. S., Roland, J. & Weis, S. Quantum weak coin flipping. In *Proc. 51st Annual ACM Symposium on Theory of Computing (STOC 2019)* 205–216 (ACM, 2019).
 13. Van Meter, R. *Quantum Networking* (Wiley, 2014).
 14. Bao, X.-H. et al. Quantum teleportation between remote atomic-ensemble quantum memories. *Proc. Natl Acad. Sci.* **109**, 20347–20351 (2012).
 15. Briegel, H.-J., Dür, W., Cirac, J. I. & Zoller, P. Quantum repeaters: the role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932–5935 (1998).

16. Cabrillo, C., Cirac, J. I., García-Fernández, P. & Zoller, P. Creation of entangled states of distant atoms by interference. *Phys. Rev. A* **59**, 1025–1033 (1999).
17. Bose, S., Knight, P. L., Plenio, M. B. & Vedral, V. Proposal for teleportation of an atomic state via cavity decay. *Phys. Rev. Lett.* **83**, 5158–5161 (1999).
18. Pompili, M. et al. Realization of a multinode quantum network of remote solid-state qubits. *Science* **372**, 259–264 (2021).
19. Humphreys, P. C. et al. Deterministic delivery of remote entanglement on a quantum network. *Nature* **558**, 268–273 (2018).
20. Legero, T., Wilk, T., Kuhn, A. & Rempe, G. Time-resolved two-photon quantum interference. *Appl. Phys. B* **77**, 797–802 (2003).
21. Bradley, C. et al. A ten-qubit solid-state spin register with quantum memory up to one minute. *Phys. Rev. X* **9**, 031045 (2019).
22. Cramer, J. et al. Repeated quantum error correction on a continuously encoded qubit by real-time feedback. *Nat. Commun.* **7**, 11526 (2016).
23. Robledo, L. et al. High-fidelity projective read-out of a solid-state spin quantum register. *Nature* **477**, 574–578 (2011).
24. Jiang, L. et al. Repetitive readout of a single electronic spin via quantum logic with nuclear spin ancillae. *Science* **326**, 267–272 (2009).
25. van Enk, S. J., Lütkenhaus, N. & Kimble, H. J. Experimental procedures for entanglement verification. *Phys. Rev. A* **75**, 052318 (2007).
26. Broadbent, A., Fitzsimons, J. & Kashefi, E. Universal blind quantum computation. In *Proc. 2009 50th Annual IEEE Symposium on Foundations of Computer Science* 517–526 (IEEE, 2009).
27. Rose, B. C. et al. Observation of an environmentally insensitive solid-state spin defect in diamond. *Science* **361**, 60–63 (2018).

28. Nguyen, C. et al. Quantum network nodes based on diamond qubits with an efficient nanophotonic interface. *Phys. Rev. Lett.* **123**, 183602 (2019).
29. Trusheim, M. E. et al. Transform-limited photons from a coherent tin-vacancy spin in diamond. *Phys. Rev. Lett.* **124**, 023602 (2020).
30. Son, N. T. et al. Developing silicon carbide for quantum spintronics. *Appl. Phys. Lett.* **116**, 190501 (2020).
31. Lukin, D. M., Guidry, M. A. & Vučković, J. Integrated quantum photonics with silicon carbide: challenges and prospects. *PRX Quantum* **1**, 020102 (2020).
32. Kindem, J. M. et al. Control and single-shot readout of an ion embedded in a nanophotonic cavity. *Nature* **580**, 201–204 (2020).
33. Chen, S., Raha, M., Phenicie, C. M., Ourari, S. & Thompson, J. D. Parallel single-shot measurement and coherent control of solid-state spins below the diffraction limit. *Science* **370**, 592–595 (2020).
34. Ruf, M., Wan, N. H., Choi, H., Englund, D. & Hanson, R. Quantum networks based on color centers in diamond. *J. Appl. Phys.* **130**, 070901 (2021).
35. Grein, M. E., Stevens, M. L., Hardy, N. D. & Benjamin Dixon, P. Stabilization of long, deployed optical fiber links for quantum networks. In *Proc. 2017 Conference on Lasers and Electro-Optics (CLEO 2017)* 1–2 (IEEE, 2017).
36. Dahlberg, A. et al. A link layer protocol for quantum networks. In *Proc. ACM Special Interest Group on Data Communication (SIGCOMM’19)* 159–173, (ACM, 2019).
37. Hensen, B. et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).

Acknowledgements

We thank S. Wehner, T. Taminiau, C. Bradley and H. de Riedmatten for discussions. We acknowledge financial support from the EU Flagship on Quantum Technologies through the project Quantum Internet Alliance (EU Horizon 2020, grant agreement no. 820445); from the European Research Council (ERC) through an ERC Consolidator Grant (grant agreement no. 772627 to R.H.); from the Netherlands Organisation for Scientific Research (NWO) through a VICI grant (project no. 680-47-624) and the Zwaartekracht program Quantum Software Consortium (project no. 024.003.037/3368). S.B. acknowledges support from an Erwin-Schrödinger fellowship (QuantNet, no. J 4229-N27) of the Austrian National Science Foundation (FWF).

Author information

Author notes

1. S. Baier

Present address: Institut für Experimentalphysik, Universität Innsbruck, Innsbruck, Austria

2. These authors contributed equally: S. L. N. Hermans, M. Pompili

Authors and Affiliations

1. QuTech and Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands

S. L. N. Hermans, M. Pompili, H. K. C. Beukers, S. Baier, J. Borregaard & R. Hanson

Contributions

S.L.N.H., M.P. and R.H. devised the experiment. S.L.N.H., M.P. and H.K.C.B. carried out the experiments and collected the data. S.L.N.H., M.P., H.K.C.B. and S.B. prepared the experimental apparatus. J.B. developed the quantum-optical model. S.L.N.H. and R.H. wrote the main manuscript, with

input from all authors. S.L.N.H., M.P. and J.B. wrote the supplementary materials, with input from all authors. S.L.N.H. and M.P. analysed the data and discussed with all authors. R.H. supervised the research.

Corresponding author

Correspondence to [R. Hanson](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Florian Kaiser and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. [Peer reviewer reports](#) are available.

Additional information

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

Extended data figures and tables

[Extended Data Fig. 1 PSB-flagged correlations Alice–Bob.](#)

Top, histograms of the detected PSB photons conditioned on a simultaneous ZPL detection in the entanglement generation attempt, for Alice (left) and Bob (right). Bottom, corresponding measured correlations in all bases. The grey bars in the Z basis represent the simulated values. For the X and Y

bases, one would expect a probability of 0.25 for all outcomes. All error bars represent one standard deviation.

Extended Data Fig. 2 PSB-flagged correlations Bob–Charlie.

Top, histograms of the detected PSB photons conditioned on a simultaneous ZPL detection in the entanglement generation attempt, for Bob (left) and Charlie (right). Bottom, corresponding measured correlations in all bases. The grey bars in the Z basis represent the simulated values. For the X and Y bases, one would expect a probability of 0.25 for all outcomes. All error bars represent one standard deviation.

Extended Data Fig. 3 Basis-alternating repetitive readout.

Basis-alternating repetitive (BAR) readout results for Charlie’s memory qubit. **a**, Readout fidelity for each readout repetition, for states $|0\rangle$ and $|1\rangle$. **b**, Readout fidelity of the BAR readout scheme for different number of readout repetitions. **c**, Fraction of inconsistent readout patterns for different number of readout repetitions. The dashed lines represent a numerical model using measured parameters, which can be found at <https://doi.org/10.4121/16645969>. All error bars represent one standard deviation.

Extended Data Fig. 4 Experimental rates.

Experimental rates of the conditional and unconditional teleportation protocol for different detection window lengths in the two-node entanglement generation.

Extended Data Fig. 5 Memory qubit coherence.

a, Coherence of Bob’s memory qubit for superposition states (triangles and circles) and eigenstates (squares and diamonds). We perform the sequence as described in the main text with and without the decoupling pulse π_M on the memory qubit, the dark blue and purple points, respectively. Furthermore, we perform the sequence with a wait time instead of entanglement attempts with (pink points) and without (yellow points) the decoupling pulse. The

grey dashed line indicates the timeout of the entanglement generation process used in the teleportation protocol. **b**, Fitted parameters for the memory coherence decay of the superposition states. All error bars represent one standard deviation.

Extended Data Fig. 6 Communication qubit coherence.

a, Decoupling of the communication qubits. The average state fidelity is plotted for different decoupling times for each setup. The shaded area represents the decoupling times used in the teleportation protocol. **b**, Fitted parameters for average state fidelity during communication qubit decoupling. All error bars represent one standard deviation.

Extended Data Table 1 Teleported state fidelities

Extended Data Table 2 Average teleported state fidelities per BSM outcome

Extended Data Table 3 Memory qubit characteristics

Extended Data Table 4 Two-node and teleportation simulation parameters

Supplementary information

Supplementary Information

This file contains Supplementary Text; equations, figures, tables and references.

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The

images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Hermans, S.L.N., Pompili, M., Beukers, H.K.C. *et al.* Qubit teleportation between non-neighbouring nodes in a quantum network. *Nature* **605**, 663–668 (2022). <https://doi.org/10.1038/s41586-022-04697-y>

- Received: 05 October 2021
- Accepted: 29 March 2022
- Published: 25 May 2022
- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04697-y>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Breakthrough in teleportation furthers quantum network development

- Oliver Slattery
- Yong-Su Kim

Nature News & Views 25 May 2022

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04697-y>.

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

[Download PDF](#)

- Article
- Open Access
- [Published: 25 May 2022](#)

Enhanced silica export in a future ocean triggers global diatom decline

- [Jan Taucher](#) [ORCID: orcid.org/0000-0001-9944-0775¹](#),
- [Lennart T. Bach](#) [ORCID: orcid.org/0000-0003-0202-3671²](#),
- [A. E. Friederike Prowe](#) [ORCID: orcid.org/0000-0001-5281-4696¹](#),
- [Tim Boxhammer](#) [ORCID: orcid.org/0000-0002-9632-5947¹](#),
- [Karin Kvale](#) [ORCID: orcid.org/0000-0001-8043-5431^{1,3}](#) &
- [Ulf Riebesell](#) [ORCID: orcid.org/0000-0002-9442-452X¹](#)

[Nature](#) volume **605**, pages 696–700 (2022)

- 2244 Accesses
- 273 Altmetric
- [Metrics details](#)

Subjects

- [Biooceanography](#)
- [Climate-change impacts](#)
- [Element cycles](#)
- [Marine biology](#)
- [Marine chemistry](#)

Abstract

Diatoms account for up to 40% of marine primary production^{1,2} and require silicic acid to grow and build their opal shell³. On the physiological and ecological level, diatoms are thought to be resistant to, or even benefit from, ocean acidification^{4,5,6}. Yet, global-scale responses and implications for biogeochemical cycles in the future ocean remain largely unknown. Here we conducted five in situ mesocosm experiments with natural plankton communities in different biomes and find that ocean acidification increases the elemental ratio of silicon (Si) to nitrogen (N) of sinking biogenic matter by 17 ± 6 per cent under $\backslash(\{ \{ p \} _ \{ \{ \backslash rm{CO} \} \} _ \{ 2 \} \} \backslash)$ conditions projected for the year 2100. This shift in Si:N seems to be caused by slower chemical dissolution of silica at decreasing seawater pH. We test this finding with global sediment trap data, which confirm a widespread influence of pH on Si:N in the oceanic water column. Earth system model simulations show that a future pH-driven decrease in silica dissolution of sinking material reduces the availability of silicic acid in the surface ocean, triggering a global decline of diatoms by 13–26 per cent due to ocean acidification by the year 2200. This outcome contrasts sharply with the conclusions of previous experimental studies, thereby illustrating how our current understanding of biological impacts of ocean change can be considerably altered at the global scale through unexpected feedback mechanisms in the Earth system.

Main

Global phytoplankton biogeography is tightly linked to the nutrient availability in the surface ocean^{7,8}. Diatoms, the dominant group of silicifiers, sustain some of the most productive marine ecosystems and are a major driver of biological CO₂ sequestration in the oceans^{1,9}. In contrast to most other phytoplankton taxa, they require silicic acid (Si(OH)₄) for biominerization of their opaline shells, called frustules. Therefore, the availability of Si compared to other major nutrients such as N determines the large-scale distribution of diatoms^{10,11}. By gravitational sinking of biogenic particles (known as the biological pump)¹², Si and N are stripped out of the surface ocean and transported to deeper water layers, where remineralization of organic matter and chemical dissolution of biogenic opal convert them back to their dissolved forms. Thus, the stoichiometric Si:N ratio of particulate matter export (Si:N_{export}), in combination with physical transport

by global ocean circulation, determines the large-scale distribution of inorganic nutrients over long timescales and the prevalence of diatoms in the world ocean^{10,13,14}. However, potential future changes in global nutrient distributions of Si and N and how they may affect diatoms in the oceans are presently unknown.

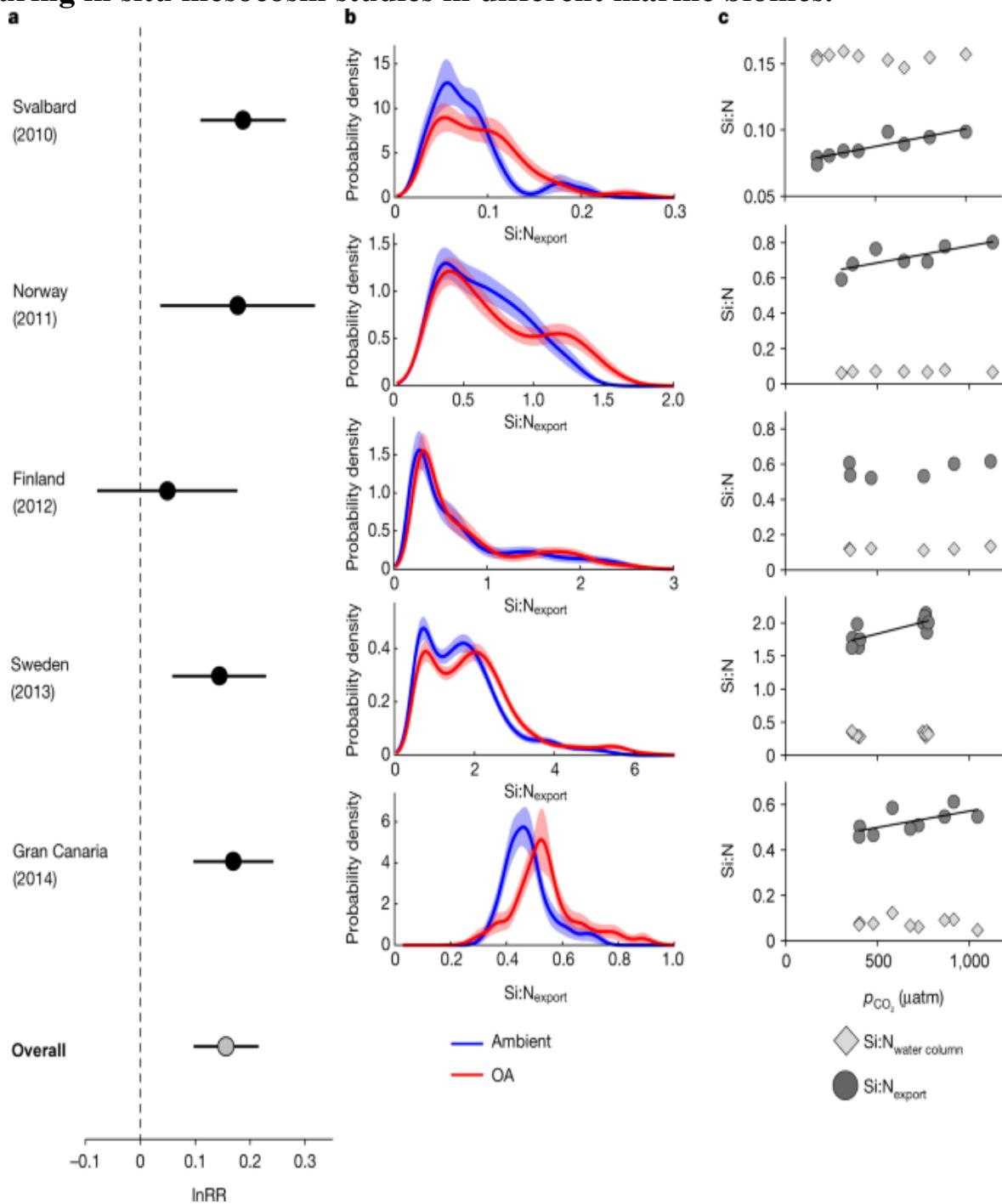
The current state of ocean acidification (OA) research suggests that diatoms will be primarily affected on the physiological level (for example, by changing the availability of CO₂ for photosynthesis) or by altered ecological interactions within the plankton community. Numerous physiological and ecological studies have made it evident that OA effects on diatoms are very variable⁴, probably owing to the diversity in relevant traits (for example, cell size or carbon-uptake mechanism), as well as interactions with other environmental factors such as light, nutrients and temperature⁶. Overall, a growing body of evidence suggests that positive effects on diatoms are more common than negative effects, both on the species and community level⁵. Consequently, diatoms are considered to be among the ‘winners’ of OA.

Impacts of OA on Si:N_{export}

We analysed data on the Si:N composition of vertical particle fluxes from five in situ mesocosm experiments, in which natural plankton communities (including diatoms) in different biomes were exposed to simulated OA (see [Methods](#) and Extended Data Fig. 1). This experimental approach enabled us to differentiate between OA effects on production (for example, responses of the diatom community) and on export/degradation (for example, changes in the composition of biogenic particles during sinking). To test for a systemic influence of OA on Si:N_{export}, we quantified OA effects using log-response ratios (lnRR) and probability densities for end-of-century CO_2 concentrations according to Representative Concentration Pathway (RCP) 6.0 to 8.5 scenarios of the fifth Intergovernmental Panel on Climate Change (IPCC) assessment report¹⁵. Our analysis reveals an overall increase of 17% (lnRR = 0.16) in the Si:N_{export} ratio under OA conditions (Fig. 1). Significant effects occurred in 4 out of 5 studies, with very similar magnitudes (lnRR ranging from 0.14 in Sweden to 0.19 in Svalbard; see Methods for calculation of effect sizes).

Observed shifts in $\text{Si:N}_{\text{export}}$ were surprisingly consistent across the different biomes with variable contributions of diatoms to overall plankton biomass, as indicated by the range of baseline Si:N ratios (<0.1 in Svalbard to approximately 1.8 in Sweden; Fig 1b).

Fig. 1: Impacts of OA on Si:N of sinking particulate matter observed during in situ mesocosm studies in different marine biomes.



a, Effect size calculated as the log-transformed response ratio (lnRR) of the treatment averages with 95% confidence intervals, including overall effect across studies (light grey). **b**, Probability density estimates of Si:N_{export} under ambient conditions (blue) and OA (red) with shaded areas denoting confidence intervals. Non-overlapping density estimates indicate statistically significant differences in the probability distributions of Si:N. Note the different scaling of the x axes, due to the variable ranges of Si:N values. **c**, Differences between OA effects on Si:N of biogenic material suspended in the water column (Si:N_{water column}, light grey) and sinking particulate matter collected in sediment traps (Si:N_{export}, dark grey). Black lines indicate that linear regression analysis was significant ($p < 0.05$). In the majority of studies (4 out of 5), we found a distinct influence of increasing $\{\{p\}\}_{\{\{\rm CO\}\}}_{\{2\}}$ on the Si:N_{export}, but no influence on Si:N_{water column}. See Methods and Extended Data Table 1 for an overview of mesocosm experiments.

Notably, the influence of OA on Si:N was only detectable for sinking particles (collected in sediment traps), but not for particulate matter suspended in the water column (Fig. 1c). This suggests that the observed OA effects on Si:N emerged primarily while the biogenic detrital particles were sinking and not due to biotic effects during their production (see [Methods](#)). Most probably, lower pH accelerated the dissolution of sinking opal and thereby increased Si:N_{export}. Seawater is generally undersaturated in silicic acid and thus corrosive to the amorphous silica of diatom shells¹⁶. Thus, diatoms protect their cell wall silica against chemical dissolution with an organic coating surrounding the cell¹⁷. Once this protective coating is degraded by bacteria (with the onset of senescence), dissolution rates of diatom silica increase drastically¹⁸. This usually coincides with the termination of diatom blooms and sinking of this biomass (for example, as marine snow). This mechanism explains why an OA effect on opal dissolution in our mesocosm data only becomes apparent in sinking material but not in suspended particulate matter. This interpretation is supported by both theory and empirical studies, which have demonstrated that chemical dissolution rates of biogenic silica are reduced at lower pH^{17,19,20}. The observed OA effect in our experiments was a 17% increase in Si:N_{export} for a pH decrease of around 0.3. Assuming that this was solely driven by the pH

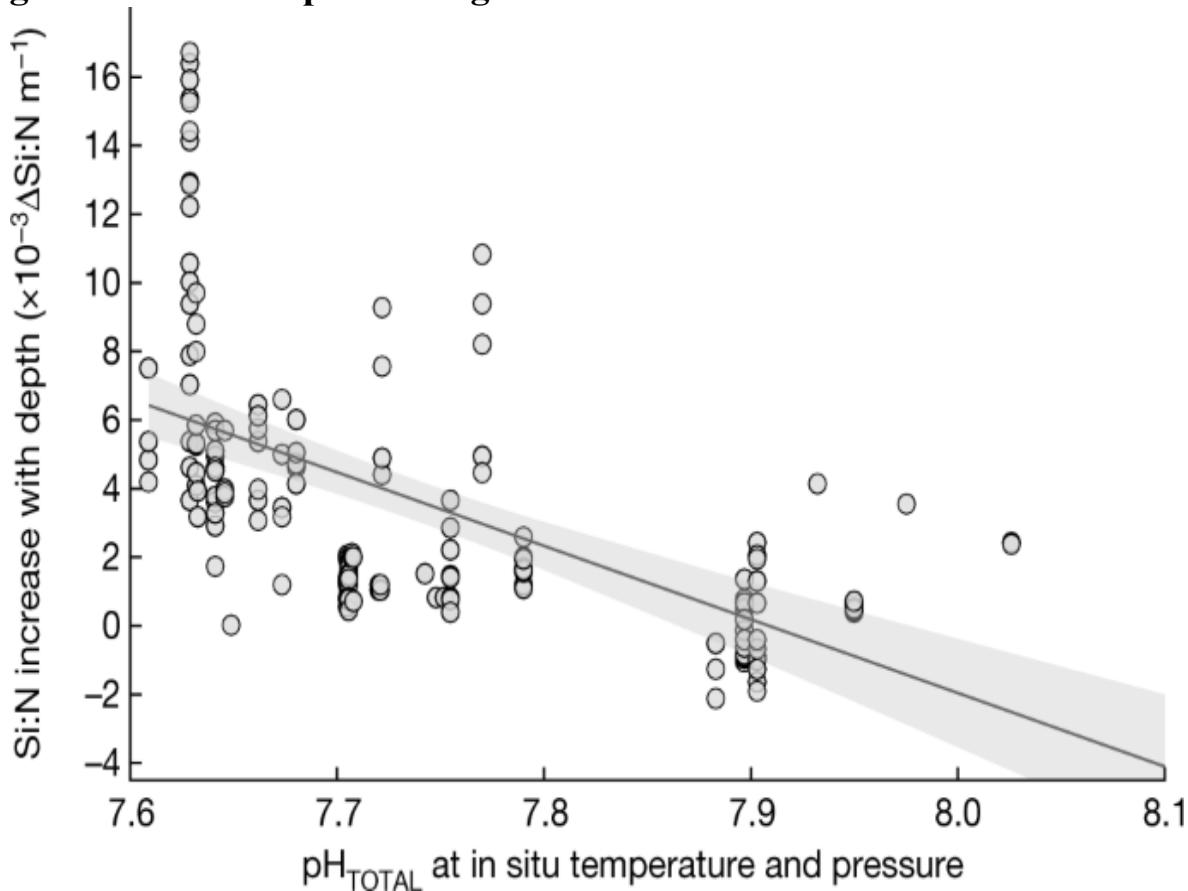
sensitivity of opal dissolution, this corresponds to a decrease in opal dissolution rate of 57% per unit pH. This number agrees very well with results from chemical studies that found 60% to 70% changes in silica dissolution per unit pH^{20,21}. Furthermore, these results are in line with theoretical dissolution kinetics: biogenic silica mainly consists of tetrahedrally bonded Si–O–Si silica, but also contains hydroxylated silica (for example, silanol: Si–OH). The chemical dissolution of biogenic silica is controlled by the breaking of bridging Si–O bonds, which is facilitated at higher pH owing to increased deprotonation of silanol groups at the reactive particle surface^{19,22}. Although alternative explanations for the observed OA impact on Si:N cannot be fully excluded (see [Methods](#)), the various independent lines of evidence and the consistency of datasets (mesocosms, ocean sediment traps, previous chemical studies) suggest that the pH dependence of biogenic silica dissolution is the most probable explanation for our findings.

This mechanism can also explain why such an effect of OA on Si:N_{export} has not been observed before, as few experimental studies have the capacity to differentiate between suspended and sinking material (that is, production and degradation of sinking particles). Furthermore, opal dissolution kinetics are probably also the reason why Finland is the only study site where no OA effect on Si:N was detected (Fig. 1): because salinity is another relevant factor controlling rates of chemical opal dissolution—higher salinity enhances dissolution rates^{19,23}—the low salinity of approximately 6 psu (practical salt units) in Finland probably obscured any potential influence of OA. It should be noted that response patterns of Si:C were more variable than those detected for Si:N, owing to additional and variable shifts in C:N under OA (Extended Data Fig. 2).

To further test our interpretation that the observed OA impacts on Si:N_{export} were driven by the pH sensitivity of opal dissolution, we analysed a global compilation of sediment trap data to examine the influence of pH on Si:N ratios of sinking particles in the ocean. It is widely known that opal dissolution in the water column is slower than organic matter remineralization, resulting in particulate matter that is progressively Si-enriched while sinking to depth^{24,25}. Our analysis reveals that this preservation of Si compared to N (given as the increase in Si:N with depth)

in the oceanic water column is enhanced at lower pH (Fig. 2). This result is consistent with our findings from our OA experiments, demonstrating that pH has a relevant role in controlling opal dissolution and the remineralization depth of silica in the ocean. Yet, this effect has so far been overlooked in the context of global-scale OA research, because the primary factor controlling the dissolution of biogenic silica in the water column is known to be temperature^{16,26} (see [Methods](#)).

Fig. 2: Influence of pH on biogenic silica dissolution in the ocean.



Shown is the relationship between pH and the increase in Si:N_{export} with depth ($\Delta \text{Si:N m}^{-1}$) as an indicator for Si preservation and dissolution. Data are from a global compilation of sediment trap data³⁵. Results from linear regression analysis reveal a significant relationship between pH and change in Si:N ratios with depth (slope = -0.0214 , $R^2 = 0.24$, $p < 10^{-11}$, $f = 58.1$, $n = 190$). Shaded area denotes 95% confidence intervals of the regression.

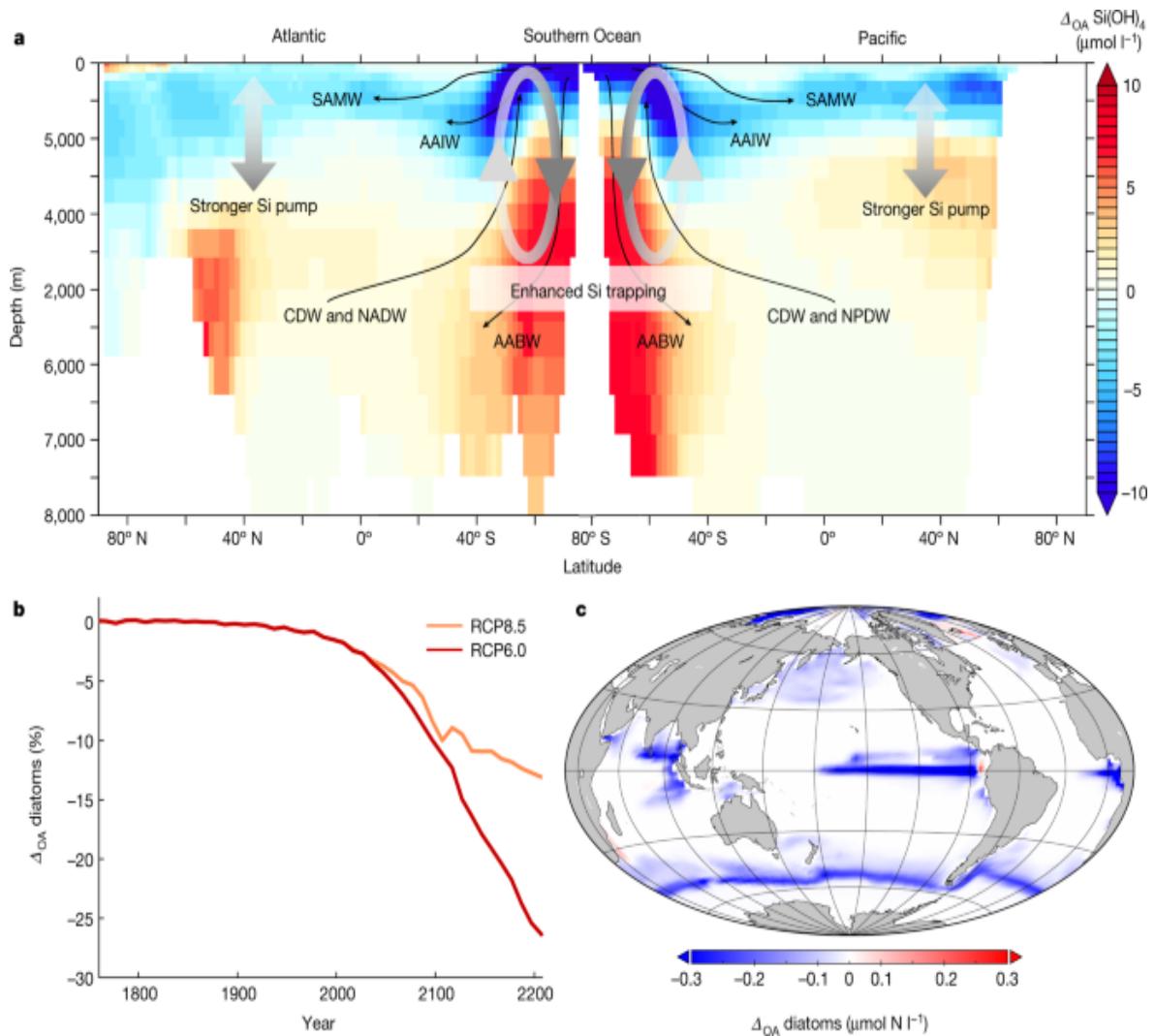
Global biogeochemical implications

The consistency of our experimental results across plankton biomes, together with the evidence for pH-dependent opal dissolution as the major driving mechanism, provide a well founded mechanistic explanation for our findings, suggesting that the sensitivity of Si:N_{export} to OA is a globally relevant mechanism. We applied an Earth system model (the University of Victoria Earth System and Climate Model, UVic ESCM) to assess how changes in Si:N_{export}—generated by pH-sensitive opal dissolution—affect global plankton biogeography and nutrient distributions under sustained OA until the year 2200 (following extended RCP8.5 and RCP6.0 scenarios)²⁷. Consistent with other Earth system models, our simulations predict profound impacts of climate change on the world ocean, for example, sea surface warming and decreased nutrient supply, which in turn lead to a decline in phytoplankton biomass (including diatoms^{28,29}; see [Methods](#)). On top of that, our simulations with pH effects on opal dissolution and Si:N_{export} reveal additional biogeochemical impacts of OA that entail far-reaching consequences for the ocean silica cycle and diatoms over the coming centuries. In this context, it is noteworthy that future changes in temperature and pH have antagonistic effects on opal dissolution—that is, the well known acceleration due to warming and the here-revealed slowdown in response to OA, with the effect of decreasing pH overcompensating for effects of warming (see [Methods](#) and Extended Data Fig. 3).

Slower opal dissolution under OA amplifies the silicate pump, that is, enhancing the transfer of Si from the surface to the deep ocean (Fig. 3a). This loss of Si(OH)₄ from the productive surface layer is strongest in the Southern Ocean, where present diatom-driven productivity and opal export are highest (Extended Data Fig. 4), and to a lesser extent in the North Pacific and Atlantic oceans. Generally, the Southern Ocean has a key role in the ocean Si cycle (i) by acting as a ‘silicon trap’ (about half of the global silicic acid inventory is recycled here)¹⁴, and (ii) through the formation of mode and intermediate waters, which constitute major sources of nutrients for primary production in large parts of the global ocean via the meridional overturning circulation^{10,30}. Both mechanisms are altered by OA in our simulations. Elevated Si:N_{export} in effect redistributes Si to depths greater

2,000–3,000 m south of the Sub-Antarctic Front ($\sim 55^\circ$ S), where it enters mainly Antarctic bottom water and partly returns towards the divergence zone of the Southern Ocean through upwelling with circumpolar deep water. This creates a cycle of particle export with elevated Si:N and slower opal dissolution and upwelling of Si-enriched circumpolar deep water in the Antarctic divergence zone, which is, however, not enough to counteract the stronger silicate pump (Fig. 3a). Together, this gradually depletes Si(OH)_4 in the upper 1,000–2,000 m and enhances permanent accumulation of Si ('trapping') in the deep Southern Ocean. As a consequence, Si is diminished in Antarctic mode and intermediate waters, which distribute nutrients northward within the upper limb of the global overturning circulation^{10,30}. Once the Antarctic mode and intermediate waters reach the surface in the northern hemisphere (mainly in the Atlantic), these already Si-deficient water masses fuel diatom production and elevated Si:N_{export} further redistributes Si from the surface to deep ocean.

Fig. 3: Global impacts of slower silica dissolution under OA.



a, Differences in the global vertical distribution of $\text{Si}(\text{OH})_4$ (zonally averaged for the Atlantic and Pacific Ocean), including a depiction of the OA-driven downward transport of Si and its trapping in the deep ocean. AABW, Antarctic bottom water; AAIW, Antarctic intermediate water; CDW, circumpolar deep water; NADW, North Atlantic deep water; NPDW, North Pacific deep water; SAMW, sub-Antarctic mode water. **b**, OA effect on global diatom biomass under RCP8.5 and RCP6.0 emission scenarios. **c**, Spatial distribution of the difference in diatom biomass. Results show the net effect of the OA-driven slowdown of silica dissolution (Δ_{OA}), given as the difference to the standard model configuration (that is, including impacts of climate change, but excluding OA effects on Si dissolution). Panels **a** and **c** show results from year 2200 of the RCP8.5 simulation. See Extended Data Table 3 and Extended Data Fig. 5 for additional results and visualizations.

The combination of these processes results in an OA-driven global loss of Si from the surface ocean and its trapping in the deep ocean. Global average Si(OH)_4 concentrations in the surface ocean decrease by -11% to -27% owing to OA until the year 2200 (RCP6.0 and RCP8.5, respectively). The progressive decline in Si(OH)_4 in the euphotic zone reduces the availability of this essential resource for diatoms, resulting in an OA-driven decline of global diatom biomass of -13% to -26% by 2200 (for RCP6.0 and RCP8.5, respectively; Fig. 3b). Almost half of this decline occurs at the northern boundary of their Southern Ocean habitat around 55°S (Fig. 3c), where Si(OH)_4 was already limiting under preindustrial conditions (Extended Data Fig. 4) and becomes progressively depleted owing to OA-enhanced $\text{Si:N}_{\text{export}}$. We note that the above numbers refer to the net effect of OA-driven changes in silica dissolution that occur on top of other impacts of climate change—for example, reduced nutrient supply due to enhanced thermal stratification (see Methods and Extended Data Table 3 for additional results).

Although these results should be considered as a first estimate of how OA impacts on $\text{Si:N}_{\text{export}}$ may affect the future biogeography of diatoms, we emphasize that the simulated global decline of diatoms is driven by an abiotic mechanism, thereby substantially reducing the uncertainty that is usually inherent to the complex and variable biotic responses to OA^{31,32,33}. Therefore, it is reasonable to assume that the OA impacts on $\text{Si:N}_{\text{export}}$ and opal dissolution would yield similar results if incorporated into other global models. However, associated consequences for ecosystem functioning and carbon cycling are more difficult to assess, as our simulations (similar to most Earth system models) do not account for potentially relevant physiological and ecological mechanisms that might trigger knock-on effects on food-web structure and the biological pump³⁴ (see Methods for further discussion).

Taken together, our findings from the mesocosm experiments and Earth system model simulations suggest that OA will probably induce a systemic increase in $\text{Si:N}_{\text{export}}$, thereby causing a decrease in Si(OH)_4 in the surface ocean and a global decline of diatoms over the next centuries. This result contrasts the findings of physiological and ecological experiments that

suggest that diatoms may be resistant to or benefit from OA^{4,5,6}. Notably, the outcome in our study arises owing to a previously overlooked biogeochemical feedback in the marine silicon cycle induced by OA, not by an effect of elevated CO₂ on diatom physiology. Interestingly, this feedback mechanism via the silicon pump could be further exacerbated should OA favour the competitiveness of diatoms in phytoplankton communities, as indicated by aforementioned studies. In this case, diatoms would even more efficiently deplete Si in the surface ocean and accelerate their eventual decline. Our study demonstrates how expected responses of marine biota to OA can be altered or even turned upside down at the global scale through unforeseen biogeochemical feedback mechanisms. In a broader context, this exemplifies that our current understanding of biological impacts of ocean change, which is largely based on single-species and small-scale ecological studies, might be deceptive when not considering the complexity of the Earth system.

Methods

Mesocosm experiments

Si:N_{export} measurements

Between 2010 and 2014, we conducted five in situ mesocosm experiments to assess impacts of OA on natural plankton communities. Study sites covered a large latitudinal gradient (28 °N–79 °N) and diverse oceanic environments/ecosystems (Extended Data Fig. 1 and Extended Data Table 1). Sample collection and processing was conducted every 1 or 2 days throughout the experiments. Sinking particulate matter was obtained from sediment traps attached to the bottom of each mesocosm, thereby collecting the entire material sinking down in the enclosed water column³⁶. Processing of sediment trap samples followed a previous protocol³⁷. Samples for particulate matter suspended in the water column were collected with depth-integrating water samplers (HYDRO-BIOS) and filtered following standard procedures. Biogenic silica was leached from the sediment trap samples and filters by alkaline pulping (0.1 M NaOH at 85 °C). After 135 min the leaching process was terminated with 0.05 M H₂SO₄ and dissolved silica

was measured spectrophotometrically³⁸. Carbon and nitrogen content were determined using an elemental CN analyser (EuroEA)³⁹.

Analysis of OA impacts

To test for a systemic influence of OA on Si:N_{export}, we synthesized the datasets from the different experiments and (i) conducted a meta-analysis to quantify effect sizes, and (ii) computed probability density estimates. Because the experimental design, the range of CO₂ treatments, and the time periods for our analysis of Si:N_{export} varied to some extent among experiments (Extended Data Table 1), we pooled mesocosms for ambient conditions and in the $\{p\}_{\{CO\}}_2$ range of ~700–1,000 μatm ('OA treatment'), corresponding to end-of-century values according to RCP 6.0 and 8.5 emission scenarios¹⁵. Effect sizes were calculated as log-transformed response ratios lnRR, an approach commonly used in meta-analysis⁴⁰:

$$\ln RR = \ln \left(\frac{\bar{X}_{OA}}{\bar{X}_{Ambient}} \right) = \ln \left(\frac{\sum n_{OA} \ln(n_{OA})}{\sum n_{Ambient} \ln(n_{Ambient})} \right)$$

where X is the arithmetic mean of Si:N_{export} ratios under OA and ambient conditions (Extended Data Table 1). Effect sizes <0 denote a negative effect of OA and effect sizes >0 indicate that the effect was positive. Effects are considered statistically significant when 95% confidence intervals (calculated from pooled standard deviations) do not overlap with zero. The overall effect size across all studies was computed by weighing individual effect sizes according to their variance, following the common methodology for meta-analyses⁴⁰. In addition, we computed probability densities of Si:N_{export} based on kernel density estimation, which better accounts for data with skewed or multimodal distributions⁴¹. Another advantage of this approach is that it does not require the calculation of temporal means. Instead, the entire data timeseries can be incorporated into the analysis, thus retaining information about temporal variability. Confidence intervals of the density estimates were calculated with a bootstrapping approach using data resampling (1,000 permutations)⁴¹. The resulting probability density plots

can be interpreted analogously to histograms. Differences among ambient and OA conditions are considered statistically significant when confidence intervals of the probability density distributions do not overlap. Numbers for suspended and sinking Si, C and N (and their respective ratios) for the analysis period are given in Extended Data Table 2.

Analysis of pH effects on Si:N in global sediment trap data

We analysed a recent compilation of global sediment trap data (674 locations collected between 1976 and 2012)³⁵. The aim of this analysis was to assess the influence of pH on opal dissolution in the world ocean. In contrast to the mesocosm experiments, where export fluxes were measured only at one depth, the global dataset provides depth-resolved information, enabling us to examine the vertical change in the Si:N ratio of sinking particulate matter and how this correlates with pH. It has long been known that the silica content of sinking particles increases with depth, as opal dissolution is less efficient than organic matter remineralization^{25,42}. The resulting accumulation of Si relative to N can be quantified as the change in Si:N with increasing depth, that is, the slope of the relationship of depth versus Si:N ($\Delta\text{Si:N}$, in units of m^{-1}). Our approach is analogous to previous studies, which used vertical profiles of Si:C as a proxy for differential dissolution/remineralization of opal and organic matter, and its regional variability in the ocean^{24,42}. We extracted all data that (I) included simultaneous measurements of Si and N, and (II) contained vertical profiles with at least three depth levels (so that $\Delta\text{Si:N}$ [m^{-1}] can be calculated). We then calculated linear regressions for individual Si:N profiles and subsequently extracted those for which Si:N displayed a statistically significant relationship with depth ($p < 0.05$). Thereby, profiles with no clear depth-related pattern of Si:N—for example, due water mass advection—were excluded. In total, 190 profiles of Si:N flux matched those criteria and were used for further analysis.

To assess the influence of pH on $\Delta\text{Si:N}$, we applied linear regression analysis, using the $\Delta\text{Si:N}$ for each vertical flux profile and the average pH (from the GLODAP database)⁴³ over the corresponding depth range. Because temperature is considered the primary factor driving opal dissolution^{16,26} and has been shown to influence the preservation of silica

compared to organic matter⁴⁴, we additionally conducted a multiple regression analysis for the same dataset, by including temperature as a second factor. Results confirm that temperature also has an influence on $\Delta\text{Si:N}$ [m^{-1}], with a comparable effect size as identified for pH. However, because the role of temperature has been discussed extensively in earlier studies (see Methods section ‘The present and future role of pH for opal dissolution in the ocean’), here we focus on the previously overlooked effect of pH.

The present and future role of pH for opal dissolution in the ocean

It is noteworthy that pH effects on opal dissolution have so far been mostly neglected in oceanographic research and earlier work almost exclusively focused on temperature as the factor controlling opal dissolution^{16,24}. Although chemical dissolution experiments have demonstrated the pH dependence of biogenic silica dissolution^{20,21}, it has so far been considered a minor factor in oceanic silica cycling. The reason probably lies in present-day gradients of temperature and pH in the ocean, and the resulting influence on opal dissolution: A temperature gradient of ~15–20 °C (average difference between surface and deep ocean), corresponds to a three- to fourfold change in the opal dissolution rate^{26,44}. By comparison, the effect of pH described in our study is much smaller, indicating an ~20% difference in opal dissolution for present-day pH gradients of ~0.3–0.35. Accordingly, in the present ocean, the effect of temperature is roughly 10-fold larger compared to that of pH—which is probably why the latter has been so far neglected. In addition, the present-day vertical gradients of both temperature and pH both work in the same direction, that is, towards a decrease in opal dissolution rates with depth (decreasing temperature and decreasing pH).

However, the situation changes completely under future scenarios of OA and warming. Surface ocean pH is predicted to decrease by 0.2–0.4 until the year 2100, whereas sea surface warming may reach 1–3 °C (refs. ^{45,46}). For these changes, the effect size of both factors on opal dissolution is on a similar order of magnitude. More importantly, in contrast to present-day vertical gradients, their future changes have antagonistic effects on opal dissolution, that is, warming-driven acceleration and a pH-driven slowdown. Notably,

our results suggest that the effect of decreasing pH even overcompensates for the effect of warming on a global average (Extended Data Fig. 3). This illustrates that pH becomes an increasingly relevant factor for opal dissolution and the pelagic Si cycle in the context of ongoing climate change and OA.

Possible other contributing factors to OA impacts on Si:N

Although our results provide strong evidence for a chemical effect of decreasing pH on opal dissolution, additional or alternative explanations for the observed OA impacts on Si:N in the mesocosm studies cannot be fully excluded. On the basis of previous findings, one would expect that OA impacts on Si:N_{export} can be explained by responses of diatoms. From a physiological perspective, lower pH may theoretically facilitate silicification by diatoms. The solubility of Si in seawater decreases with decreasing pH, promoting precipitation and inhibiting dissolution of opal. Diatoms are known to utilize this physicochemical property to precipitate opal in a cellular compartment with low pH conditions^{47,48}. However, experimental evidence is scarce and partly controversial, with indications for either enhanced or reduced silica production under lower pH^{49,50}. From an ecological perspective, higher Si:N_{export} may have arisen from shifts in phytoplankton community composition, with a greater proportion of particle export driven by diatoms compared to other (non-silicifying) taxa, or by more heavily silicified species within the diatom community. However, our data do not support either of these two potential explanations, as the influence of OA on Si:N is only detectable for vertical particle fluxes (collected in sediment traps), but not for freshly produced particulate matter in the water column (Fig. 1c). This suggests that OA effects on Si:N emerged primarily while the biogenic detrital particles were sinking and not due to biotic effects during their production. Another possibility is that changes in N remineralization under simulated OA additionally contributed to the increase in Si:N. However, the current consensus is that bacterial communities and organic matter remineralization are mostly resilient to OA⁵¹. Results from studies that reported effects are very variable and, in most cases, it was not possible to separate direct pH effects (for example, on bacterial activity) from indirect effects mediated through pH-driven changes in quality and/or quantity of the organic matter substrate. Thus, there are

currently no indications that OA will enhance N consumption of sinking organic matter. Altogether, the various independent lines of empirical evidence (mesocosms, ocean sediment traps, chemical studies) and the consistency of their results suggest that the pH effect on Si dissolution is the most probable explanation for our findings.

Global impact assessment from Earth system model simulations

We incorporated the effects of simulated OA on Si:N_{export} observed in the in situ mesocosm experiments into an Earth system model to assess the global scale impacts on nutrient availability and plankton biogeography over the coming centuries. Consequently, we applied a modified version of the University of Victoria Earth System and Climate Model (UVic ESCM), which simulates silica cycling and diatoms, as well as other functional groups of phytoplankton as described in an earlier work⁵². Biogenic opal is produced by diatoms, including a parameterization for iron dependency of silicification, resulting in elevated Si:N ratios of production under iron limitation⁵². Vertical profiles of opal fluxes and dissolved silica are instantaneously computed based on biogenic silica production in the surface ocean and dissolution throughout the water column, and silica dissolution is parameterized as an exponential, temperature-dependent rate. Simulated present-day spatial distributions of Si(OH)₄ in the surface ocean agree well with observational data (Extended Data Fig. 4).

To simulate OA effects on silica dissolution and Si:N_{export}, we parameterized the specific silica dissolution rate to scale with changes in pH throughout the water column relative to preindustrial conditions for each box in the three-dimensional model grid, thereby accounting for the vertical characteristics of future pH changes (see Extended Data Fig. 3). Therefore, we assumed that the pH sensitivity of biogenic silica dissolution derived from the observed OA effect on Si:N_{export} (17% for ΔpH of around 0.3) is linear, corresponding to a decrease in opal dissolution rate of 57% per unit pH. This estimate agrees remarkably well with published rates from chemical dissolution experiments^{20,21}. We note that the model also accounts for effects of warming on silica dissolution (using a temperature dependence that is

similar to other global models) that work in the opposite direction as the OA effect (see Extended Data Fig. 3).

Model simulations were run for the period 1750 to 2200 using extended IPCC scenarios (RCP 8.5 and RCP 6.0) for atmospheric CO₂ concentrations²⁷. The reason why we conducted the simulations until the year 2200 (instead of 2100 as commonly done in other climate change studies) is that we expected impacts of the OA-driven slowdown of opal dissolution, such as Si trapping in the deep ocean, to emerge only on the long timescale of global circulation. However, as can be seen in Fig. 3b and Extended Data Fig. 5, OA-amplified Si trapping in the deep ocean and the resulting decline in diatoms become apparent by 2100. Thus, the reference year (2100 or 2200) only affects that magnitude of the effect; qualitatively, the results are very similar.

Generally, simulated impacts of climate change are consistent with other models, for example, reduced nutrient supply to the surface ocean and an associated decrease in phytoplankton biomass. The most important results that are relevant for the interpretation of OA effects reported here (via slower silica dissolution) are presented in Extended Data Table 3. More details on model behaviour in climate change simulations can be found elsewhere⁵².

We emphasize that the simulated OA effect on silica dissolution occurs on top of other climate change impacts that are already known, for example, ecosystem responses to ocean warming and lower nutrient supply. Because these impacts have been extensively discussed in previous studies, we focus here on the global-scale implications of the insights from our work: the slowdown of silica dissolution under OA as revealed by our analysis of mesocosm and ocean sediment trap data. Thus, visualization and interpretation of results from the climate change simulations mostly refer to the net effect of OA-sensitive opal dissolution (Δ_{OA}), which is quantified as the difference between (a) the model including OA effects on opal dissolution and (b) the standard model configuration. Additional results are presented in Extended Data Fig. 5 and Extended Data Table 3.

Limitations of the global model

The model we used (UVic ESCM) is similar to other common models in terms of its ecosystem component, its skill in reproducing present-day conditions of biogeochemical quantities, and its behaviour in climate change simulations^{52,53}. Thus, it is reasonable to assume that the simulated OA impacts on opal dissolution and Si:N_{export} would yield similar results if incorporated into other global models. The driving mechanism, namely the slowdown of opal dissolution under OA, alters the vertical profile of particulate opal fluxes and regenerated Si(OH)₄. Because this is a chemical effect, it should be largely insensitive to specifics of the ecosystem model structure—instead, the most important factor is probably how the spatial distribution of diatoms and opal production are reproduced by different models. Our model shows good skill in reproducing observational data of Si(OH)₄ in the surface ocean (Extended Data Fig. 4), indicating that simulated spatial patterns of diatoms and opal production are reasonably realistic. However, as with most Earth system models, our simulations do not account for some potentially relevant mechanisms that may arise, for example, through ecological competition or complex food-web interactions, which may also yield potential repercussions for the global carbon cycle and will be discussed in the following.

In the model including OA-sensitive opal dissolution, a large part of the loss in diatom biomass is compensated by an increase in productivity of other phytoplankton groups. Accordingly, global primary productivity and carbon export remain largely unaffected by the OA effect on opal dissolution and Si:N_{export}, despite the sharp decline in diatoms. This is largely attributable to the degree of competition in the model, which depends on the choice of zooplankton prey selectivity and grazing formulations⁵⁴. By contrast, a recent modelling study focusing on functional diversity and ecological redundancy has demonstrated that changes in phytoplankton composition can entail knock-on effects on primary productivity, trophic transfer and carbon export³⁴. Accordingly, it is possible that the OA-driven loss of diatoms (owing to slower opal dissolution) could trigger additional ecological changes, which might in turn modify carbon cycling and export fluxes. However, the low degree of ecological detail in most Earth system models (including ours) does not enable an assessment of such complex knock-on effects.

Another important factor to consider is iron. It is well known that iron limitation enhances silicification of diatoms, increase cellular Si:N by up to two- or threefold^{55,56,57}. Thus, future changes in iron supply may cause shifts in the Si:N of particle flux from the surface ocean, which could theoretically counteract/enhance the consequences of OA-enhanced Si:N_{export} to some extent. In this context, the largely iron-limited Southern Ocean is of key interest, as iron supply might increase in the future, owing to increased aerosol dust deposition and melting ice^{45,58}. This would alleviate iron limitation and reduce the Si:N of diatoms, thereby possibly counteracting the OA-induced increase in Si:N_{export}. Our model includes a parameterization for iron-limitation effects on opal production by diatoms, thereby also controlling Si:N_{export}. However, as with most other Earth system models⁵⁹, it does not account for the complex mechanisms that may lead to future changes in iron deposition (for example, aeolian dust deposition, input from ice melting). Simulated iron inputs to the ocean are fixed at preindustrial rates. Thus, it is not possible to directly assess how future changes in iron limitation might interact with OA-driven changes in opal dissolution. However, in this context it is important to differentiate between effects on Si:N in the surface ocean (for example, iron effects on during production) and those occurring throughout the water column (for example, due to OA effects on dissolution): changes in the total amount of produced biogenic silica would alter the magnitude of Si flux from the euphotic zone, whereas the pH effect on Si dissolution alters flux attenuation throughout the water column. The relative size of the OA effect (that is, the proportional decrease in dissolution of sinking opal) is thus independent of the magnitude of Si flux from the surface. Accordingly, it can be assumed that future changes in dust deposition and Si:N (for example, in the Southern Ocean) would be superimposed by the OA-driven decrease in opal dissolution.

Furthermore, effects of opal ballasting on sinking velocities and remineralization rates of in diatom-derived organic matter are not accounted for in the model. In theory, the OA-driven decrease in opal dissolution may increase particle sinking velocities owing to enhanced (that is, prolonged) mineral ballasting by opal⁶⁰. At the same time, slower opal dissolution may enhance the protection of organic matter against remineralization to some extent⁶¹. On the basis of such considerations, it is often suggested that a

decrease in diatoms (such as the OA-driven loss of diatoms reported here) may reduce the efficiency of the biological carbon pump owing to slower sinking speed and/or faster remineralization of sinking organic matter of non-diatom origin¹. However, as the mechanisms on the particle scale described above are very complex and evidence on mineral ballasting in observational data is controversial, they are not included in our model (and neither in other, ecologically more complex, models)^{34,62}. Nevertheless, these mechanisms should be kept in mind when interpreting our results, as they could potentially have repercussions for the efficiency of the biological carbon pump.

Altogether, we note that our model results should be considered as an important estimate of how OA impacts on opal dissolution and Si:N_{export} may affect the marine silica cycle and the future biogeography of diatoms. As the OA-driven decrease in opal dissolution is a purely abiotic mechanism, we consider the main findings from our model simulations (decline in Si(OH)₄ and diatoms) to be robust and valid on large spatiotemporal scales, and expect them to be reproducible with other Earth system models. However, the model properties outlined above result in only minor knock-on effects of OA-driven changes on primary productivity and carbon export. Whether this holds true for the real ocean is uncertain and other Earth system models, for example, those with a higher degree of ecological complexity, may yield somewhat different results. For instance, although the OA effect on Si(OH)₄ and diatoms might be similar, possible responses of plankton community structure and biogeochemical processes may differ depending on the ecosystem model structure. Therefore, we hope that our study will be an incentive for the scientific community to explore OA effects on opal dissolution with different global models, and thereby assessing the potential variability of this effect among models.

Data availability

The data from the mesocosm experiments used for this study are provided as [Supplementary Data](#), and are also archived on the PANGAEA database (<https://doi.org/10.1594/PANGAEA.940756>). Furthermore, additional data

from the individual mesocosm experiments can be found on the PANGAEA database using the keyword ‘KOSMOS’.

References

1. Tréguer, P. et al. Influence of diatom diversity on the ocean biological carbon pump. *Nat. Geosci.* **11**, 27–37 (2018).
2. Tréguer, P. et al. The silica balance in the world ocean: a reestimate. *Science* **268**, 375–379 (1995).
3. Del Amo, Y. & Brzezinski, M. A. The chemical form of dissolved Si taken up by marine diatoms. *J. Phycol.* **35**, 1162–1170 (1999).
4. Dutkiewicz, S. et al. Impact of ocean acidification on the structure of future phytoplankton communities. *Nat. Clim. Change* **5**, 1002–1006 (2015).
5. Bach, L. T. & Taucher, J. CO₂ effects on diatoms: a synthesis of more than a decade of ocean acidification experiments with natural communities. *Ocean Sci.* **15**, 1159–1175 (2019).
6. Gao, K. S. & Campbell, D. A. Photophysiological responses of marine diatoms to elevated CO₂ and decreased pH: a review. *Funct. Plant Biol.* **41**, 449–459 (2014).
7. Moore, C. M. et al. Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**, 701–710 (2013).
8. Weber, T. S. & Deutsch, C. Ocean nutrient ratios governed by plankton biogeography. *Nature* **467**, 550–554 (2010).
9. Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. & Quéguiner, B. Production and dissolution of biogenic silica in the ocean. Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Glob. Biogeochem. Cycles* **9**, 359–372 (1995).

10. Sarmiento, J. L., Gruber, N., Brzezinski, M. A. & Dunne, J. P. High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature* **427**, 56–60 (2004). erratum **479**, 556 (2011).
11. Dugdale, R. C., Wilkerson, F. P. & Minas, H. J. The role of a silicate pump in driving new production. *Deep Sea Res. I* **42**, 697–719 (1995).
12. Volk, T. & Hoffert, M. I. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 99–110 (AGU, 1985).
13. Brzezinski, M. A. et al. A switch from Si(OH)₄ to NO₃[−] depletion in the glacial Southern Ocean. *Geophys. Res. Lett.* **29**, 5-1–5-4 (2002).
14. Holzer, M., Primeau, F. W., DeVries, T. & Matear, R. The Southern Ocean silicon trap: data-constrained estimates of regenerated silicic acid, trapping efficiencies, and global transport paths. *J. Geophys. Res. Oceans* **119**, 313–331 (2014).
15. IPCC. *Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2014).
16. Van Cappellen, P., Dixit, S. & van Beusekom, J. Biogenic silica dissolution in the oceans: reconciling experimental and field-based dissolution rates. *Glob. Biogeochem. Cycles* **16**, 23-21–23-10 (2002).
17. Lewin, J. C. The dissolution of silica from diatom walls. *Geochim. Cosmochim. Acta* **21**, 182–198 (1961).
18. Bidle, K. D. & Azam, F. Accelerated dissolution of diatom silica by marine bacterial assemblages. *Nature* **397**, 508–512 (1999).
19. Loucaides, S., Van Cappellen, P. & Behrends, T. Dissolution of biogenic silica from land to ocean: role of salinity and pH. *Limnol. Oceanogr.* **53**, 1614–1621 (2008).

20. Greenwood, J. E., Truesdale, V. W. & Rendell, A. R. Toward an understanding of biogenic-silica dissolution in seawater – an initial rate approach applied between 40 and 90 °C. *Aquat. Geochem.* **11**, 1–20 (2005).
21. Dove, P. M. & Elston, S. F. Dissolution kinetics of quartz in sodium chloride solutions: analysis of existing data and a rate model for 25 °C. *Geochim. Cosmochim. Acta* **56**, 4147–4156 (1992).
22. Gendron-Badou, A. C., Coradin, T., Maquet, J., Fröhlich, F. & Livage, J. Spectroscopic characterization of biogenic silica. *J. Non-Cryst. Solids* **316**, 331–337 (2003).
23. Roubeix, V., Becquevort, S. & Lancelot, C. Influence of bacteria and salinity on diatom biogenic silica dissolution in estuarine systems. *Biogeochemistry* **88**, 47–62 (2008).
24. Ragueneau, O., Schultes, S., Bidle, K., Claquin, P. & Moriceau, B. Si and C interactions in the world ocean: importance of ecological processes and implications for the role of diatoms in the biological pump. *Glob. Biogeochem. Cycles* **20**, GB4S02 (2006).
25. Nelson, D. M., DeMaster, D. J., Dunbar, R. B. & Smith, W. O. Jr. Cycling of organic carbon and biogenic silica in the Southern Ocean: estimates of water-column and sedimentary fluxes on the Ross Sea continental shelf. *J. Geophys. Res. Oceans* **101**, 18519–18532 (1996).
26. Kamatani, A. Dissolution rates of silica from diatoms decomposing at various temperatures. *Mar. Biol.* **68**, 91–96 (1982).
27. Meinshausen, M. et al. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Clim. Change* **109**, 213 (2011).
28. Bopp, L. et al. Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences* **10**, 6225–6245 (2013).
29. Kwiatkowski, L. et al. Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary

- production decline from CMIP6 model projections. *Biogeosciences* **17**, 3439–3470 (2020).
30. Primeau, F. W., Holzer, M. & DeVries, T. Southern Ocean nutrient trapping and the efficiency of the biological pump. *J. Geophys. Res. Oceans* **118**, 2547–2564 (2013).
 31. Kroeker, K. J., Kordas, R. L., Crim, R. N. & Singh, G. G. Meta-analysis reveals negative yet variable effects of ocean acidification on marine organisms. *Ecol. Lett.* **13**, 1419–1434 (2010).
 32. Boyd, P. W. & Brown, C. J. Modes of interactions between environmental drivers and marine biota. *Front. Mar. Sci.* **2**, 9 (2015).
 33. Gaylord, B. et al. Ocean acidification through the lens of ecological theory. *Ecology* **96**, 3–15 (2015).
 34. Dutkiewicz, S., Boyd, P. W. & Riebesell, U. Exploring biogeochemical and ecological redundancy in phytoplankton communities in the global ocean. *Global Change Biol.* **27**, 1196–1213 (2021).
 35. Mouw, C. B., Barnett, A., McKinley, G. A., Gloege, L. & Pilcher, D. Global ocean particulate organic carbon flux merged with satellite parameters. *Earth Syst. Sci. Data* **8**, 531–541 (2016).
 36. Riebesell, U. et al. A mobile sea-going mesocosm system – new opportunities for ocean change research. *Biogeosciences* **10**, 1835–1847 (2013).
 37. Boxhammer, T., Bach, L. T., Czerny, J. & Riebesell, U. Sampling and processing of mesocosm sediment trap material for quantitative biogeochemical analysis. *Biogeosciences* **13**, 2849–2858 (2016).
 38. Hansen, H. P. & Koroleff, F. in *Methods of Seawater Analysis* (eds Grasshoff, K. et al.) 159–228 (Wiley, 1999).
 39. Sharp, J. H. Improved analysis for “particulate” organic carbon and nitrogen from seawater. *Limnol. Oceanogr.* **19**, 984–989 (1974).

40. Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156 (1999).
41. Schartau, M., Landry, M. R. & Armstrong, R. A. Density estimation of plankton size spectra: a reanalysis of IronEx II data. *J. Plankton Res.* **32**, 1167–1184 (2010).
42. Ragueneau, O., Dittert, N., Pondaven, P., Tréguer, P. & Corrin, L. Si/C decoupling in the world ocean: is the Southern Ocean different? *Deep Sea Res. II* **49**, 3127–3154 (2002).
43. Olsen, A. et al. The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean. *Earth Syst. Sci. Data* **8**, 297–323 (2016).
44. Bidle, K. D., Manganelli, M. & Azam, F. Regulation of oceanic silicon and carbon preservation by temperature control on bacteria. *Science* **298**, 1980–1984 (2002).
45. IPCC. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (Cambridge Univ. Press, 2019).
46. Orr, J. C. et al. Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* **437**, 681–686 (2005).
47. Vrieling, E. G., Gieskes, W. W. C. & Beelen, T. P. M. Silicon deposition in diatoms: control by the pH inside the silicon deposition vesicle. *J. Phycol.* **35**, 548–559 (1999).
48. Martin-Jézéquel, V., Hildebrand, M. & Brzezinski, M. A. Silicon metabolism in diatoms: implications for growth. *J. Phycol.* **36**, 821–840 (2000).
49. Petrou, K. et al. Acidification diminishes diatom silica production in the Southern Ocean. *Nat. Clim. Change* **9**, 781–786 (2019).

50. Hervé, V. et al. Multiparametric analyses reveal the pH-dependence of silicon biomineralization in diatoms. *PLoS One* **7**, e46722 (2012).
51. Hutchins, D. A. & Fu, F. Microorganisms and ocean global change. *Nat. Microbiol.* **2**, 17058 (2017).
52. Kvale, K. et al. Explicit silicate cycling in the Kiel Marine Biogeochemistry Model, version 3 (KMBM3) embedded in the UVic ESCM version 2.9. *Geosci. Model Dev.* **14**, 7255–7285 (2021).
53. Keller, D. P., Oschlies, A. & Eby, M. A new marine ecosystem model for the University of Victoria Earth System Climate Model. *Geosci. Model Dev.* **5**, 1195–1220 (2012).
54. Prowe, A. E. F., Pahlow, M. & Oschlies, A. Controls on the diversity–productivity relationship in a marine ecosystem model. *Ecol. Modell.* **225**, 167–176 (2012).
55. Hutchins, D. A. & Bruland, K. W. Iron-limited diatom growth and Si:N uptake ratios in a coastal upwelling regime. *Nature* **393**, 561–564 (1998).
56. Takeda, S. Influence of iron availability on nutrient consumption ratio of diatoms in oceanic waters. *Nature* **393**, 774–777 (1998).
57. Brzezinski, M. A., Dickson, M.-L., Nelson, D. M. & Sambrotto, R. Ratios of Si, C and N uptake by microplankton in the Southern Ocean. *Deep Sea Res. II* **50**, 619–633 (2003).
58. Hutchins, D. A. & Boyd, P. W. Marine phytoplankton and the changing ocean iron cycle. *Nat. Clim. Change* **6**, 1072–1079 (2016).
59. Séférian, R. et al. Tracking improvement in simulated marine biogeochemistry between CMIP5 and CMIP6. *Curr. Clim. Change Rep.* **6**, 95–119 (2020).
60. Ploug, H., Iversen, M. H., Koski, M. & Buitenhuis, E. T. Production, oxygen respiration rates, and sinking velocity of copepod fecal pellets:

direct measurements of ballasting by opal and calcite. *Limnol. Oceanogr.* **53**, 469–476 (2008).

61. Moriceau, B. et al. Si–C interactions during degradation of the diatom *Skeletonema marinoi*. *Deep Sea Res. II* **56**, 1381–1395 (2009).
62. Aumont, O., Ethé, C., Tagliabue, A., Bopp, L. & Gehlen, M. PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies. *Geosci. Model Dev.* **8**, 2465–2513 (2015).
63. Taucher, J. et al. Changing carbon-to-nitrogen ratios of organic-matter export under ocean acidification. *Nat. Clim. Change* **11**, 52–57 (2021).
64. Garcia, H. E. et al. *World Ocean Atlas 2013. Volume 4: Dissolved Inorganic Nutrients (Phosphate, Nitrate, Silicate)* (National Oceanographic Data Center, 2013).
65. Schulz, K. G. et al. Temporal biomass dynamics of an Arctic plankton bloom in response to increasing levels of atmospheric carbon dioxide. *Biogeosciences* **10**, 161–180 (2013).
66. Riebesell, U. et al. Competitive fitness of a predominant pelagic calcifier impaired by ocean acidification. *Nat. Geosci.* **10**, 19–23 (2017).
67. Paul, A. J. et al. Effect of elevated CO₂ on organic matter pools and fluxes in a summer Baltic Sea plankton community. *Biogeosciences* **12**, 6181–6203 (2015).
68. Bach, L. T. et al. Influence of ocean acidification on a natural winter-to-summer plankton succession: first insights from a long-term mesocosm study draw attention to periods of low nutrient concentrations. *PLoS One* **11**, e0159068 (2016).
69. Taucher, J. et al. Influence of ocean acidification and deep water upwelling on oligotrophic plankton communities in the subtropical North Atlantic: insights from an in situ mesocosm study. *Front. Mar. Sci.* **4**, 85 (2017).

70. Moore, J. K. et al. Sustained climate warming drives declining marine biological productivity. *Science* **359**, 1139 (2018).
71. Marinov, I., Doney, S. C. & Lima, I. D. Response of ocean phytoplankton community structure to climate change over the 21st century: partitioning the effects of nutrients, temperature and light. *Biogeosciences* **7**, 3941–3959 (2010).
72. Dutkiewicz, S., Scott, J. R. & Follows, M. J. Winners and losers: ecological and biogeochemical changes in a warming ocean. *Glob. Biogeochem. Cycles* **27**, 463–477 (2013).
73. Laufkötter, C. et al. Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences* **12**, 6955–6984 (2015).

Acknowledgements

This study was supported by the German Federal Ministry of Science and Education (BMBF) in the framework of the projects BIOACID III ('Biological Impacts of Ocean Acidification', FKZ 03F0728) and SOPRAN III ('Surface Ocean Processes in the Anthropocene', FKZ 03F0662). Funding for the different mesocosm studies was also provided by the 'European Project on Ocean Acidification' (EPOCA, grant no. 211384) from the European Community's Seventh Framework Programme (FP7/2007–2013), the EU project MESOAQUA (grant no. 228224), as well as BIOACID II (FKZ 03F06550) and SOPRAN II (03F0611). Furthermore, we thank all participating scientists and technicians for their efforts in realizing the different studies. We also thank the staff of the marine biological stations in the different study locations for providing logistics, technical assistance and support at all times. We thank the captains and crews of RV *Viking Explorer*, MV *Esperanza* of Greenpeace, RV *Håkon Mosby* (2011609), RV *Alkor* (AL376, AL394, AL397, AL406, AL420), RV *Heincke* (HE360), RV *Poseidon* (POS463), and RV *Hesperides* (29HE20140924) for support during transport, deployment and recovery of the mesocosm facilities.

Funding

Open access funding provided by GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel.

Author information

Authors and Affiliations

1. GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

Jan Taucher, A. E. Friederike Prowe, Tim Boxhammer, Karin Kvale & Ulf Riebesell

2. Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

Lennart T. Bach

3. GNS Science, Lower Hutt, New Zealand

Karin Kvale

Contributions

This study was conceived by J.T. Mesocosm experiments were coordinated and implemented by U.R. with support and data acquisition by T.B., L.T.B. and J.T. Analysis of global sediment trap data was conducted by J.T. Model implementation was realized by K.K., A.E.F.P. and J.T. Analysis and visualization of model results was done by J.T. The manuscript was written by J.T. with contributions from all co-authors.

Corresponding author

Correspondence to [Jan Taucher](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks David Hutchins and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. [Peer reviewer reports](#) are available.

Additional information

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

Extended data figures and tables

[Extended Data Fig. 1 Mesocosm experiments on OA.](#)

a, Locations of in situ mesocosm experiments with the ‘Kiel Off-Shore Mesocosms for Ocean Simulations’ (KOSMOS). **b**, Schematic drawing of a pelagic mesocosm enclosing the natural plankton community and collecting sinking organic matter in a full-diameter sediment trap at the bottom (15 to 25 m water depth, depending on the study site). Different OA scenarios were achieved by adding filtered, CO₂-saturated seawater equally distributed into the mesocosms³⁶ at the beginning of the experiment, and usually several more times throughout the study period to maintain CO_2 levels within target levels. Illustration by R. Erven and T.B. (GEOMAR), reprinted with permission from the AGU and Springer Nature, ref. ⁶³.

[Extended Data Fig. 2 Comparison of OA impacts on different elemental ratios \(Si:N, C:N and Si:C\) of sinking biogenic matter.](#)

Note that effects on Si:C display a larger variability than Si:N owing to additional (but variable) responses of C:N ratios. In contrast to the OA effect on Si (slower chemical dissolution), the shifts in C:N are biotically driven, resulting from responses of primary producers and heterotrophic consumers, which together leave an imprint on the C:N sinking out of the productive surface layer⁶³. Depending on the direction and magnitude of the OA effect on C:N, this offsets or amplifies the effect on Si:C compared to Si:N.

However, because OA effects on C:N are driven by biotic processes and linked to freshly produced material, it can be assumed that this effect is mostly restricted to the surface ocean, whereas the OA effect on Si occurs throughout the entire water column (see Extended Data Fig. 3). It can thus be expected that the OA-driven decrease in opal dissolution will enhance the preservation of Si compared to C, thereby increase Si:C ratios of particles sinking through the water column.

Extended Data Fig. 3 Comparison of warming and OA impacts on silica dissolution.

One-dimensional representation of warming and OA effects on silica dissolution and export to depth in our model, using globally averaged vertical profiles of temperature (left) and pH (middle) under preindustrial conditions and for the year 2200 (values from the RCP8.5 simulation). The right panel depicts the model parameterization for opal dissolution, applied to the corresponding temperature and pH profiles. Averaged over the upper 1,000 m, the specific opal dissolution rate ($\text{Diss}_{\text{opal}}$) under preindustrial conditions amounts to 0.0008 (in units of m^{-1}), that is, 0.08% of sinking opal is dissolved per metre, resulting in a characteristic ‘Martin curve’ of opal flux attenuation with depth (black line). $\text{Diss}_{\text{opal}}$ in the model is temperature-dependent and roughly doubles for an increase in temperature of 10 °C. Warming of ~3 °C in the upper ocean until the year 2200 (averaged over the upper 1,000 m) increases $\text{Diss}_{\text{opal}}$ by about 1.3-fold (from 0.08% to 0.11%), thereby leading to a stronger flux attenuation with depth (red dashed line) compared to preindustrial conditions. At the same time, pH decreases by around 0.6 units until year 2200 (averaged over the upper 1,000 m). According to our findings, opal dissolution rates are 57% lower per unit decrease in pH, (that is, a decrease by a factor of ~2.5). This effect of pH

works in the opposite direction as the temperature effect, that is, slowing down opal dissolution and thus weakening the decline of opal with depth. When considering this pH effect in isolation in our example (that is, not accounting for the effects of warming; dashed blue line), this drop in pH would decrease $\text{Diss}_{\text{opal}}$ from 0.08% to 0.045% (that is, reducing it by a factor of 1.8). When considering both warming and acidification at the same time, the pH effect is still strong enough to over-compensate the temperature effect. $\text{Diss}_{\text{opal}}$ amounts to 0.059%, which is notably lower than under preindustrial (0.08%) and warming-only (0.11%) conditions. Accordingly, opal dissolution in a warmer and lower-pH ocean is slower than under preindustrial conditions, thereby leading to an enhanced efficiency of silica export to the deep ocean (green line).

Extended Data Fig. 4 Model versus data for silicic acid in the surface ocean.

a, b, Present-day distribution of sea surface $\text{Si}(\text{OH})_4$ from the World Ocean Atlas⁶⁴ (**a**) and simulated by our model (**b**). **c**, Simulated distribution of diatom biomass.

Extended Data Fig. 5 Simulated OA impacts on silicic acid and diatoms.

a, b, Future changes in $\text{Si}(\text{OH})_4$ (**a**) and diatom biomass (**b**) in the surface ocean relative to preindustrial conditions. Shown are results from simulations with the standard model configuration, that is, excluding OA-effects on silica dissolution (dashed lines), and the model including OA-effects on silica dissolution (solid lines) for RCP6.0 and RCP8.5 emission scenarios. See [Methods](#) and Extended Data Table 3 for more details.

Extended Data Table 1 Overview of study locations, experimental design and environmental conditions of the different mesocosm experiments

Extended Data Table 2 Mass fluxes of Si, N and C and their elemental ratios

Extended Data Table 3 Summary of results from model simulations

Supplementary information

Supplementary Data

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Taucher, J., Bach, L.T., Prowe, A.E.F. *et al.* Enhanced silica export in a future ocean triggers global diatom decline. *Nature* **605**, 696–700 (2022). <https://doi.org/10.1038/s41586-022-04687-0>

- Received: 06 May 2021
- Accepted: 24 March 2022
- Published: 25 May 2022

- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04687-0>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

[Sinking diatoms trap silicon in deep seawater of acidified oceans](#)

- David A. Hutchins

Nature News & Views 25 May 2022

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04687-0>

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

[Download PDF](#)

- Article
- Open Access
- [Published: 11 May 2022](#)

Extricating human tumour immune alterations from tissue inflammation

- [Florian Mair](#) [ORCID: orcid.org/0000-0001-6732-5449](#)^{1 na1},
- [Jami R. Erickson](#) [ORCID: orcid.org/0000-0002-4549-5217](#)^{1,2 na1},
- [Marie Frutoso](#) [ORCID: orcid.org/0000-0002-7848-1049](#)¹,
- [Andrew J. Konecny](#)^{1,2},
- [Evan Greene](#)¹,
- [Valentin Voillet](#)^{1,3},
- [Nicholas J. Maurice](#) [ORCID: orcid.org/0000-0002-2667-472X](#)^{1,4},
- [Anthony Rongvaux](#) [ORCID: orcid.org/0000-0001-7142-6533](#)^{2,5},
- [Douglas Dixon](#) [ORCID: orcid.org/0000-0003-0911-3717](#)^{6 nAff10},
- [Brittany Barber](#)⁷,
- [Raphael Gottardo](#)^{1,8 nAff11} &
- [Martin Prlic](#) [ORCID: orcid.org/0000-0002-0685-9321](#)^{1,9}

[Nature](#) volume **605**, pages 728–735 (2022)

- 20k Accesses
- 162 Altmetric
- [Metrics details](#)

Subjects

- [Adaptive immunity](#)
- [Immunotherapy](#)
- [Inflammation](#)
- [Tumour immunology](#)

Abstract

Immunotherapies have achieved remarkable successes in the treatment of cancer, but major challenges remain^{1,2}. An inherent weakness of current treatment approaches is that therapeutically targeted pathways are not restricted to tumours, but are also found in other tissue microenvironments, complicating treatment^{3,4}. Despite great efforts to define inflammatory processes in the tumour microenvironment, the understanding of tumour-unique immune alterations is limited by a knowledge gap regarding the immune cell populations in inflamed human tissues. Here, in an effort to identify such tumour-enriched immune alterations, we used complementary single-cell analysis approaches to interrogate the immune infiltrate in human head and neck squamous cell carcinomas and site-matched non-malignant, inflamed tissues. Our analysis revealed a large overlap in the composition and phenotype of immune cells in tumour and inflamed tissues.

Computational analysis identified tumour-enriched immune cell interactions, one of which yields a large population of regulatory T (T_{reg}) cells that is highly enriched in the tumour and uniquely identified among all haematopoietically-derived cells in blood and tissue by co-expression of ICOS and IL-1 receptor type 1 (IL1R1). We provide evidence that these intratumoural $IL1R1^+$ T_{reg} cells had responded to antigen recently and demonstrate that they are clonally expanded with superior suppressive function compared with $IL1R1^-$ T_{reg} cells. In addition to identifying extensive immunological congruence between inflamed tissues and tumours as well as tumour-specific changes with direct disease relevance, our work also provides a blueprint for extricating disease-specific changes from general inflammation-associated patterns.

Main

Antigen-presenting cells (APCs) and T cells residing in non-lymphoid tissues adapt distinct phenotypic and functional properties relative to their circulating counterparts in the peripheral blood^{5,6,7}. These immune cells are also present in many solid tumour types, where they are thought to be critical determinants of tumour development and disease outcome^{1,8}. One hallmark of immune-infiltrated human tumour tissues is the presence of an inflammatory microenvironment—this has been extensively scrutinized during the past decade^{3,9}. However, since there is a paucity of studies on human non-malignant, inflamed tissues, it remains unclear which immune cell subsets and signalling pathways in the human tumour microenvironment are distinct from general inflammatory processes.

One of the best studied immune populations in tumour tissues is functionally exhausted (dysfunctional) T cells and T_{reg} cells, both of which are considered pivotal for inefficient anti-tumour immune responses^{10,11}. These T cell subsets express immuno-inhibitory molecules such as programmed death 1 (PD-1) or cytotoxic T-lymphocyte-associated antigen 4 (CTLA4), which are the targets of various immunotherapeutic approaches¹². However, expression of PD-1 and CTLA4 is not limited to tumour-infiltrating T cells and is also found on T cells in non-malignant tissues during homeostasis and inflammation^{6,13}.

Notably, the effector program of T cells and their expression of immuno-regulatory molecules is closely linked to the function of (APCs), including dendritic cells, macrophages and other monocyte-derived cells¹⁴. APCs integrate tissue-specific and inflammation-dependent cues from the tissue environment, and can enhance or suppress local T cell responses¹⁵. Thus, functional alteration of APCs in the human tumour microenvironment has been suggested as an additional promising therapeutic target^{16,17}.

We hypothesized that comparing the human tumour microenvironment with non-malignant, inflamed tissues could identify tumour-unique immune alterations that are distinct from general inflammatory responses. We thus combined several single-cell analysis pipelines to generate a comprehensive immune landscape of human head and neck squamous cell carcinoma (HNSCC) with site-matched non-malignant inflamed tissues from the oral cavity. Our data revealed substantial congruence of the immune phenotypes

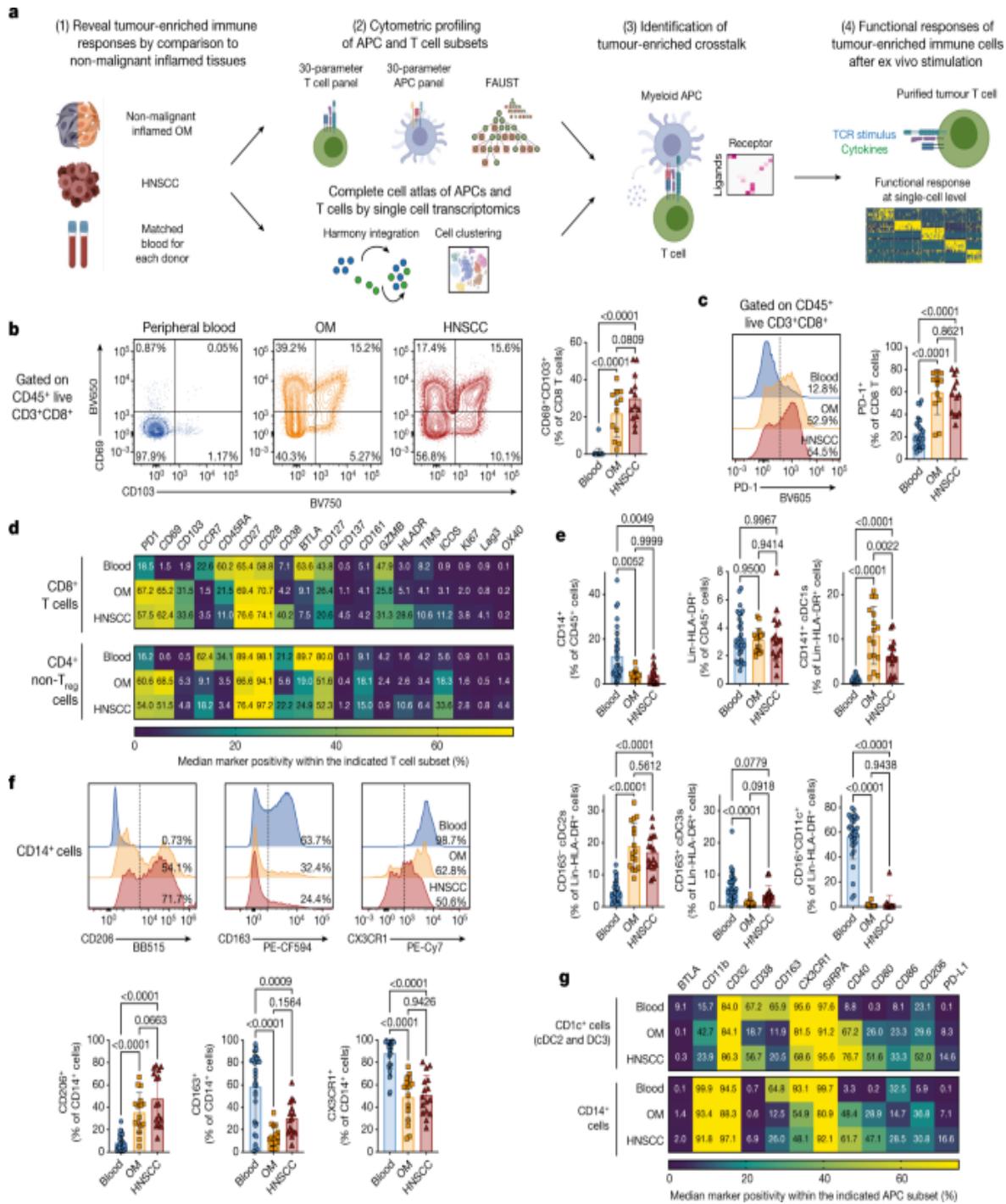
between these tissue groups. Computational analysis approaches identified tumour-specific changes in subsets of activated APCs and T_{reg} cells, including predicted major histocompatibility complex (MHC)–T cell receptor (TCR) and IL-1–IL1R signalling. Follow-up experiments confirmed these computational predictions: IL1R1⁺ T_{reg} cells in the tumour showed substantial clonal expansion, superior immunosuppressive function and hallmarks of recent TCR stimulation. Finally, this T_{reg} population could be identified among all haematopoietic cells by the combined expression of IL1R1 and ICOS, thus providing a unique opportunity to specifically target a large population of intratumoural T_{reg} cells.

Phenotypic congruence of OM and HNSCC

Surgery is typically the first line of treatment for HNSCC, of which oral and oropharyngeal squamous cell carcinoma are subsites¹⁸. Non-malignant inflamed oral mucosal (OM) tissues (typically without prior anti-inflammatory treatment) from oral surgeries served as our reference. Together, this enabled us to compare the immune infiltrate of human inflamed with that of tumour tissues without therapeutic interventions as a confounding variable (sample list in Supplementary Table 1).

First, we extensively catalogued the immune landscape in OM and HNSCC tissues and matched blood by using 2 flow cytometry panels comprising 30 parameters (Fig. 1a, Supplementary Tables 2, 3) (adapted from ref. 19). The frequency of CD3⁺ T cells, CD19⁺ B cells and CD56⁺ natural killer (NK) cells among total CD45⁺ live cells as well as the CD4/CD8 ratio was essentially equivalent between OM and HNSCC tissues (Extended Data Fig. 1a, b). However, we observed a significant increase of CD4⁺CD25⁺CD127⁻Foxp3⁺ T_{reg} cells in HNSCC compared with OM tissues²⁰ (Extended Data Fig. 1c, d).

Fig. 1: Similar immune phenotypes in inflamed non-malignant OM tissues and HNSCC.



(top) and CD4⁺ helper T cells (without CD25⁺CD127⁻ T_{reg} cells, bottom) across peripheral blood, OM and HNSCC. **e**, Quantification of the indicated antigen-presenting cell (APC) populations. **f**, Representative histograms and quantification for CD206, CD163 and CX3CR1 on CD14⁺ cells. **g**, Heat map representing the expression pattern for all the indicated molecules within CD1c⁺ cDC2s and cDC3s (top) and CD14⁺ cells (bottom). All summary graphs are represented as mean ± s.d. ($n = 12$ for OM and $n = 13$ for HNSCC samples for T cell data, $n = 16$ for OM and HNSCC for APC data). Statistical analyses were performed using one-way ANOVA with Tukey's multiple comparisons test.

Recent findings suggest that cytotoxic CD8⁺ T cells with a tissue-resident memory phenotype can be a principal predictor of tumour progression^{8,21,22}. The expression patterns of the tissue residency markers CD69 and CD103 were very similar between OM and HNSCC tissues (Fig. [1b](#)). PD-1, a biomarker of exhausted T cells²³ was expressed by approximately 50% of total CD8⁺ T cells in both OM and HNSCC tissue samples (Fig. [1c](#)). The transcription factors TCF1 and TOX, as well as CD39 showed similar expression patterns in OM- and HNSCC-infiltrating CD8⁺ T cells (Extended Data Fig. [1e,f](#)). The same was true for the majority of markers for CD4⁺ and CD8⁺ T cells across both tissues (Fig. [1d](#), Extended Data Fig. [1g,h](#)).

Next, we quantified subsets of APCs in the tumour microenvironment: CD14⁺ monocyte/macrophage-like cells, CD11c⁺CD141⁺ cross-presenting type 1 classical dendritic cells (cDC1s), CD11c⁺CD1c⁺CD163⁻ cDC2s, CD11c⁺CD1c⁺CD163⁺ DC3s (previously referred to as inflammatory dendritic cells^{24,25,26}) and CD16⁺ non-conventional monocytes (Fig. [1e](#), Extended Data Fig. [1i](#)). Whereas the relative abundance of CD14⁺ cells and total CD14⁻CD3⁻CD19⁻ (hereafter referred to as Lin⁻) HLADR⁺ cells was indistinguishable between OM and HNSCC tissues, we noted a slight decrease in the frequency of CD141⁺ cDC1s in HNSCC²⁷ (Fig. [1e](#)). cDC2s, DC3s and CD16⁺ cells were present in OM and HNSCC tissues with similar frequencies (Fig. [1e](#)). Of note, we observed similar expression patterns for the mannose receptor CD206 (commonly used as a marker for alternatively activated, M2 macrophages) on CD14⁺ cells across the different tissues (Fig. [1f](#)), indicating that M2-like phenotypes are not a specific hallmark of the

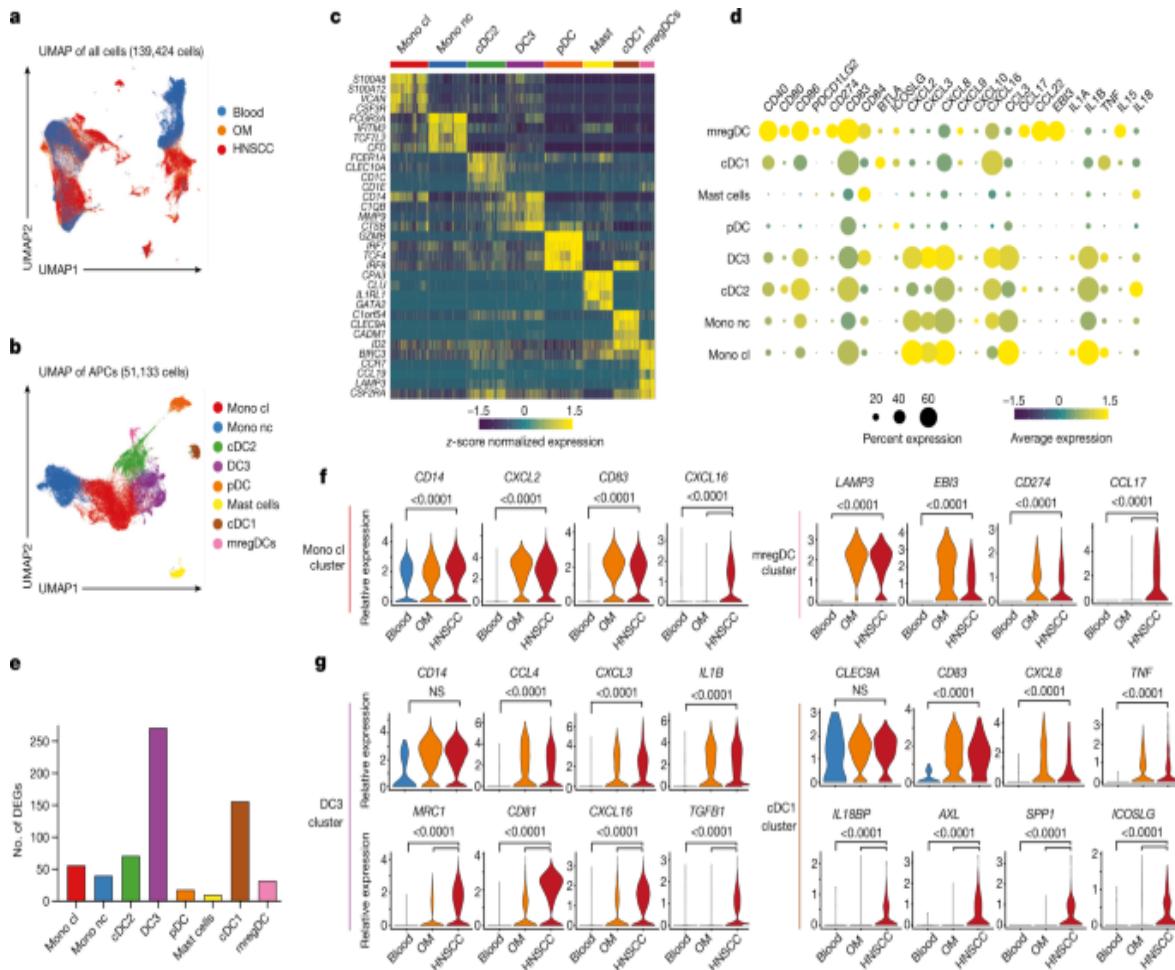
tumour microenvironment²⁸. A comparison of all biomarkers showed that tissue-infiltrating APC subsets were phenotypically relatively similar between OM and HNSCC tissues (Fig. 1f,g, Extended Data Fig. 1k–m). Together, these data indicate that the immune infiltrates in inflamed OM and HNSCC show substantial congruence in composition and phenotype.

To identify specific immune subset differences, we performed computational analysis using full annotation using shape-constrained trees (FAUST)²⁹, a machine learning algorithm that discovers and annotates statistically relevant cellular phenotypes in an unsupervised manner. For the T cell panel, FAUST identified a single subcluster of CD8⁺ T cells, and four CD4⁺ T_{reg} (CD25⁺ CD127⁻) phenotypes marked by expression of ICOS with combinations of PD-1, TIM3 and HLA-DR as being enriched in HNSCC (Extended Data Fig. 2a–c). For the APC panel, FAUST identified several tumour-enriched phenotypes of CD14⁺ cells, cDC2s and cDC3s marked by co-expression of CD40 and PD-L1 (Extended Data Fig. 2d,e,g), resembling an activated APC phenotype. To further interrogate congruencies and differences in these immune phenotypes, we next used a single-cell RNA-seq (scRNA-seq) approach.

Tumour-enriched cytokine modules in APCs

To ensure analysis of rare T cell and APC subsets, we sorted pan CD3⁺ T cells and Lin⁻HLA-DR⁺ cells from fresh OM and HNSCC tissues with matched blood (Extended Data Fig. 3a–c). After filtering for quality control and data integration using Harmony³⁰, we obtained approximately 140,000 cells from 8 donors (Extended Data Fig. 3d). After dimensionality reduction using uniform manifold approximation and projection³¹ (UMAP) and cellular annotation using SingleR³², canonical T cell and APC populations separated clearly on the UMAP plot (Extended Data Fig. 3e). Cells derived from OM and HNSCC mostly grouped together, but separate from peripheral blood (Fig. 2a), in line with the phenotypic overlap found in our flow cytometry data.

Fig. 2: Comprehensive scRNA-seq analysis of OM and HNSCC immune infiltrates.



a, UMAP of the combined scRNA-seq data after quality filtering and Harmony integration, coloured by tissue origin (more details in Extended Data Fig. 3). **b**, UMAP plot of the APC populations after subsetting and reclustering, coloured by cluster. Mono cl, classical monocyte; Mono nc, non-classical monocyte. **c**, Key DEGs in each APC cluster. **d**, Scaled dot plot showing the transcript expression across APC clusters from combined OM and HNSCC data (excluding blood). **e**, Number of DEGs between HNSCC and OM-derived cells per APC cluster as determined by MAST. **f**, Violin plots showing the expression of selected transcripts for the monocyte cluster (left) and the mregDC cluster (right). **g**, Violin plots showing the expression of selected transcripts for the DC3 cluster (left) and the cDC1 cluster (right). All graphs are showing combined data for $n = 4$ for OM samples and $n = 4$ for HNSCC samples, with a total of 139,424 cells after filtering for quality control criteria. Violin plots show adjusted P -values

(Bonferroni correction) as calculated by the Seurat implementation of MAST.

Next, we re-clustered APCs and mapped these populations to established lineages (Fig. 2b, Extended Data Fig. 4a). We also found a population of HLA-DR-expressing mast cells mainly in HNSCC tissues, marked by the signature genes *CLU* (encoding mast cell carboxypeptidase A) and *GATA2* (Fig. 2c, Extended Data Fig. 4a). In line with our flow cytometry data, cDC1s were reduced in HNSCC tissues, while the other APC subsets appeared similarly distributed in both tissues across all donors (Extended Data Fig. 4b,c), including a population expressing high levels of *CCR7*, *CCL19* and *CSF2RA* (encoding the GM-CSF receptor) (Fig. 2c), resembling mature dendritic cells enriched in immunoregulatory molecules³³ (mregDCs). When we analysed the abundance of transcripts of key co-regulatory genes, chemokines and cytokines across all APC subsets from OM and HNSCC (Fig. 2d), we noted that shared modules of chemokine transcripts were detected in the monocyte, cDC2 and DC3 clusters, whereas mregDCs were the dominant CCL17- and CCL22-expressing population. CXCL8 and CXCL16 expression was detected in all subsets, albeit to varying degrees.

To determine whether the transcriptional activity of these APC clusters changed in HNSCC relative to the inflamed OM, we identified differentially expressed genes (DEGs) in each cluster using model-based analysis of single-cell transcriptomics³⁴ (MAST). MAST revealed a pronounced alteration of DC3 and cDC1 transcriptomes in HNSCC compared with OM (that is, 150–250 genes), whereas the remaining APC clusters showed more congruent profiles (Fig. 2e). Manually selected genes that were either shared between tissues or enriched in HNSCC are shown in Fig. 2f,g, and heat maps are shown in Extended Data Fig. 4d,e. Overall, this comparison highlights that altered transcriptional signatures in HNSCC tissues are detected mainly in cDC1 and DC3 subsets and include differential expression of gene encoding cytokines (*TGFB1* and *IL18BP*) and co-receptors (*ICOSLG*).

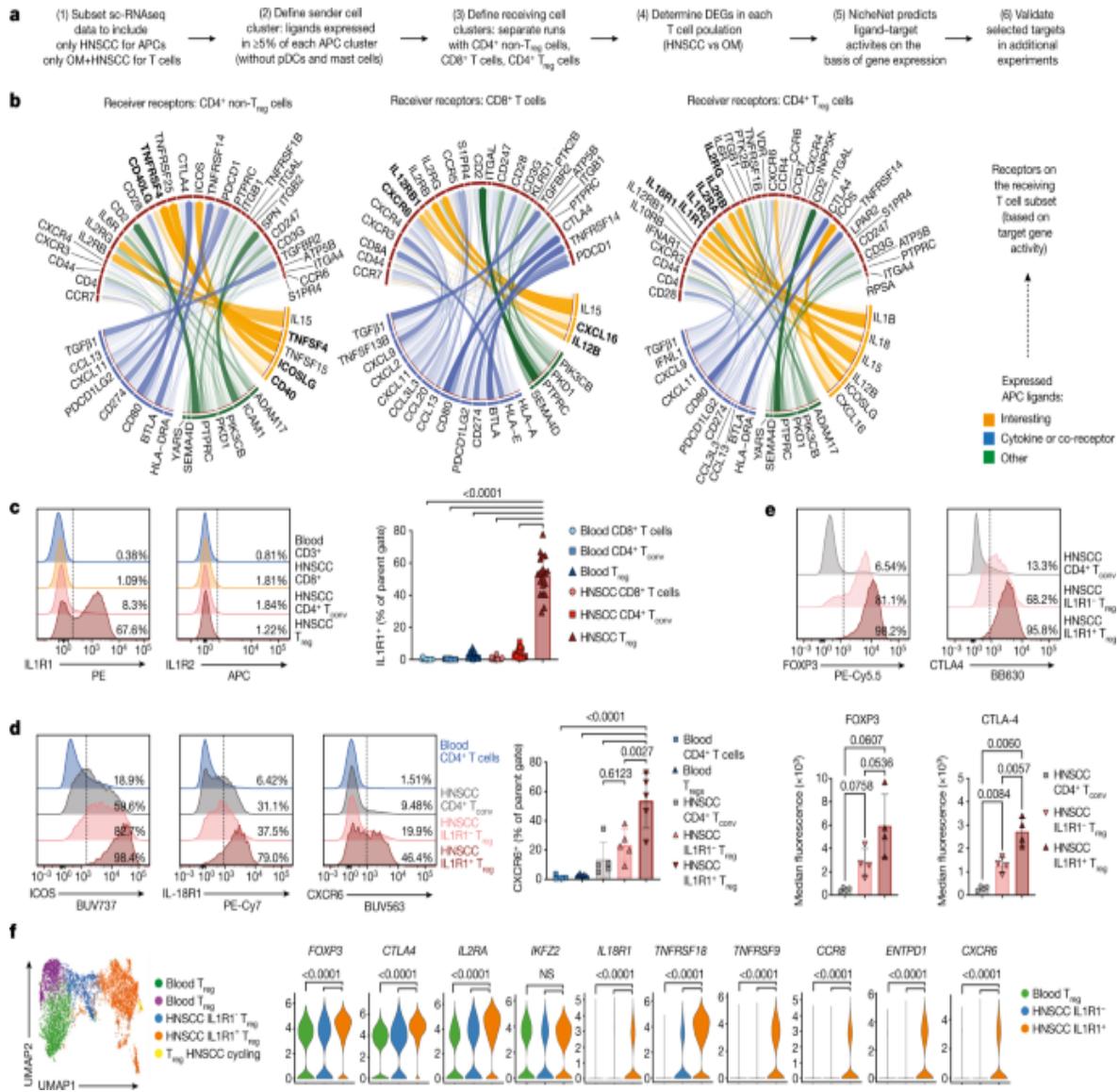
Finally, re-clustering of the T cells suggested that there were eight transcriptionally separate populations, which we annotated manually (Extended Data Fig. 5a). Again, cluster distribution across donors and

tissues was remarkably constant, with the HNSCC samples showing an expansion of the CD4⁺ T_{reg} cluster (Extended Data Fig. 5b). We noted that the number of DEGs between HNSCC and OM-infiltrating cells showed the largest change in two CD8⁺ T cell clusters and in the T_{reg} cluster (Extended Data Fig. 5c), with the other clusters being rather similar (Extended Data Fig. 5d). Of note, the DEGs in these two CD8 T cell clusters were associated with cytotoxic properties (Extended Data Fig. 5e). Calculating T cell lineage scores for the T cell clusters further highlighted the increase in the T_{reg} signature (Extended Data Fig. 5f,g), in line with our flow cytometry data.

Tumour-enriched APC-T cell crosstalk

To better understand the relevance of the immune alterations that we had observed, we investigated potential tumour-specific cross-talk between T cells and APCs using NicheNet³⁵ (workflow in Fig. 3a). We set the APC clusters derived from scRNA-seq (excluding plasmacytoid dendritic cells (pDCs) and mast cells) as the sender population, and the CD4⁺ conventional T cell, CD8⁺ T cell and CD4⁺ T_{reg} clusters as separate receiver populations. For each T cell subset, we focused our analysis on the top 20 ligand–receptor pairs (Fig. 3b, Extended Data Fig. 6a). Several signalling axes were shared by T cell subsets in the tumour microenvironment (for example, the ligands CD80, CD274 and PDCD1LG2, and the cytokines TGFβ1 and IL-15). NicheNet predicted several enriched ligand–receptor interactions between APCs and T_{reg} cells in HNSCC: ICOS ligand (ICOSLG) via ICOS, IL-18 via the IL18-R1 and IL-1B via the IL-1 receptors type 1 and type 2 (Fig. 3b, right). Of note, when we assessed which APC population expressed transcripts for these ligands, some were found in multiple subsets in both OM and HNSCC (for example, *TGFB1*, *IL1B* and *CXCL16*), whereas *ICOSLG* was found only on cDC1s from the HNSCC (Fig. 2g, Extended Data Fig. 6b).

Fig. 3: NicheNet analysis predicts tumour-enriched APC-T cell crosstalk.



a, The NicheNet workflow was applied to the scRNA-seq data shown in Fig. 2. **b**, Circos plots showing the top ligand–receptor pairs identified by NicheNet. Transparency of the connection represents the interaction strength. APC ligands are on the bottom, TCRs are on top. **c**, Representative plots and quantification for the surface protein expression of IL1R1 ($n = 19$). **d**, Representative plots showing the expression of ICOS, IL-18R1 and the chemokine receptor CXCR6 on the indicated T cell subsets. Right, quantification for CXCR6 ($n = 5$). **e**, Representative plots (top) and mean fluorescence intensity (bottom) for FOXP3 and CTLA4 on T cell subsets from HNSCC ($n = 4$). **f**, UMAP plot of T_{reg} cells sorted from blood and tumour of $n = 3$ donors with HNSCC after targeted transcriptomics, coloured

by cluster. Violin plots show the expression of selected transcripts across T_{reg} clusters. All summary graphs are represented as mean ± s.d. Statistical analyses of cytometry data was performed using one-way ANOVA with Tukey's multiple comparisons test, analysis of targeted transcriptomics used the Seurat implementation of MAST (adjusted *P*-values after Bonferroni correction).

Since our flow cytometry and scRNA-seq approaches both indicated T_{reg} changes in the HNSCC, we further explored the predicted T_{reg}-APC interactions. We first tested whether the receptors predicted by NicheNet were present as proteins, and found that IL1R1 was expressed specifically by tumour-infiltrating T_{reg} cells, but not by tumour-infiltrating CD4⁺ T cells or CD8⁺ T cells, nor by T cells in the peripheral blood (Fig. 3c). Notably, up to 70% of the T_{reg} cells expressed IL1R1, whereas IL1R2, a decoy receptor for IL-1 signalling, was detectable on less than 2% of cells. Nearly all IL1R1⁺ T_{reg} cells co-expressed ICOS and IL-18R1 in conjunction with higher levels of the chemokine receptor CXCR6 (Fig. 3d), matching the NicheNet predictions. Finally, both IL1R1⁺ and IL1R1⁻ intratumoural T_{reg} cells expressed FOXP3, but CTLA4 expression was significantly increased only in the IL1R1⁺ fraction (Fig. 3e).

We next addressed whether intratumoural APCs had the capacity to express IL-1. Following ex vivo culture, a majority of HNSCC-derived CD14⁺ cells and some pan cDCs expressed IL-1 β and IL-1 α protein (Extended Data Fig. 6c). Analysis of flash-frozen whole-tumour lysate revealed substantial levels of IL-1 α , IL-1 β and IL-18 (Extended Data Fig. 6d). Together, these data indicate that intratumoural IL-1 is probably available for IL1R1⁺ T_{reg} cells, but also raise the question of the biological and clinical relevance of the IL1R1⁺ T_{reg} population.

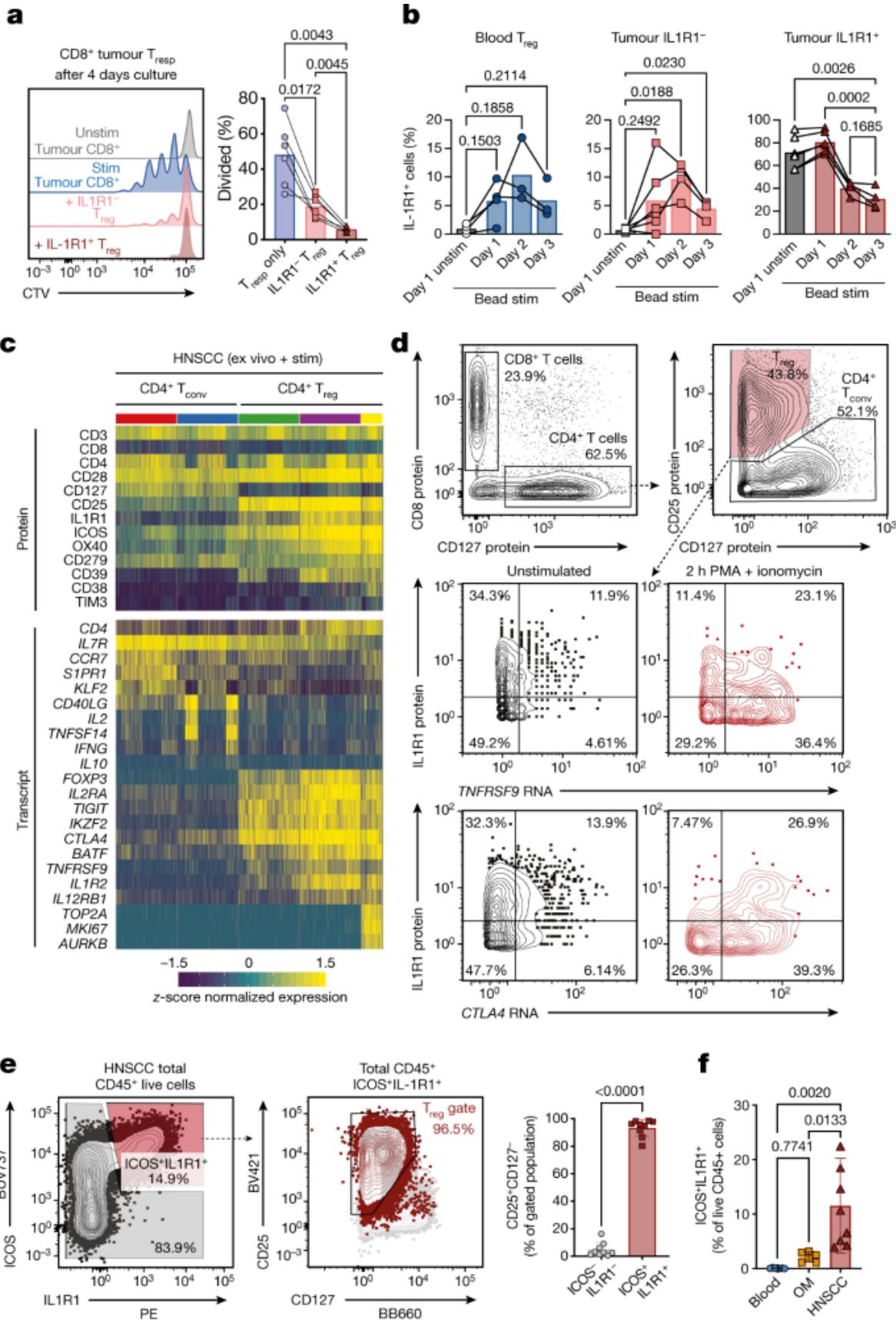
To determine how IL1R1⁺ T_{reg} cells differ from their IL1R1⁻ counterparts, we used a targeted transcriptomics approach³⁶ to measure the expression of genes with known relevance for immune function (Supplementary Table 4). Intratumoural IL1R1⁺ and IL1R1⁻ T_{reg} cells (from *n* = 3 tumours) formed separate clusters (Fig. 3f, orange and blue, respectively), distinct from blood

T_{reg} cells. More than 50 transcripts were selectively enriched in the IL1R1⁺ cluster, including *TNFRSF18* (which encodes GITR) and *TNFRSF9* (which encodes 4-1BB), which has been suggested as a pan-cancer T_{reg} target³⁷ (Fig. 3f). Overall, these data suggest that IL1R1⁺ T_{reg} cells represent a transcriptionally distinct population of intratumoural T_{reg} cells.

IL1R1⁺ T_{reg} cells are highly suppressive

To directly assess the functional capacity of intratumoural IL1R1⁺ T_{reg} cells, we established classic suppression assays suitable for the low cell numbers from HNSCC tissues. We found that IL1R1⁺ T_{reg} cells were more effective than their IL1R1⁻ counterparts at suppressing proliferation of CD8⁺ (Fig. 4a) and CD4⁺ responder T (T_{resp}) cells isolated from tumours as well as from peripheral blood (Extended Data Fig. 7a). Furthermore, we observed a decrease in the concentration of effector molecules in the culture supernatant in the presence of IL1R1⁺ T_{reg} cells (Extended Data Fig. 7b), and suppression was dependent on the ratio of T_{reg} cells to T_{resp} cells (Extended Data Fig. 7c).

Fig. 4: IL1R1-expressing T_{reg} cells represent a functionally distinct population.



a, Proliferation of HNSCC-derived CD8⁺ T responder (T_{resp}) cells ($n = 6$) in an in vitro suppression assay with IL1R1⁻ T_{reg} cells (light red) and IL1R1⁺ T_{reg} cells (dark red). Representative histograms show dilution of Cell Trace Violet. Stim, stimulated; unstim, unstimulated. **b**, Expression kinetics of IL1R1 after in vitro culture in the presence of anti-CD3/CD28/CD2 beads for T_{reg} cells sorted from peripheral blood (left, $n = 3$), IL1R1⁻ T_{reg} cells (middle, $n = 4$) and IL1R1⁺ T_{reg} cells (right, $n = 5$) from HNSCC. **c**, Tumour-infiltrating T cells from two donors with HNSCC after performing short-term stimulation and targeted transcriptomics with AbSeq (Extended Data Fig. 8). Heat maps show top differentially expressed proteins (top) and transcripts (bottom) across the selected clusters. **d**, TNFRSF9 and CTLA4 transcript expression by T_{reg} cells left unstimulated (left) and after short-term stimulation with PMA and ionomycin (right). The y-axis shows IL1R1 protein expression. **e**, Representative plots and quantification ($n = 9$) showing that within total CD45⁺ cells in HNSCC nearly all ICOS⁺ IL1R1⁺ cells are T_{reg} cells. **f**, Quantification of total ICOS⁺ IL1R1⁺ cells in peripheral blood ($n = 7$), OM ($n = 6$) and HNSCC samples ($n = 8$). All summary graphs are represented as mean \pm s.d. Statistical analyses were performed using one-way ANOVA with Tukey's multiple comparisons test or using a two-tailed paired *t*-test (e).

Since NicheNet predicted that T_{reg} cells were receiving TCR signals (Fig. 3b), we tested whether a TCR signal is sufficient to induce IL1R1 expression. We sorted T_{reg} cells from peripheral blood, and IL1R1⁻ and IL1R1⁺ T_{reg} cells from HNSCC (Extended Data Fig. 7d). IL1R1 expression on blood T_{reg} cells and IL1R1⁻ tumour T_{reg} cells was induced after 24 h of stimulation with anti-CD3, anti-CD28 and anti-CD2 beads (anti-CD3/CD28/CD2 beads) (Fig. 4b). IL1R1 expression on IL1R1⁺ tumour T_{reg} cells was sustained on day 1, but declined by day 2 and day 3. Stimulation of T_{reg} cells with anti-CD3 and anti-CD28 coated beads optimized for T cell activation led to a more pronounced increase in IL1R1 expression (up to 50% of blood T_{reg} cells and IL1R1⁻ tumour T_{reg} cells, Extended Data Fig. 7e). These data indicate that a TCR combined with a costimulation signal is sufficient to induce IL1R1 expression.

To assess the functional impact of IL-1 signalling on T_{reg} cells, we performed bulk-RNA sequencing of blood T_{reg} cells and IL1R1 $^{-}$ and IL1R1 $^{+}$ HNSCC T_{reg} cells after culture with anti-CD3/CD28/CD2 beads with or without IL-1 for 2 days. The addition of IL-1 led to significant upregulation of more than 50 genes in T_{reg} cells isolated from blood, including *IL2RA*, but had minimal effects on the transcriptional profile of intratumoural T_{reg} cells (Extended Data Fig. 7f). Expression of *FOXP3*, *CTLA4* and *IKZF2* (which encodes Helios) was unaffected by IL-1 signals across all T_{reg} populations, indicating stable maintenance of the T_{reg} phenotype (Extended Data Fig. 7g). The increase in *IL2RA* was mirrored by an increase of IL2RA (also known as CD25) protein expression (Extended Data Fig. 7h). A comparison of the transcriptional profiles of the T_{reg} populations after TCR stimulation revealed that both blood and HNSCC IL1R1 $^{-}$ T_{reg} cells showed high transcriptional activity (more than 900–1,200 transcripts were upregulated), whereas the changes in IL1R1 $^{+}$ T_{reg} cells were much less pronounced (only around 70 transcripts were upregulated; Extended Data Fig. 7i). Together, these data suggest that IL-1 may increase the suppressive function of T_{reg} cells that are not yet fully activated by increasing CD25 expression, whereas intratumoural IL1R1 $^{+}$ T_{reg} cells are already highly activated and less responsive to additional stimulatory signals.

To further assess the responsiveness of these T_{reg} cells to activating signals we used AbSeq for combined protein and transcript profiling of HNSCC-infiltrating T cells with and without short-term stimulation with phorbol myristate acetate (PMA) and ionomycin (Extended Data Fig. 8a). Focusing our analysis on CD4 $^{+}$ T cells and T_{reg} cells (Extended Data Fig. 8b,c), unsupervised clustering revealed three T_{reg} clusters: IL1R1 $^{-}$, IL1R1 $^{+}$ and an actively proliferating IL1R1 $^{+}$ T_{reg} cluster marked by *MKI67* and *TOP2A* (Fig 4c). T_{reg} cells stimulated with PMA and ionomycin showed increased expression of *TNFRSF9* and *CTLA4* transcripts, and IL1R1 $^{+}$ T_{reg} cells showed the largest increase in CTLA4 protein expression (Fig. 4d). These data further underline the notion that IL1R1 $^{+}$ T_{reg} cells are activated and functional, while encompassing a subset that is actively proliferating. Thus,

we next assessed TCR diversity of IL1R1⁺ T_{reg} cells using single-cell VDJ sequencing. We found that 10–20% of IL1R1⁺ T_{reg} cells in HNSCC tumours were clonally expanded, with the top 10 clones in the tumour making up more than 5% of the total number of cells recovered for some donors (Extended Data Fig. [8d–f](#)). The top 3 expanded clones showed distinct gene expression profiles relative to the remaining cells—for example, enrichment for *TNFRSF9* (Extended Data Fig. [8g](#)).

The data collected so far indicate that IL1R1 expression marks a highly suppressive and expanded subset of intratumoural T_{reg} cells. We therefore tested whether it was possible to uniquely identify these cells with a small set of biomarkers. We found that nearly all cells in the CD45⁺IL1R1⁺ICOS⁺ gate were T_{reg} cells (Fig. [4e](#)). Thus, the combined expression of just two cell surface-expressed proteins—IL1R1 and ICOS—could uniquely identify these T_{reg} cells among all haematopoietic (CD45⁺) cells in HNSCC and blood. Finally, in a set of follow-up experiments we assessed whether IL1R1⁺ICOS⁺ T_{reg} cells were truly enriched in HNSCC over inflamed OM tissues and found that IL1R1⁺ICOS⁺ cells were significantly enriched in HNSCC compared with OM (Fig. [4f](#)).

To determine whether IL1R1⁺ T_{reg} cells occur specifically in HNSCC, we analysed and found IL1R1⁺ T_{reg} cells in a small set of additional solid tumour types by flow cytometry (Extended Data Fig. [9a, b](#)). Furthermore, we mined publicly available scRNA-seq datasets^{[38,39](#)}, and found that among intratumoural T cells, expression of *IL1R1* transcript was largely restricted to T_{reg} cells (Extended Data Fig. [9c, d](#)) and was detected across 19 different tumour types (Extended Data Fig. [9e, f](#)). Together, these analyses indicate that IL1R1⁺ T_{reg} cells are not unique for HNSCC, but are also present to varying degrees in other solid tumours.

Discussion

Overall, our data reveal that immune phenotypes typically associated with the human tumour microenvironment are also present in non-malignant, inflamed tissues. The expression pattern of PD-1 was essentially identical in

T cells from both tumour tissues and non-malignant, inflamed tissues, which could offer an explanation for sometimes severe side-effects of systemic anti-PD-1 treatment^{2,40}. Of note, PD-1 expression is typically considered to be driven by TCR signals, but is also upregulated by pro-inflammatory cytokines^{41,42}, which may explain the high expression levels in inflamed tissues. Furthermore, our data indicate that mregDCs³³ are also present in non-malignant, inflamed tissues with minimal transcriptional changes in tumour tissues.

NicheNet analysis of receptor–ligand interactions³⁵ predicted that T_{reg} cells in the tumour actively received TCR signals and responded to IL-1 and IL-18. IL-18 has been implicated in inducing a tissue-repair program by secretion of amphiregulin from IL-18R⁺ T_{reg} cells⁴³. Less is known about the effects of IL-1, traditionally considered a pro-inflammatory cytokine, on T_{reg} cells^{44,45,46}. Whereas IL1R1 expression on T_{reg} cells has been reported in vitro⁴⁷, our data revealed that a transcriptionally and functionally distinct subset of IL1R1⁺ T_{reg} cells is present in vivo in human HNSCC. Of note, we considered using the mouse model system for additional mechanistic studies, but mouse T_{reg} cells did not express IL1R1 in response to TCR-mediated stimulation (Extended Data Fig. 10a–d). We also tried to recapitulate the IL1R1⁺ T_{reg} phenotype in a humanized mouse model of squamous cell carcinoma (Extended Data Fig. 10e). The intratumoural immune infiltrate of the humanized mice was remarkably similar to that in primary human HNSCC tumours, including expression of PD-1, tissue-resident memory T (T_{RM}) cell markers (Extended Data Fig. 10f) and increased T_{reg} infiltration (Extended Data Fig. 10g). However, we only observed low levels of IL1R1 expression by intratumoural T_{reg} cells in humanized mice (Extended Data Fig. 10h). One possible explanation for this difference could be that the tumour microenvironment in humanized mice is sterile, whereas HNSCC (and other human tumours) can contain a viable microbiome⁴⁸. Overall, these data highlight how our human tissue comparison approach identified a T_{reg} population that is otherwise missed in various mouse models.

Of note, compared with IL1R1[−] T_{reg} cells or CD4⁺ T cells, IL1R1⁺ T_{reg} cells expressed higher levels of CXCR6, which was recently reported to be

critical for anti-tumour activity of cytotoxic T cells in a mouse model⁴⁹. It is tempting to speculate that this chemokine–receptor pair could also regulate T_{reg} migration and co-localization with cytotoxic CD8⁺ T cells in the tumour, on the basis of the expression of the ligand CXCL16 in DC3s (Fig. 2g).

The depletion of tumour-infiltrating T_{reg} cells is considered a promising anti-tumour therapy^{50,51,52}. However, therapeutic manipulation of T_{reg} cells in the tumour without affecting other T_{reg} populations has proved difficult. We show that the co-expression of IL1R1 and ICOS uniquely identifies an intratumoural T_{reg} population from all other haematopoietically-derived (CD45⁺) cells in the tumour or peripheral blood. This could provide a possible pathway for tumour-specific depletion of a large intratumoural T_{reg} population using bi-specific antibodies or logic-gated chimeric antigen receptor (CAR) T cells. Overall, our approach could serve as a blueprint to identify immunological congruencies and differences across other tissue types and disease states to improve our understanding of disease-specific processes.

Methods

Primary cells

The head and neck squamous cell carcinoma (HNSCC) tissue samples were obtained after informed consent from otherwise treatment-naïve patients undergoing surgical resection of their primary tumour, ensuring that the immune infiltrate was not influenced by prior therapeutic interventions such as radiotherapy. Inflamed OM tissue biopsies were obtained from individuals undergoing routine dental surgeries for a variety of inflammatory conditions such as perimplantitis, periodontitis or osseous surgery. Matched peripheral blood samples were collected from each tissue donor if possible. All study participants signed a written informed consent before inclusion in the study, and the protocols were approved by the institutional review board (IRB) at the Fred Hutchinson Cancer Research Center (IRB#6007-972 and IRB#8335). A detailed list of the samples and relevant procedure information, together with the panels and/or sequencing experiment

performed is provided in Supplementary Table 1. Furthermore, cryopreserved peripheral blood mononuclear cells (PBMCs) from healthy controls (Seattle Area Control Cohort (SAC)) were obtained via the HIV Vaccine Trial network (HVTN) and used for titrations, panel development and as a longitudinal technical control for all flow cytometry acquisitions (data not shown). The human squamous cell carcinoma line SCC-15 was obtained and validated from ATCC (tested negative for mycoplasma).

Isolation of leukocytes from solid human tissues and peripheral blood

After surgical procedures, fresh tissue samples were placed immediately into a 50-ml conical tube with complete media (RP10: RPMI1640 supplemented with penicillin, streptomycin and 10% fetal bovine serum (FBS)) and kept at 4 °C. Samples were processed within 1–4 h after collection based on optimized protocols adapted from ref. 53. In brief, tissue pieces were minced using a scalpel into small pieces and incubated with Collagenase II (Sigma-Aldrich, 0.7 mg ml⁻¹) and DNase (5 U ml⁻¹) in RPMI1640 with 7.5% FBS for 30–45 min depending on sample size. Subsequently, any remaining tissue pieces were mechanically disrupted by repeated resuspension with a 30 ml syringe with a large bore tip (16 × 1.5 blunt). The cell suspension was filtered using a 70-µm cell strainer, washed in RPMI1640 and immediately used for downstream procedures.

Peripheral blood samples (1–10 ml) were collected in ACD tubes and then processed using SepMate tubes (StemCell Technologies, 85450) and Lymphoprep (Stem Cell Technologies, 07851) according to manufacturer protocols. In brief, whole blood samples were centrifuged for 10 min at 400g, and the plasma supernatant was collected separately and immediately frozen at –80 °C. Remaining cells were resuspended in 30 ml PBS and pipetted on top of 13.5 ml Lymphoprep in a SepMate tube. After centrifugation for 16 min at 1,200g, the mononuclear cell fraction in the supernatant was poured into a fresh 50-ml tube, washed with PBS and immediately used for downstream procedures. For blood samples from dental surgery patients, red blood cells were lysed using ACK-lysis buffer (Thermo Fisher, A10492-01), and the remaining white blood cells were directly used for downstream staining.

If required, cells isolated from tissue samples or from peripheral blood were frozen using either a 90% FBS/10% DMSO mixture or Cell Culture Freezing Medium (Gibco, 12648010), and stored in liquid nitrogen until used for downstream procedures.

Flow cytometry and cell sorting

For flow cytometric analysis good practices were followed as outlined in the guidelines for use of flow cytometry⁵⁴ and consensus suggestions for data analysis⁵⁵. Directly following isolation or thawing, cells were incubated with Fc-blocking reagent (BioLegend Trustain FcX, 422302) and fixable UV Blue Live/Dead reagent (ThermoFisher, L34961) in PBS (Gibco, 14190250) for 15 min at room temperature. After this, cells were incubated for 20 min at room temperature with 50 µl total volume of antibody master mix freshly prepared in Brilliant staining buffer (BD Biosciences, 563794), followed by two washes in fluorescence-activated cell sorting (FACS) buffer (PBS with 2% FBS). All antibodies were titrated and used at optimal dilution, and staining procedures were performed in 96-well round-bottom plates (for cell sorting in 5-ml polystyrene tubes). A detailed list of the main panels used, including fluorochromes, antibody catalogue numbers and final dilutions is provided in Supplementary Table 2 (panels designed according to best practices as described⁵⁶) and Supplementary Table 3. For sorting, cells were immediately used after staining, and for analysis, the stained cells were fixed with 4% PFA (Cytofix/Cytoperm, BD Biosciences, 554722) for 20 min at room temperature, washed, resuspended in FACS buffer and stored at 4 °C in the dark until acquisition. If necessary, intracellular (CD68, granzyme B (GZMB) or CTLA4) or intranuclear staining (FOXP3, KI67, TCF1, TOX, T-bet or EOMES) was performed following the appropriate manufacturer protocols (eBioscience FOXP3/Transcription Factor Staining Buffer Set, Thermo Fisher 00-5532-00).

Single-stained controls were prepared with every experiment using antibody capture beads (BD Biosciences anti-mouse (552843) or anti-mouse Plus, and anti-rat (552844)) diluted in FACS buffer, or cells for Live/Dead reagent, and treated exactly the same as the samples (including fixation procedures). For each staining of experimental samples, a PBMC sample from the same

healthy donor (SAC) was stained with the same panel as a longitudinal reference control (data not shown).

All samples were acquired using a FACSymphony A5 (BD Biosciences), equipped with 30 detectors and 355 nm (65 mW), 405 nm (200 mW), 488 nm (200 mW), 532 nm (200 mW) and 628 nm (200 mW) lasers and FACSDiva acquisition software (BD Biosciences). Full details on the optical configuration of the instruments used are as described¹⁹. Detector voltages were optimized using a modified voltage titration approach⁵⁷ and standardized from day to day using MFI target values and 6-peak Ultra Rainbow Beads⁵⁶ (Spherotec, URCP-38-2K). After acquisition, data was exported in FCS 3.1 format and analysed using FlowJo (version 10.6.x, and 10.7.x, BD Biosciences). Samples were analysed using a combination of manual gating and computational analyses approaches⁵⁵, with doublets being excluded by FSC-A vs FSC-H gating. For fresh samples acquired on different experimental days with the T cell or APC panel, files were exported as compensated data and analysed combined together in a new workspace (see deposited data on www.flowrepository.org). Gates were kept the same across all samples except where changes in the density distribution of populations clearly indicated the need for sample-specific adjustment. For the APC panel, PD-L1 (V450 channel) as well as CD85k (V510 channel) were excluded from analysis because of interference or high variability from highly auto-fluorescent myeloid cells in some samples. For the T cell panel, granzyme B and TIM3 staining showed donor-specific shifts in intensity, requiring sample-specific adjustments of gates.

All cell sorting was performed either on a FACSaria III (BD Biosciences), equipped with 20 detectors and 405 nm, 488 nm, 532 nm and 628 nm lasers or on a FACSymphony S6 cell sorter (BD Biosciences), equipped with 50 detectors and 355 nm, 405 nm, 488 nm, 532 nm and 628 nm lasers. For all sorts involving myeloid cells, an 85-µm nozzle operated at 45 psi sheath pressure was used, for sorts exclusively targeting T cells, a 70-µm nozzle at 70 psi sheath pressure was used. Unless stated otherwise, cells were sorted into chilled Eppendorf tubes containing 500–1,000 µl complete RPMI, washed once in PBS and immediately used for subsequent processing.

Whole-transcriptome single-cell library preparation and sequencing

cDNA libraries were generated using the 10x Genomics Chromium Single Cell 3' Reagent Kits v2 protocol or the v3 protocol, or using the 10x Genomics Chromium Single Cell 5' Reagent Kit v1 protocol (see Supplementary Table 1). In brief, after sorting single cells were isolated into oil emulsion droplets with barcoded gel beads and reverse transcriptase mix using the Chromium controller (10x Genomics). cDNA was generated within these droplets, then the droplets were dissociated. cDNA was purified using DynaBeads MyOne Silane magnetic beads (ThermoFisher, 370002D). cDNA amplification was performed by PCR (10 cycles) using reagents within the Chromium Single Cell 3' Reagent Kit v2 or v3 (10x Genomics) or the VDJ and GEX reagent kit v1 (see list of samples in Supplementary Table 1). Amplified cDNA was purified using SPRIselect magnetic beads (Beckman Coulter) according to the respective protocol. cDNA was enzymatically fragmented and size selected prior to library construction. Libraries were constructed by performing end repair, A-tailing, adaptor ligation, and PCR (12 cycles). Quality of the libraries was assessed by using Agilent 2200 TapeStation with High Sensitivity D5000 ScreenTape (Agilent). Quantity of libraries was assessed by performing digital droplet PCR (ddPCR) with Library Quantification Kit for Illumina TruSeq (BioRad, 1863040) or determined by Qubit with the dsDNA HS Assay (Q32851). Pooled Libraries were diluted to 2 nM or 3 nM and paired-end sequencing was performed on a HiSeq 2500 (Illumina) or a NovaSeq 6000 (Illumina) utilizing S1 or S2 flow cells, targeting between 25,000–50,000 reads per cell.

Targeted transcriptomics single-cell library preparation and sequencing

cDNA libraries were generated as described in detail⁵⁸. In brief, after sorting, single cells were stained with Sample-Tag antibodies (if required, see Extended Data Fig. 8a) and or AbSeq antibodies (if required), washed three times, pooled and counted and subsequently loaded onto a nano-well cartridge (BD Rhapsody), lysed inside the wells followed by mRNA capture on cell capture beads according to manufacturer instructions⁵⁸. Cell Capture

Beads were retrieved and washed prior to performing reverse transcription and treatment with Exonuclease I. cDNA underwent targeted amplification using the Human Immune Response Panel primers and a custom supplemental panel (listed in Supplementary Table 3) via PCR (10–11 cycles). PCR products were purified, and mRNA PCR products were separated from Sample-Tag (and AbSeq, where applicable) PCR products with double-sided size selection using SPRIselect magnetic beads (Beckman Coulter). mRNA and Sample Tag products were further amplified using PCR (ten cycles). PCR products were then purified using SPRIselect magnetic beads. Quality of PCR products was determined by using an Agilent 2200 TapeStation with High Sensitivity D5000 ScreenTape (Agilent) in the Fred Hutch Genomics Shared Resource laboratory. The quantity of PCR products was determined by Qubit with Qubit dsDNA HS Assay (Q32851). Targeted mRNA product was diluted to $2.5 \text{ ng } \mu\text{l}^{-1}$, and the Sample Tag and AbSeq PCR products were diluted to $1 \text{ ng } \mu\text{l}^{-1}$ to prepare final libraries. Final libraries were indexed using PCR (6 cycles). Index PCR products were purified using SPRIselect magnetic beads. Quality of all final libraries was assessed by using Agilent 2200 TapeStation with High Sensitivity D5000 ScreenTape and quantified using a Qubit Fluorometer using the Qubit dsDNA HS Kit (ThermoFisher). Final libraries were diluted to 3 nM and multiplexed for paired-end (100 bp) sequencing on a NovaSeq 6000 (Illumina) using S1 and S2 flow cells. For the gene expression library, we targeted 5,000–20,000 reads per cell, for the AbSeq library 10,000–15,000 reads per cell, and for the Sample-Tag libraries 500–2,000 reads per cell.

Ex vivo stimulation assays

Cells were isolated from tissues or blood as described above. For some of the stimulation assays cryo-preserved cell suspensions were used after assessing good cellular viability. For the 2 h short-term stimulation assays with targeted transcriptomics (Fig. 4), CD3⁺ T cells (live CD45⁺CD19⁻CD3⁺ events) were isolated using FACS using a BD FACSAria II. Five-thousand cells were placed into each well of a V-bottom 96-well plate with 200 μl complete media. Cells were then left untreated (control), or stimulated with IL-12, IL-15 and IL-18 (each at 1 nM), or with PMA (50 ng ml^{-1}) and ionomycin (500 ng ml^{-1}) for 2 h at 37 °C. Cells were

then washed with 1× PBS and prepared for targeted transcriptomics and staining with oligonucleotide-conjugated antibodies as described⁵⁸. For the 1- to 3-day stimulation assays (Fig. 4, Extended Data Fig. 7), CD4⁺CD25⁺CD127⁻IL1R1⁺ and IL1R1⁻ T_{reg} cells were isolated from blood and HNSCC tissues using a FACSsymphony S6 sorter (BD Biosciences), and cultured either in RP10 alone or with anti-CD3/CD28 Dynabeads (Gibco, 11161D, used at a 1:1 bead-to-cell ratio) or with anti-CD3/CD28/CD2 beads (Miltenyi, 130-092-909, T_{reg} Suppression Inspector, also used at a 1:1 bead-to-cell ratio), either with or without recombinant IL-1 β (Peprotech, 200-01B) at 50 ng ml⁻¹. For some experiments, culture cells were subsequently stained and 250–500 viable cells were sorted on an BD S6 sorter followed by bulk RNA-sequencing (RNA-seq) analysis using the SMART-Seq v4 kit (Takara) as described further below.

Suppression assays

For suppression assays, IL1R1⁺ and IL1R1⁻CD4⁺CD25⁺CD127⁻ regulatory T cells and CD4⁺CD25⁻ and CD8⁺ T_{resp} cells were sorted from cryopreserved HNSCC samples. For some experiments, matched peripheral blood was included. T_{resp} cells were labelled with Cell Trace Violet (CTV) according to the manufacturer instructions (Thermo Fisher, C34571). In brief, 10⁶ sorted T_{resp} cells were washed with PBS after the sort, and then incubated in pre-warmed PBS containing a final concentration of 5 μ M freshly diluted CTV for 15 min. The reaction was quenched with prewarmed RP10. Both T_{resp} and T_{reg} cells were counted twice on a BioRad TC20 cell counter. 10,000 (20,000 for some experiments) CTV-labelled T_{resp} cells were cultured alone, or with 10,000 T_{reg} cells (or titred amounts of T_{reg} cells) in a 96-well round-bottom plate at 37 °C for 4 days together with anti-CD3/CD28/CD2 beads (Miltenyi, 130-092-909, T_{reg} Suppression Inspector). An unstimulated control well was included with every experiment. Where indicated, recombinant IL-1 β (Peprotech, 200-01B) was added to achieve a final concentration of 50 ng ml⁻¹. On the read-out day, supernatants were collected and frozen at -80 °C, and the cells were stained with a 14-colour readout panel including Live/Dead reagent (Supplementary Table 2), fixed and acquired on a BD FACSsymphony A5, as described above. Cell

proliferation was assessed by using the proliferation platform in FlowJo 10.7 (BD Biosciences), with percentage of divided cells (modelled, not gated) as the main readout. Supernatants were processed for Luminex analysis by the Immunomonitoring Core of the Fred Hutchinson Cancer Research Center.

Luminex analysis of tumour lysates

Luminex analysis was performed on lysates of tissues. To obtain lysates from tumour tissues, a 2×2 mm piece was incubated for one minute in PBS/0.1% tween. After incubation, the tissue piece was minced in the buffer and then centrifuged at 10,000 rpm for 5 min. The supernatant was collected and immediately flash-frozen on dry ice. Processing for Luminex was performed by the Immunomonitoring Core of the Fred Hutchinson Cancer Research Center.

Isolation and stimulation of mouse cells

Mouse protocols were approved by and in compliance with the ethical regulations of Fred Hutchinson Cancer Research Center's IACUC. All animals were maintained in specific pathogen-free facilities and euthanized in accordance with institutional protocols. We received thymus, spleen, and lymph node (LN) from male *Foxp3^{eGFP-cre-ERT2}* mice (age ≥ 8 weeks) (from J. Lund), and mechanically dissociated thymus, spleen or lymph node through a 70- μm strainer. To enrich T cells from spleen–lymph node single-cell suspensions, we used a T cell-negative isolation based magnetic enrichment (Stemcell Technologies). For TCR stimulations, we prepared plate-bound anti-CD3 and anti-CD28 by incubating 96-well V-bottom tissue culture plates with 100 μl of 1 $\mu\text{g ml}^{-1}$ anti-CD3 (clone: 145-2C11) and 2 $\mu\text{g ml}^{-1}$ anti-CD28 (clone 37.51) in 1 \times PBS for 3 h at 37 °C. We decanted and washed residual anti-CD3/anti-CD28 solution and plated 1×10^6 isolated T cells per well in 96-well V-bottom tissue culture plates. We cultured cells in modified RP10 media (RPMI1640 supplemented with 10% FBS, 2mM l-glutamine, 100 U ml^{-1} penicillin-streptomycin, 1 mM sodium pyruvate, 0.05 mM β -mercaptoethanol and 1 mM HEPES). We collected cells for flow analysis at 0-, 1- and 2-day time points for flow cytometric analysis as described above. The following panel was used: anti-TCR $\gamma\delta$ –PerCPe710 (clone eBioGL3), anti-CD4–BV786 (clone GK1.5), anti-CD8a–

V500 (clone 53-6.7), anti-CD44–AF700 (clone IM7), anti-CD69–PECy7 (clone H1.2F3), anti-PD-1–BV605 (clone 28F.1A12), anti-ICOS–AF647 (clone C398.4A), anti-IL1R1–PE (clone 35F5), anti-IL1R2–BV421 (clone 4E2), anti-CD3–BUV805 (clone 17A2) and anti-FOXP3–FITC (clone FJK-16s, intranuclear post fixation).

Humanized mouse experiments

MISTRG mice ($\text{M-CSF}^{\text{h/h}}\text{IL-3/GM-CSF}^{\text{h/h}}\text{SIRP}\alpha^{\text{h/m}}\text{TPO}^{\text{h/h}}\text{RAG2}^{-/-}\text{IL2R}\gamma^{-/-}$) were previously reported⁵⁹. All animal experiments were approved by Fred Hutchinson Cancer Research Center’s Institutional Animal Care and Use Committee (protocol 50941). De-identified human fetal liver tissues, obtained with informed consent from the donors, were procured by Advanced Bioscience Resources and their use was determined as non-human subject research by Fred Hutch’s Institutional Review Board (6007-827). Fetal livers were cut in small fragments, treated for 45 min at 37 °C with collagenase D (Roche, 100 ng ml⁻¹), and a single-cell suspension was prepared. Hematopoietic cells were enriched by density gradient centrifugation in Lymphocyte Separation Medium (MP Biomedicals) followed by positive immunomagnetic selection with anti-human CD34 microbeads (Miltenyi Biotec). Purity (>90% CD34⁺ cells) was confirmed by flow cytometry and cells were frozen at -80 °C in FBS containing 10% DMSO. Newborn MISTRG mice (day 1–3) were sublethally irradiated (80 cGy gamma rays in a Caesium-137 irradiator) and ~20,000 CD34⁺ cells in 20 µl PBS were injected into the liver with a 22-gauge needle (Hamilton Company), as described⁵⁹. Engraftment levels were measured as the percentage of human CD45⁺ cells among total (mouse and human combined) CD45⁺ cells in the blood.

The human squamous cell carcinoma line SCC-15 was obtained and verified from ATCC. Cells were grown to ~80% confluence in DMEM/F12 supplemented with 12.5 mM l-glutamine, 15 mM HEPES, 0.5 mM sodium pyruvate and 400 ng ml⁻¹ hydrocortisone. Approximately 0.5 million cells per mouse were resuspended in 75 µl PBS, mixed with 25 µl growth-factor-reduced Matrigel (Corning, 354230) and then injected subcutaneously under anaesthesia in the flank of humanized mice. The size of the tumours was

measured weekly for 7 weeks with a calliper. SCC15 tumour tissues were processed for leukocyte isolation as described above for human tissues.

Bulk RNA-seq experiments and analysis

Bulk RNA-seq was performed on 250 sort-purified IL1R1⁺ and IL1-R1⁻ T_{reg} cells derived from either cryopreserved blood or HNSCC tissues samples after culture in conditions of no stimulation, stimulation with anti-CD3/CD28/CD2 beads, and stimulation with anti-CD3/CD28/CD2 beads and IL-1 β (50 ng ml⁻¹) for days 1, 2, and 3. In total, 88 samples were sequenced, and each condition was represented by at least 3 or more biological replicates.

Cells were sorted directly into lysis buffer from the SMART-Seq v4 Ultra Low Input RNA Kit for sequencing (Takara), immediately snap frozen on dry ice, and transferred to -80 °C storage until processed into cDNA. All samples were thawed, cells were lysed, and cDNA was synthesized and amplified per the manufacturer's instruction. After amplification, sequencing libraries were constructed using the NexteraXT DNA sample preparation kit with unique dual indexes (Illumina) to generate Illumina-compatible barcoded libraries. Libraries were pooled and quantified using a Qubit Fluorometer (Life Technologies). Sequencing of pooled libraries was carried out on a NextSeq 2000 sequencer (Illumina) with paired-end 59-base reads, using a NextSeq P2 sequencing kit (Illumina) with a target depth of 5 million reads per sample.

Base calls were processed to FASTQs on BaseSpace (Illumina), and a base call quality-trimming step was applied to remove low-confidence base calls from the ends of reads. Reads were processed using workflows managed on the Galaxy platform. Reads were trimmed by 1 base at the 3' end then trimmed from both ends until base calls had a minimum quality score of at least 30. Any remaining adapter sequence was removed as well. To align the trimmed reads, STAR aligner (v2.4.2a) was used with the GRCh38 reference genome and gene annotations from ensembl release 91. Gene counts were generated using HTSeq-count (v0.4.1). Quality metrics were compiled from PICARD (v1.134), FASTQC (v0.11.3), Samtools (v1.2), and HTSeq-count (v0.4.1).

A quality filter was applied to retain libraries in which the fraction of aligned reads examined compared to total FASTQ reads was >70%, the median coefficient of variation of coverage was less than 0.85, and the library had at least 1 million reads. All sequenced samples passed these quality filters. Non-protein coding genes and genes expressed at less than 1 count per million in fewer than 10% of samples were filtered out.

Expression counts were normalized using the TMM algorithm. For differential gene expression analysis, the linear models for microarray data (Limma) R package after Voom transformation was used; this approach either outperforms or is highly concordant with other published methods. Linear models were generated, and donor identity was included as a random effect. For differential gene expression comparisons, genes with a false discovery rate (FDR) of less than 0.1 and an absolute expression fold-change greater than 1 were considered differentially expressed.

Pre-processing for whole transcriptome analysis (WTA) and targeted transcriptomics data

Raw base call (BCL) files were demultiplexed to generate Fastq files using the Cell Ranger mkfastq pipeline within Cell Ranger (10x Genomics).

Whole-transcriptome Fastq files were processed using the standard Cell Ranger pipeline (10x genomics) within Cell Ranger 2.1.1 or Cell Ranger 3.0.2. In brief, Cell Ranger count performs read alignment, filtering, barcode and unique molecular identifier (UMI) counting, and determination of putative cells. The final output of Cell Ranger (the molecule per cell count matrix) was then analysed in R using the package Seurat^{60,61} (3.0) as described below. For targeted transcriptomics data, Fastq files were processed via the standard Rhapsody analysis pipeline (BD Biosciences) on Seven Bridges (www.sevenbridges.com). In brief, after read filtering, reads are aligned to a reference genome and annotated, barcodes and UMIs are counted, followed by determining putative cells. The final output (molecule per cell count matrix) was also analysed in R using Seurat^{60,61} (version 3.0) as described below. For 5' VDJ sequencing experiments, the output after Cell Ranger vdj was analysed using the Loupe VDJ browser v3 (10x Genomics). For the SMART-Seq v4 experiments, Fastq files were aligned to the GRCh38 reference genome as described in more detail above.

Seurat workflow for targeted transcriptomics and WTA data

The R package Seurat^{60,61} was used for all downstream analysis, with custom scripts based on the following general guidelines for analysis of scRNA-seq data⁶².

In brief, for whole-transcriptome data, only cells that had at least 200 genes (v2 kits) or 800 genes (v3 kits), and depending on sample distribution less than 7–15% mitochondrial genes were included in analysis. All acquired samples were merged into a single Seurat object, followed by a natural log normalization using a scale factor of 10,000, determination of variable genes using the vst method, and a z-score scaling. Principal component analysis was used to generate 75 principal components, followed by data integration using Harmony³⁰. The dimensionality reduction generated by Harmony was used to calculate UMAP, and graph-based clustering with a resolution between 0.2 and 0.6. For cell annotation, we applied SingleR as a purely data-driven approach³², and used the expression of typical lineage transcripts to verify the cell label annotation. For all subsequent analysis steps, the integrated Seurat object was separated into two objects containing all T cells or all APCs, respectively, and UMAP calculation as well as clustering steps were repeated.

For targeted transcriptomics data³⁶, separate cartridges from the same experiment were merged (if applicable), and only cells that had at least 30 genes were included in downstream analysis. After generating a Seurat object, a natural log normalization using a scale factor of 10,000 was done, followed by determination of variable genes using the vst method, and a z-score scaling. Principal component analysis was used to generate 75 principal components, followed by data integration using Harmony³⁰. The dimensionality reduction from Harmony was used for subsequent UMAP calculation and graph-based clustering with tuned resolution. Protein phenotyping data was stored in a separate slot as described in the Seurat tutorial for CITE-seq data, and normalized using the centred log ratio (CLR) method³⁶. For some figures, the count matrices were exported as FCS files using the package Premessa, and then imported into FlowJo 10.7.x. Appropriate arcsinh transformations were applied in a channel-specific

manner, and transcript or protein expression was plotted and quantified using two-dimensional plots.

For all differential gene expression analyses we utilized the Seurat implementation of MAST (model-based analysis of single-cell transcriptomes) with the number of UMIs included as a covariate (proxy for cellular detection rate (CDR)) in the model³⁴. For calculating the T helper scores (Extended Data Fig. 5f,g) we used the AddModuleScore function of Seurat (see Github script on https://github.com/MairFlo/Tumor_vs_Inflamed/blob/main/OM_HNSCC_sc_RNAseq_Harmony). The genes used were as follows: T_{H1}: *IFNG*, *TBX21*, *IL12RB1* and *IL12RB2*; T_{H2}: *TNFSF11*, *GATA3* and *IL4*; T_{H17}: *RORC*, *CCR6*, *IL17A*, *IL17F*, *IL23R*, *IL22*, *AHR*, *IL26*, *CCL20*; T_C: *CD8B*, *CD8A*, *TNF*, *IFNG*, *IL2*, *GZMB*, *PRF1*, *GZMA* and *FAS*. T_{ex}: *TCF7*, *TOX*, *HAVCR2*, *PDCD1* and *LAG3*; T_{reg}: *FOXP3*, *CTLA4*, *IL2RA*, *IL2RB* and *ENTPD1*.

NicheNet workflow

NicheNet analysis was adapted from the vignette described at <https://github.com/saeyslab/nichennetr>³⁵. In brief, the separate Seurat objects containing APCs (described above) were subsetted to contain only HNSCC-derived cells, and the Seurat object containing T cells only HNSCC and OM-derived cells. During multiple separate NicheNet runs, different T cell subsets were set as ‘receiver’ (that is, CD4 non-T_{reg} clusters 0 and 2, CD8 T cell clusters 1, 3 and 4 and T_{reg} cluster 5; Extended Data Fig. 5a) and all myeloid cell clusters (except the pDC and mast cell cluster; Fig 2b) as ‘sender’ populations. For the receiver cell population, a DEG test was performed to find genes enriched in HNSCC vs OM samples, with the key parameters being set as follows: genes expressed in at least 10% of the cells of the respective T cell clusters, and filtered after the DEG test for an adjusted *P*-value of less than 0.05 and average log fold change more than 0.25. For the sending cell population, all ligands expressed in at least 5% of the cells in the respective APC cluster were considered. NicheNet analysis was performed based on the vignette to infer receptors, filter for documented links and generate a circus plot of the top ligand-receptor interactions for the

respective cellular populations. Scoring of the predicted targets was based on a Pearson correlation coefficient as described in the NicheNet vignette.

Circos plots were generated as described in the vignette³⁵ to visualize links between ligands on APCs and receptors on the T cell subsets.

FAUST analysis

For the T cell panel, FAUST was used to discover and annotate phenotypes in 22 samples (11 HNSCC and 11 OM). FAUST was applied to CD45⁺ live lymphocytes identified through manual gating. The MR1-tetramer, CD45 and the Live/Dead marker were excluded from the FAUST analysis to account for the manual analysis. After tuning, FAUST selected the markers CD8, CD4, CD3, CD45RA, CD27, CD19, CD103, CD69, CD28, HLADR, GZMB, PD-1, CD25, ICOS, TCR $\gamma\delta$, CD38 and TIM3 for discovery and annotation of phenotypes. Counts of the discovered phenotypes labelled CD3⁺ and CD19⁻ were tested for association with tissue type using a binomial generalized linear mixed-effects model with a subject level random effect. Fifty phenotypes were associated with tissue type at the FDR-adjusted 0.05 level.

For the APC panel, FAUST was used to discover and annotate phenotypes in 32 samples (16 HNSCC and 16 OM). FAUST was applied to CD45⁺ live CD19⁻CD3⁻ cells identified through manual gating. The markers CD3, CD19, CD45, PD-L2 and CD85k and the Live/Dead marker were excluded from the FAUST analysis to account for the manual analysis as well as observed autofluorescence in the detectors used for PD-L2 and CD85k. After tuning, FAUST selected the markers CD1c, CD11b, CD11c, CD14, CD16, CD32, CD38, CD40, CD68, CD80, CD86, CD123, CD141, CD163, CD206, CX3CR1, HLADR, PDL1 and SIRPA for discovery and annotation of phenotypes. Counts of the discovered phenotypes annotated as HLADR⁺ were tested for association with tissue type using a binomial generalized linear mixed-effects model with a subject level random effect. 21 phenotypes were associated with tissue type at the FDR-adjusted 0.05 level.

Statistical analyses

Unless stated otherwise, all data are represented as mean \pm s.d. Statistical analyses between blood, OM and HNSCC samples were performed using one-way ANOVA with Tukey's multiple comparisons test. *P*-values are shown in full, except if smaller than 0.0001. Statistical analysis was performed using GraphPad Prism (v9).

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

The scRNA-seq data as well as the bulk RNA-seq data discussed in this publication have been deposited in the NCBI Gene Expression Omnibus under accession [GSE163633](#). Alignment was based on the GRCh38 reference genome. Flow cytometry raw data have been deposited at <https://www.flowrepository.org> using the Identifiers [FR-FCM-Z4UX](#), [FR-FCM-Z4UP](#) and [FR-FCM-Z4UQ](#) or can be requested from F.M (fmair@fredhutch.org). The data used in Extended Data Fig. 9c–e are from the scRNA-seq Data Portal for T cells in Pan Cancer at http://cancer-pku.cn:3838/PanC_T.

Code availability

The main R scripts used for data processing of the scRNA-seq data are available on https://github.com/MairFlo/Tumor_vs_Inflamed or can be requested from F.M. The scripts used for processing of the bulk RNA-seq data are available on https://github.com/akonecny/Bulk-RNAseq-Tumor_vs_Inflamed.

References

1. Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**, 1–10(2013).

2. Martins, F. et al. Adverse effects of immune-checkpoint inhibitors: epidemiology, management and surveillance. *Nat. Rev. Clin. Oncol.* **16**, 563–580 (2019).
3. Greten, F. R. & Grivennikov, S. I. Inflammation and cancer: triggers, mechanisms, and consequences. *Immunity* **51**, 27–41 (2019).
4. Mujal, A. M. & Krummel, M. F. Immunity as a continuum of archetypes. *Science* **364**, 28–29 (2019).
5. Fan, X. & Rudensky, A. Y. Hallmarks of tissue-resident lymphocytes. *Cell* **164**, 1198–1211 (2016).
6. Kumar, B. V. et al. Human tissue-resident memory T cells are defined by core transcriptional and functional signatures in lymphoid and mucosal sites. *Cell Rep.* **20**, 2921–2934 (2017).
7. Szabo, P. A. et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, 4706–4716 (2019).
8. Amsen, D., van Gisbergen, K. P. J. M., Hombrink, P. & van Lier, R. A. W. Tissue-resident memory T cells at the center of immunity to solid tumors. *Nat. Immunol.* **19**, 538–546 (2018).
9. Ji, A. L. et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* **182**, 497–514.e422 (2020).
10. Scott, A. C. et al. TOX is a critical regulator of tumour-specific T cell differentiation. *Nature* **571**, 270–274 (2019).
11. Chao, J. L. & Savage, P. A. Unlocking the complexities of tumor-associated regulatory T cells. *J. Immunol.* **200**, 415–421 (2018).
12. Sharma, P. & Allison, J. P. Dissecting the mechanisms of immune checkpoint therapy. *Nat. Rev. Immunol.* **20**, 75–76 (2020).

13. Woodward Davis, A. S. et al. The human tissue-resident CCR5⁺ T cell compartment maintains protective and functional properties during inflammation. *Sci. Transl. Med.* **11**, eaaw8718 (2019).
14. Guilliams, M. et al. Dendritic cells, monocytes and macrophages: a unified nomenclature based on ontogeny. *Nat. Rev. Immunol.* **14**, 571–578 (2014).
15. Cabeza-Cabrero, M., Cardoso, A., Minutti, C. M., Pereira da Costa, M. & Reis, E. S. C. dendritic cells revisited. *Annu. Rev. Immunol.* **39**, 131–166 (2021).
16. Binnewies, M. et al. Unleashing type-2 dendritic cells to drive protective antitumor CD4⁺ T cell immunity. *Cell* **177**, 556–571.e516 (2019).
17. Wculek, S. K. et al. Dendritic cells in cancer immunology and immunotherapy. *Nat. Rev. Immunol.* **144**, 646–618 (2019).
18. Chow, L. Q. M. Head and neck cancer. *New Engl. J. Med.* **382**, 60–72 (2020).
19. Mair, F. & Prlic, M. OMIP-044: 28-color immunophenotyping of the human dendritic cell compartment. *Cytometry A* **106**, 255 (2018).
20. Suzuki, S. et al. Immune-checkpoint molecules on regulatory T-cells as a potential therapeutic target in head and neck squamous cell cancers. *Cancer Sci.* **111**, 1943–1957 (2020).
21. Duhen, T. et al. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* **9**, 2724 (2018).
22. Savas, P. et al. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat. Med.* **24**, 986–993 (2018).
23. Blank, C. U. et al. Defining ‘T cell exhaustion’. *Nat. Rev. Immunol.* **19**, 665–674 (2019).

24. Bourdely, P. et al. Transcriptional and functional analysis of CD1c⁺ human dendritic cells identifies a CD163⁺ subset priming CD8⁺CD103⁺ T cells. *Immunity* **53**, 335–352.e338 (2020).
25. Segura, E. & Amigorena, S. Inflammatory dendritic cells in mice and humans. *Trends Immunol.* **34**, 440–445 (2013).
26. Dutertre, C.-A. et al. Single-cell analysis of human mononuclear phagocytes reveals subset-defining markers and identifies circulating inflammatory dendritic cells. *Immunity* **51**, 573–589.e578 (2019).
27. Salmon, H. et al. Expansion and activation of CD103⁺ dendritic cell progenitors at the tumor site enhances tumor responses to therapeutic PD-L1 and BRAF inhibition. *Immunity* **44**, 924–938 (2016).
28. Lavin, Y. et al. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. *Cell* **169**, 750–765.e717 (2017).
29. Greene, E. et al. New interpretable machine-learning method for single-cell data reveals correlates of clinical response to cancer immunotherapy. *Patterns* **2**, 100372 (2021).
30. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
31. McInnes, L., Healy, J. & Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.1802.03426> (2018).
32. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
33. Maier, B. et al. A conserved dendritic-cell regulatory program limits antitumour immunity. *Nature* **580**, 257–262 (2020).
34. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell

RNA sequencing data. *Genome Biol.* **16**, 278 (2015).

35. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
36. Mair, F. et al. A targeted multi-omic analysis approach measures protein expression and low-abundance transcripts on the single-cell level. *Cell Rep.* **31**, 107499 (2020).
37. Freeman, Z. T. et al. A conserved intratumoral regulatory T cell signature identifies 4-1BB as a pan-cancer target. *J. Clin. Invest.* **130**, 1405–1416 (2020).
38. Cillo, A. R. et al. Immune landscape of viral- and carcinogen-driven head and neck cancer. *Immunity* **52**, 183–199.e189 (2020).
39. Zheng, L. et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
40. Luoma, A. M. et al. Molecular pathways of colon inflammation induced by cancer immunotherapy. *Cell* **182**, 655–671.e622 (2020).
41. Kinter, A. L. et al. The common γ -chain cytokines IL-2, IL-7, IL-15, and IL-21 induce the expression of programmed death-1 and its ligands. *J. Immunol.* **181**, 6738–6746 (2008).
42. Maurice, N. J., Berner, J., Taber, A. K., Zehn, D. & Prlic, M. Inflammatory signals are sufficient to elicit TOX expression in mouse and human CD8 $^{+}$ T cells. *JCI Insight* **6**, e150744 (2021).
43. Arpaia, N. et al. A distinct function of regulatory T cells in tissue protection. *Cell* **162**, 1078–1089 (2015).
44. Alvarez, F. et al. The alarmins IL-1 and IL-33 differentially regulate the functional specialisation of Foxp3 $^{+}$ regulatory T cells during mucosal inflammation. *Mucosal Immunol.* **12**, 746–760 (2019).

45. Mercer, F., Kozhaya, L. & Unutmaz, D. Expression and function of TNF and IL-1 receptors on human regulatory T cells. *PLoS ONE* **5**, e8639 (2010).
46. Nikolouli, E. et al. Recirculating IL-1R2⁺ T_{reg}s fine-tune intrathymic T_{reg} development under inflammatory conditions. *Cell. Mol. Immunol.* **30**, 531 (2020).
47. Tran, D. Q. et al. Selective expression of latency-associated peptide (LAP) and IL-1 receptor type I/II (CD121a/CD121b) on activated human FOXP3⁺ regulatory T cells allows for their purification from expansion cultures. *Blood* **113**, 5125–5133 (2009).
48. De Martin, A. et al. Distinct microbial communities colonize tonsillar squamous cell carcinoma. *Oncoimmunology* **10**, 1945202 (2021).
49. Di Pilato, M. et al. CXCR6 positions cytotoxic T cells to receive critical survival signals in the tumor microenvironment. *Cell* **184**, 4512–4530.e4522 (2021).
50. Togashi, Y., Shitara, K. & Nishikawa, H. Regulatory T cells in cancer immunosuppression—implications for anticancer therapy. *Nat. Rev. Clin. Oncol.* **16**, 356–371 (2019).
51. Plitas, G. & Rudensky, A. Regulatory T cells in cancer. *Annu. Rev. Cancer Biol.* **4**, 459–477 (2020).
52. Simpson, T. R. et al. Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. *J. Exp. Med.* **210**, 1695–1710 (2013).
53. Leelatian, N. et al. Single cell analysis of human tissues and solid tumors with mass cytometry. *Cytometry B* **92**, 68–78 (2017).
54. Cossarizza, A. et al. Guidelines for the use of flow cytometry and cell sorting in immunological studies (third edition). *Eur. J. Immunol.* **51**, 2708–3145 (2021).

55. Liechti, T. et al. An updated guide for the perplexed: cytometry in the high-dimensional era. *Nat. Immunol.* **22**, 1190–1197 (2021).
56. Mair, F. & Tyznik, A. J. High-dimensional immunophenotyping with fluorescence-based cytometry: a practical guidebook. *Methods Mol. Biol.* **2032**, 1–29 (2019).
57. Perfetto, S. P., Ambrozak, D., Nguyen, R., Chattopadhyay, P. K. & Roederer, M. Quality assurance for polychromatic flow cytometry using a suite of calibration beads. *Nat. Protoc.* **7**, 2067–2079 (2012).
58. Erickson, J. R. et al. AbSeq protocol using the nano-well cartridge-based rhapsody platform to generate protein and transcript expression data on the single-cell level. *STAR Protoc.* **1**, 100092 (2020).
59. Rongvaux, A. et al. Development and function of human innate immune cells in a humanized mouse model. *Nat. Biotechnol.* **32**, 364–372 (2014).
60. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).
61. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
62. Amezquita, R. A. et al. Orchestrating single-cell analysis with Bioconductor. *Nat. Methods* **17**, 137–145 (2020).

Acknowledgements

We thank the HIV Vaccine Trial Network (HVTN) for access to their flow cytometry instrumentation, the Flow Cytometry Shared Resources of the FHCRC, the Genomics Shared Resources of the FHCRC and the Genomics Core Lab of the Benaroya Research Institute for sequencing. We thank J. Linton for humanized mouse generation; members of the Prlic laboratory for critical discussion; E. Newell, D. Zehn and A. J. Tyznik for advice; S. Bowell, P. Driskell and A. Subang for expert help with human sample

collection; and Regeneron Pharmaceuticals for the use of MISTRG mice. This work was supported by the Emerson Collective (to M.P.), the Washington Research Foundation (WRF) (to M.P.), NIH grants R01AI123323 (to M.P.) and R21DE026565 (to M.P. and D.D.) and the FHCRC President's Young Investigator Award (to M.P.). F.M. is an ISAC Marylou scholar and was supported through The American Association of Immunologists (AAI) Intersect Fellowship Program for Computational Scientists and Immunologists. J.R.E. was supported by National Institutes of Health (NIH) T32 AI007509-20. N.J.M was supported by National Cancer Institute Grant F99 CA245735. A.R. was supported by funds from the Bezos family and the FHCRC Immunotherapy integrated research center. This work used the FHCRC Scientific Computing High Performance nodes, which were supported by ORIP grant S10OD028685. The experimental overview schemes in Fig. [1a](#), Extended Data Fig. [8a](#) use templates from BioRender.com.

Author information

Author notes

1. Douglas Dixon

Present address: Department of Periodontics, University of Tennessee Health Science Center, College of Dentistry, Memphis, TN, USA

2. Raphael Gottardo

Present address: University of Lausanne and Lausanne University Hospital, Switzerland, Lausanne, Switzerland

3. These authors contributed equally: Florian Mair, Jami R. Erickson

Authors and Affiliations

1. Fred Hutchinson Cancer Research Center, Vaccine and Infectious Disease Division, Seattle, WA, USA

Florian Mair, Jami R. Erickson, Marie Frutoso, Andrew J. Konecny, Evan Greene, Valentin Voillet, Nicholas J. Maurice, Raphael Gottardo & Martin Prlic

2. Department of Immunology, University of Washington, Seattle, WA, USA

Jami R. Erickson, Andrew J. Konecny & Anthony Rongvaux

3. Cape Town HVTN Immunology Laboratory, Hutchinson Centre Research Institute of South Africa, NPC (HCRISA), Cape Town, South Africa

Valentin Voillet

4. Molecular and Cellular Biology Graduate Program, University of Washington, Seattle, WA, USA

Nicholas J. Maurice

5. Fred Hutchinson Cancer Research Center, Clinical Research Division, Seattle, WA, USA

Anthony Rongvaux

6. Department of Periodontics, School of Dentistry, University of Washington, Seattle, WA, USA

Douglas Dixon

7. Department of Otolaryngology-Head and Neck Surgery, University of Washington, Seattle, WA, USA

Brittany Barber

8. Department of Statistics, University of Washington, Seattle, WA, USA

Raphael Gottardo

9. Department of Global Health, University of Washington, Seattle, WA,
USA

Martin Prlic

Contributions

Conceptualization: F.M., J.R.E. and M.P. Methodology: F.M., J.R.E., M.F. and A.J.K. Software: F.M., J.R.E., E.G. and V.V. Formal analysis: F.M., J.R.E., A.J.K., E.G., V.V. and M.P. Investigation: F.M., J.R.E., M.F., A.J.K., N.J.M., A.R. and M.P. Writing, original draft: F.M., J.R.E. and M.P. Writing, review and editing: F.M., J.R.E., M.F., E.G., V.V. D.D., A.J.K., N.J.M., A.R., B.B., R.G. and M.P. Visualization: F.M. Funding acquisition: D.D., R.G. and M.P. Resources: D.D. and B.B. Supervision: M.P.

Corresponding author

Correspondence to [Martin Prlic](#).

Ethics declarations

Competing interests

Part of the single-cell sequencing experiments was supported with reagents received from BD Biosciences and Illumina. BD Biosciences and Illumina were not involved in data analysis and interpretation. R.G. has received speaker fees from Illumina and Fluidigm and support from Juno Therapeutics and Janssen Pharma, has consulted for Takeda Vaccines, Juno Therapeutics and Infotech Soft, and has ownership interest in Modulus Therapeutics. E.G. and R.G. declare ownership interest in Ozette Technologies. F.M., J.R.E. and M.P. hold the related patent *Specific Targeting of Tumor-infiltrating regulatory T cells (Tregs) using ICOS and IL-1R1* (US patent no. 63/092957).

Peer review

Peer review information

Nature thanks Sergio Quezada and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

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

Extended data figures and tables

[Extended Data Fig. 1 Additional immune subset quantifications and representative flow cytometry data \(related to main Figure 1\).](#)

(a) Quantification of CD3+ T cells, CD19+ B cells and CD56+ NK cells across the different tissue sources (blue: peripheral blood, orange: OM, red: HNSCC). (b) Frequency of CD4+ and CD8+ T cells within the CD3+ T cell compartment. (c) Quantification of CD4+ CD25+ CD127- regulatory T cells (Tregs) in the different tissue samples. (d) Intranuclear staining for Foxp3 and CTLA-4 on a representative HNSCC sample to confirm that CD25+ CD127- cells are bona fide Foxp3+ Tregs (red histograms: CD4+ CD25+ CD127- cells, grey histograms: CD4+ CD25- CD127+/- cells). (e) Representative histograms and quantification for the expression for TCF-1 (left, n = 7 for OM and n = 10 for HNSCC) and CD39 (right, n = 9) on CD8+ T cells. (f) Quantification for TCF-1 specifically on PD-1+ CD8+ T cells (left) and for the MFI of the transcription factor TOX on CD8+ T cells (right). (n = 7 for OM and n = 6 for HNSCC) (g) Staining patterns for phenotyping markers in the high-dimensional T cell panel, pregated on live CD3+ CD8+ T cells (HNSCC). Positivity cut-offs were the same for all samples, except where shifts in staining patterns based on density distributions indicated the need for adjustments. GrzmB staining showed donor-specific shifts, and Tim3 was impacted by autofluorescence in some donors. (h) Heatmap showing the median fluorescence intensities (MFI) for all the indicated molecules within CD8+ cytotoxic T cells (right) and CD4+

conv T cells (without CD25+ CD127- Tregs, left) across peripheral blood, OM, and HNSCC. This heatmap matches main Figure 1d, but shows MFIs instead of percentages. (i) Representative gating strategy for the identification of canonical antigen-presenting cell (APC) subsets in HNSCC. Plots are concatenated from three individual donors. (k) Staining patterns for all phenotyping markers in the high-dimensional APC panel shown on a representative HNSCC sample, pre-gated on live CD11c+ HLA-DR+ conventional DCs. Positivity cut-offs were left the same for all samples, except where shifts in staining patterns based on density distributions indicated adjustments. PD-L2 (on BV421) and CD85k (BV480) were excluded from all analyses because of significant variability due to autofluorescence between donors/experimental runs (data not shown). (m) Heatmap showing the median fluorescence intensities (MFI) for all the listed molecules within CD1c+ cDC2s/DC3s (right) and CD14+ cells (left). This heatmap matches main Figure 1g, but shows MFIs instead of percentages. Statistical analyses were performed using one-way ANOVA with Tukey's multiple comparisons.

Extended Data Fig. 2 Computational analysis using FAUST for the T cell and APC panels (related to main Figure 1).

(a) and (b) The top 10 T cell phenotypes showing differential abundance between OM (orange) and HNSCC (red) samples as identified by FAUST (FDR-adjusted level below 0.05). Population frequency is relative to the CD45+ live gate. Negative markers are not listed, all phenotypes are CD3+. (c) Example ridge plots of the marker distribution for T cell phenotype #9 (CD4+ CD3+ CD27+ CD69+ CD28+ HLADR+ PD1+ CD25+ ICOS+ CD38+ Tim3+) from a representative HNSCC sample. Grey histogram denotes marker distribution across all cells, green on the selected FAUST cluster. (d) and (e) A selection from the top 20 myeloid APC phenotypes showing differential abundance between OM (orange) and HNSCC (red) samples as identified by FAUST (FDR-adjusted level below 0.05). Population frequency is relative to the CD45+ live CD3- CD19- gate. Negative markers are not listed. (f) Simplifying the FAUST discovered Treg phenotype to just two markers, HLA-DR and ICOS. Summary plots show the relative frequency of ICOS+ HLADR+ cells (left) and ICOS+HLADR- cells (right) in the CD4+ CD25+ CD127- Treg compartment across blood,

OM and HNSCC ($n = 13$). (g) Simplifying the FAUST discovered APC phenotype to just two markers, CD40 and PD-L1. Summary plots show the relative frequency of CD40+ PD-L1+ cells on CD1c+ DCs (left) and CD14+ cells (right) in blood, OM and HNSCC ($n = 16$). For the FAUST box plots, the lower bound of the box is the 25th percentile (q25), center is the median, upper bound is the 75th percentile (q75). The lower whisker is 1.5*interquartile below the q25, and the upper whisker is 1.5*interquartile above q75. Y-axis is shown on a square-root scale. Statistical analyses in (f) and (g) were performed using one-way ANOVA with Tukey's multiple comparisons, and summary graphs are represented as mean \pm SD.

Extended Data Fig. 3 Gating strategy for the sorts and additional plots for the scRNA-seq data (related to main Figure 2).

(a) General gating of a representative OM tissue samples for CD45+ live singlets for the WTA 10x scRNA-seq experiments. (b) Gating strategy used for sorting the pan APC population and for pan T cells for WTA 10x scRNA-seq experiments. Red shaded gates were sorted. For some experiments (data not shown) MR1-Tetramer+ MAIT cells and CD56+ NK cells were sorted separately. (c) Representative re-analysis of a fraction of sorted pan T cells before loading onto the 10x Chromium controller. (d) UMAP plots of the combined scRNA-seq data after QC filtering (see Github script) and Harmony integration, colored by donor. A total of 139,424 cells is shown. Right plots depict individual donors separately, showing that cells from each donor distribute across the entire plot. Of note, for some donors only the T cell or the APC population could be sorted. (e) UMAP plots of the combined scRNA-seq data after QC filtering and Harmony integration, colored by simplified cell type calling derived from SingleR. The populations indicated in the legend were used for re-clustering and more in-depth analysis of the APC population (main Figure 2b) and T cells (Extended Data Fig. 5), respectively. (f) UMAP plots show heatmap overlays for CD3E and HLA-DRA transcripts to highlight the main lineages.

Extended Data Fig. 4 Additional analyses of the scRNA-seq data for the APC populations (related to main Figure 2).

(a) Heatmap overlays for the expression of key lineage transcripts on the UMAP plot for the APC population (see main Figure 2b). (b) Relative abundance of the APC clusters across donors and tissue origin. (c) Relative contribution of each tissue source to the indicated APC cluster (colors match the cluster description in main Fig. 2b). (d) Heatmaps showing the top 30 transcripts that were differentially expressed in HNSCC for the DC3 cluster (left) and cDC1 cluster (right), which are the two clusters showing the largest number of DE genes between OM and HNSCC (see main Figure 2e). (e) Heatmaps showing the top 30 transcripts that were shared between OM and HNSCC, but differentially expressed from matched peripheral blood, for each of the indicated APC clusters (mono classical, mono non-classical, cDC2s, DC3s, cDC1s and mreg DCs). For (d) and (e), the APC clusters identified in the scRNA-seq data (main Figure 2b) were subsetted, and either the genes shared between OM and HNSCC cells, or genes differentially expressed between HNSCC and OM cells were identified by the Seurat implementation of MAST (see material and methods and Github script for additional details).

Extended Data Fig. 5 Additional analyses of the scRNA-seq data for T cell populations (related to main Figure 2).

(a) UMAP plot of the T cell populations after subsetting and reclustering, colored by cluster. (b) Relative cluster abundance across donors and tissue origin (color code is the same as in panel a) (c) Number of DE genes between HNSCC and OM-derived cells per T cell cluster as determined by MAST. (d) Heatmaps showing the top 30 transcripts that were shared between OM and HNSCC, but differentially expressed from matched peripheral blood, for each of the indicated T cell clusters. (e) Heatmaps showing the top 30 transcripts that were differentially expressed in HNSCC for the T cell clusters 4–6, which were the two clusters showing the largest number of DE genes between OM and HNSCC. (f) Heatmap overlays showing the indicated lineage transcripts and T cell scores (see material and methods) on the UMAP plot. (g) Violin plots depicting the relative T cell score in the indicated HNSCC (red) vs OM-derived (orange) T cell clusters.

Extended Data Fig. 6 Additional plots related to the NicheNet predictions (related to main Figure 3).

(a) Plots derived from the NicheNet analysis showing the predicted ligand activity (orange) for the top ligands (as ranked by Pearson correlation coefficient), and the predicted target genes (purple) for all three separate NicheNet runs: sender APCs + receiver cells CD4+ conv T cell clusters (upper left plots), sender APCs + receiver CD8+ T cell clusters (upper right plots) and sender APCs + receiver CD4+ Treg cluster (lower plots). The full script utilized for the NicheNet analysis is available on Github (see material and methods). The ligands that are highlighted in main Figure 3b as “interesting” are highlighted in red/bold in this panel here. (b) The APC clusters from main Figure 2b were subsetted to contain only OM or HNSCC-derived cells, and Violin plots depict relative expression of the indicated ligand transcripts across all APC clusters either in OM-infiltrating cells (left columns) or HNSCC infiltrating cells (right columns). The ligands here match the ones highlighted as “interesting” in main Figure 3b. (c) Representative plots showing the protein expression for the cytokines IL-1 α and IL-1 β after ex vivo culture of bulk HNSCC-derived APCs in the presence of Brefeldin A, followed by intracellular cytokine staining. (d) Concentration of IL-1 α and IL-1 β and IL-18 as measured by Luminex analysis in flash-frozen HNSCC samples (n = 4). LOD: limit of detection.

[Extended Data Fig. 7 Additional data from Treg suppression and stimulation assays \(related to main Figure 4\).](#)

(a) Proliferation of HNSCC-derived CD4+ T responder cells from tumor (left) or from blood (right) in an in vitro suppression assay with IL-1R1-Tregs (light red) and IL-1R1+ Tregs (dark red) from tumor. Representative histograms show dilution of Cell Trace Violet (CTV) after 4 days. n = 4. (b) Concentration of Granzyme B, IL-2 and IFN-g in the culture supernatants of the 4-day suppression assays shown in main Figure 4a. (c) Proliferation of HNSCC-derived CD8+ T responder cells in an in vitro suppression assay with titrated amounts of IL-1R1- Tregs (left plot) and IL-1R1+ Tregs (right plot). Histograms show CTV dilution after 4 days of culture with Treg-to-Teff ratios of 1:4, 1:2 and 1:1 (top to bottom). (d) Plots depict representative post-sort purity of HNSCC-Treg populations used for the stimulation experiments in main Figure 4b. (e) Expression of IL-1R1 after two days of in vitro culture either unstimulated or in the presence of Gibco anti-CD3/CD28 beads for Tregs sorted from peripheral blood (blue), IL-1R1-

(light red) and IL-1R1+ Tregs (dark red) from HNSCC. n = 3. (f) Volcano plots showing differential gene expression of SMART-Seq bulk RNAseq data (250 sorted cells) of blood Tregs, HNSCC IL-1R1- and IL-1R1+ after 2 days of culture with anti-CD3/CD28/CD2 beads with (light and dark red) or without IL-1 (grey). (g) Transcripts per million for FOXP3, Helios (IKZF2) and CTLA4 from the bulk RNAseq data for the indicated time points and culture conditions. n = 4 for the d2 time point, n = 6 for the d1 time point. (h) CD25 median fluorescence intensity (MFI) on the indicated Treg populations after 2 days of culture with anti-CD3/CD28/CD2 beads +/- IL-1. n = 3 for blood and n = 4 for the HNSCC samples. (i) Volcano plots showing differential gene expression of SMART-Seq bulk RNAseq data (250 sorted cells) of blood Tregs and HNSCC IL-1R1+ Tregs after 2 days either unstimulated or with anti-CD3/CD28/CD2 beads. Number of DE genes upregulated is highlighted in bold.

[Extended Data Fig. 8 Additional analyses for the VDJ and targeted transcriptomics/Abseq data \(related to main Figure 4\).](#)

(a) Experimental outline for the targeted transcriptomics+AbSeq experiments (see also material and methods). Combined data set after Harmony integration consists of ex vivo CD45+ cells, unstimulated T cells, and PMA-Ionomycin stimulated T cells from two different donors. (b) UMAP plot of tumor-infiltrating T cells of two different HNSCC donors after targeted transcriptomics and AbSeq as described in (a), colored by clustering based on transcript. The highlighted clusters are shown in more detail in main Figure 4c. (c) Selected cytokine transcripts are shown on manually gated T cell subsets from HNSCC AbSeq data (populations as identified by surface protein). Red plots show CD8+ T cells, violet plots show CD25+ CD127- Tregs, and blue plots show CD25- CD4+ non Tregs (all after PMA/Iono for 2h). TBX21 encodes for T-bet, TNFRSF9 encodes for CD137 (4-1BB). (d) t-SNE plot of sorted Tregs from the peripheral blood (top) and tumor (bottom) from a single donor after VDJ sequencing (10x genomics 5' v1 chemistry). All cells with a complete TCR sequence are marked in light blue, and the top 10 clonotypes are marked in dark blue (encompassing 20 cells in the blood, and 239 cells in the tumor). (e) Table showing the total cell counts, barcode counts and detected clonotypes after VDJ sequencing for the indicated samples (n = 3 for the HNSCC tumor

samples, n = 2 for matched peripheral blood). Counts are derived from VDJ Loupe browser. (f) Expanded clones by single-cell VDJ sequencing within sorted IL-1R1+ Tregs from HNSCC tumors relative to total Tregs from matched peripheral blood. Every TCR sequence that was present in 2 cells or more was considered an expanded clone. (n = 2 for blood, n = 3 for HNSCC) (g) Differentially expressed (DE) genes for the top 3 expanded clones from one single donor. Heatmaps show a selection of the top DE genes as identified by MAST for each clone (right side of each heatmap) vs downsampled HNSCC IL1R1- Tregs and IL1R1+ Tregs.

Extended Data Fig. 9 IL-1R1 expression in other cancer types (related to main Figure 4).

(a) IL-1R1 expression is present on CD4+ CD25+ CD127- Tregs in tumor tissues from a cohort of human papillary carcinoma patients. Plot on the very right shows quantification of the MFI for CTLA-4 on the indicated populations (n = 4 donors, one-way ANOVA with Tukey's multiple comparisons). (b) To assess IL-1R1 protein expression in additional tumor types, we were able to collect a single sample of human breast cancer (left) and human lung cancer (right). IL-1R1 expression is present on breast cancer Tregs, but only at very low level/absent on lung cancer Tregs. (c) UMAP plot of all CD4 T cell metaclusters. There are four Treg clusters, which are highlighted on the right side of the UMAP plot. Heatmap overlays on the right side show expression of the transcripts encoding FOXP3 and IL1R1. (d) Violin plots depicting expression of FOXP3 (top panel) and IL1R1 (lower panel) across all CD4 T cell clusters of the combined data set as annotated by Zheng et al. Shaded area highlights the Treg clusters. IL1R1 expression is primarily found in cluster 20 and 21. (e) Violin plots depicting expression of IL1R1 in cells from cluster 20 CD4.c20.Treg.TNFRSF9) across different tumor types. While expression appears to be absent in AML and BCL, varying levels of IL1R1 transcript are present across all other tumor types. (f) IL1R1 transcript expression in scRNA-seq data may underestimate protein expression. Left panel shows violin plots for expression of the indicated transcripts in the Treg cluster 5 of our combined scRNaseq data from OM and HNSCC (see Extended Data Fig. 5a), right histograms show protein expression for the same genes by flow cytometry on live CD4+ CD25+ CD127- Tregs on a representative HNSCC sample.

Data in panels c-e are publicly available from the pan-cancer T cell atlas (containing 21 cancer types), Zheng et al, Science 2021 (ref. [39](#)). Figures were generated using the Shiny app on http://cancer-pku.cn:3838/PanC_T. Per the authors' description, box plot overlays show 25th percentile (lower bound), median (center) and 75th percentile (upper bound), whiskers extend +/- 1.5*interquartile range. Data in panel a-b and f are from our own tissue collections.

Extended Data Fig. 10 Differential expression kinetics of IL-1R1 in human cells, murine cells and a humanized mouse model.

(a) Overview of stimulation experiments for murine (left) and human (right) purified Treg populations. (b) Murine thymic $\gamma\delta$ T cells (lower plot) were used as a positive control to validate the anti-mouse IL-1R1-PE antibody signal. (c) Representative plots depicting IL-1R1 and CD69 expression on unstimulated (black) and TCR stimulated (red) Tregs after 48 h. (d) Quantification of IL-1R1 expression after stimulation for 1 or 2 days in murine (top panel, n = 3) and human Tregs (bottom panel, n = 5), highlighting the discrepancy in expression level and kinetics. (e) Overview and timeline of human SCC15-tumor experiment in humanized MISTRG mice. Bottom, photograph of 5 tumors after collection. (f) Representative flow cytometry plots showing similar expression patterns for a set of key T cell markers in primary human HNSCC biopsies (top) and SCC15 tumors in MISTRG mice reconstituted with human immune cells (bottom). (g) Representative plots and quantification showing a similar increase in Treg frequencies between human blood/HNSCC tumor tissue (top) and humanized mouse blood/SCC15 tumor tissue (bottom). n = 6 for the humanized mouse samples, and n = 14 for the human HNSCC samples. (h) Representative plots and quantification showing that IL-1R1 expression is detectable, but under-represented in the humanized mouse model (n = 5) compared to primary human HNSCC tissue (see main Figure [3c](#)). Error bars represent mean +/- SD. Statistical analyses were performed using one-way ANOVA with Tukey's multiple comparisons test for (h) or using a two-tailed paired t-test for (g).

Supplementary information

Supplementary Information

This file contains Supplementary Tables 2–4 and Figs. 1, 2.

Reporting Summary

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Mair, F., Erickson, J.R., Frutoso, M. *et al.* Extricating human tumour immune alterations from tissue inflammation. *Nature* **605**, 728–735 (2022). <https://doi.org/10.1038/s41586-022-04718-w>

- Received: 23 March 2021
- Accepted: 01 April 2022
- Published: 11 May 2022

- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04718-w>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04718-w>

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

[Download PDF](#)

- Article
- Open Access
- [Published: 04 May 2022](#)

Intermittent PI3K δ inhibition sustains anti-tumour immunity and curbs irAEs

- [Simon Eschweiler](#) ORCID: [orcid.org/0000-0002-6379-8588](#)¹,
- [Ciro Ramírez-Suástegui](#)¹,
- [Yingcong Li](#) ORCID: [orcid.org/0000-0002-0746-0531](#)^{1,2},
- [Emma King](#)^{3,4},
- [Lindsey Chudley](#)⁵,
- [Jaya Thomas](#)³,
- [Oliver Wood](#)³,
- [Adrian von Witzleben](#) ORCID: [orcid.org/0000-0001-6766-4758](#)^{3,6},
- [Danielle Jeffrey](#) ORCID: [orcid.org/0000-0003-3193-3420](#)³,
- [Katy McCann](#) ORCID: [orcid.org/0000-0003-4388-7265](#)³,
- [Hayley Simon](#)¹,
- [Monalisa Mondal](#)¹,
- [Alice Wang](#)¹,
- [Martina Dicker](#)¹,
- [Elena Lopez-Guadamillas](#)⁷,
- [Ting-Fang Chou](#) ORCID: [orcid.org/0000-0001-7095-3518](#)¹,
- [Nicola A. Dobbs](#)⁸,
- [Louisa Essame](#)⁸,
- [Gary Acton](#)⁸,
- [Fiona Kelly](#)⁸,
- [Gavin Halbert](#)⁹,

- [Joseph J. Sacco](#)^{5,10},
- [Andrew Graeme Schache](#) ORCID: [orcid.org/0000-0001-9466-6038](#)^{5,11},
- [Richard Shaw](#) ORCID: [orcid.org/0000-0002-5157-4042](#)^{5,11},
- [James Anthony McCaul](#)¹²,
- [Claire Paterson](#)¹³,
- [Joseph H. Davies](#)⁴,
- [Peter A. Brennan](#)¹⁴,
- [Rabindra P. Singh](#)¹⁵,
- [Paul M. Loadman](#) ORCID: [orcid.org/0000-0002-4259-8616](#)¹⁶,
- [William Wilson](#)¹⁷,
- [Allan Hackshaw](#) ORCID: [orcid.org/0000-0002-5570-5070](#)¹⁷,
- [Gregory Seumois](#) ORCID: [orcid.org/0000-0002-8164-6852](#)¹,
- [Klaus Okkenhaug](#) ORCID: [orcid.org/0000-0002-9432-4051](#)¹⁸,
- [Gareth J. Thomas](#)³,
- [Terry M. Jones](#)^{5,11},
- [Ferhat Ay](#) ORCID: [orcid.org/0000-0002-0708-6914](#)¹,
- [Greg Friberg](#)¹⁹,
- [Mitchell Kronenberg](#) ORCID: [orcid.org/0000-0001-6318-6445](#)^{1,2},
- [Bart Vanhaesebroeck](#) ORCID: [orcid.org/0000-0002-7074-3673](#)⁷,
- [Pandurangan Vijayanand](#) ORCID: [orcid.org/0000-0001-7067-9723](#)^{1,5,20 na1} &
- [Christian H. Ottensmeier](#) ORCID: [orcid.org/0000-0003-3619-1657](#)^{1,3,5,10 na1}

Nature volume 605, pages 741–746 (2022)

- 18k Accesses
- 254 Altmetric
- [Metrics details](#)

Subjects

- [Head and neck cancer](#)
- [Immunotherapy](#)

Abstract

Phosphoinositide 3-kinase δ (PI3Kδ) has a key role in lymphocytes, and inhibitors that target this PI3K have been approved for treatment of B cell malignancies^{1,2,3}. Although studies in mouse models of solid tumours have demonstrated that PI3Kδ inhibitors (PI3Kδi) can induce anti-tumour immunity^{4,5}, its effect on solid tumours in humans remains unclear. Here we assessed the effects of the PI3Kδi AMG319 in human patients with head and neck cancer in a neoadjuvant, double-blind, placebo-controlled randomized phase II trial (EudraCT no. 2014-004388-20). PI3Kδ inhibition decreased the number of tumour-infiltrating regulatory T (T_{reg}) cells and enhanced the cytotoxic potential of tumour-infiltrating T cells. At the tested doses of AMG319, immune-related adverse events (irAEs) required treatment to be discontinued in 12 out of 21 of patients treated with AMG319, suggestive of systemic effects on T_{reg} cells. Accordingly, in mouse models, PI3Kδi decreased the number of T_{reg} cells systemically and caused colitis. Single-cell RNA-sequencing analysis revealed a PI3Kδi-driven loss of tissue-resident colonic ST2 T_{reg} cells, accompanied by expansion of pathogenic T helper 17 (T_H17) and type 17 CD8⁺ T (T_C17) cells, which probably contributed to toxicity; this points towards a specific mode of action for the emergence of irAEs. A modified treatment regimen with intermittent dosing of PI3Kδi in mouse models led to a significant decrease in tumour growth without inducing pathogenic T cells in colonic tissue, indicating that alternative dosing regimens might limit toxicity.

Main

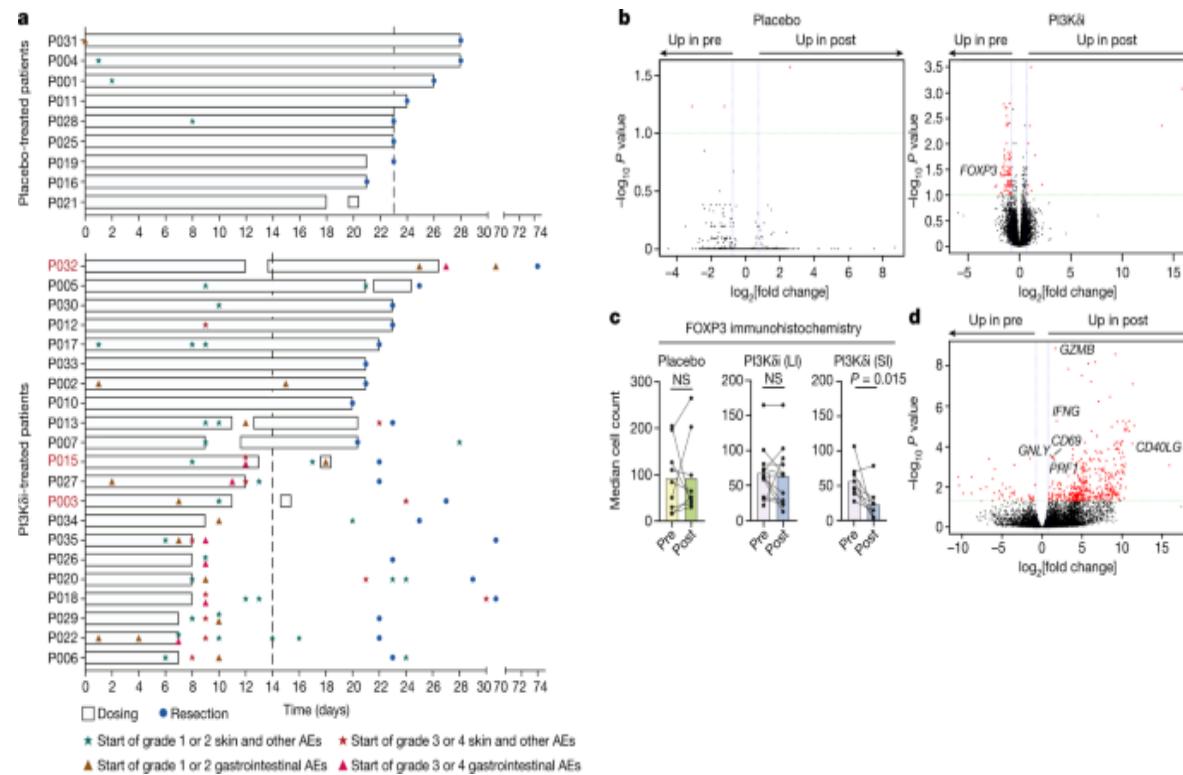
PI3K inhibitors were initially considered to target mainly PI3K activity intrinsic to cancer cells, which was the underlying rationale for testing inhibitors against the leukocyte-enriched PI3Kδ in B cell malignancies. However, subsequent studies have shown that PI3Kδ inhibition also has clear immunomodulatory activities, largely T cell-mediated, that were under-appreciated at the time of the early trials in B cell malignancies, causing irAEs that have hampered clinical progress and utility. Several lines of evidence suggest that PI3Kδi preferentially inhibit T_{reg} cells over other T

cell subsets^{4,5,6,7,8} but so far, no trials have been performed to explicitly explore this concept in humans. Here we provide an in-depth investigation of the effect of PI3K δ inhibition on immune cells in patients with solid tumours and also explore the mechanism that leads to irAEs.

PI3K δ inhibition causes irAEs

To evaluate the potential of PI3K δ i as immunotherapeutic agents in human solid cancers, we administered the PI3K δ i AMG319 to treatment-naive patients with resectable head and neck squamous cell carcinoma (HNSCC) in a neoadjuvant, double-blind, placebo-controlled randomized phase II trial (Extended Data Fig. [1a, b](#)). We measured target inhibition (using phosphorylated AKT (pAKT) levels in B cells) (Extended Data Fig. [1c](#)) and drug levels (Extended Data Fig. [1d](#)) to verify drug administration. Thirty-three patients were randomized in a 2:1 ratio (AMG319:placebo) for the trial and 30 patients received at least one dose of AMG319 or placebo. Fifteen patients received 400 mg daily of AMG319 (range of 7–24 days per patient). Unexpectedly, at the 400 mg dose, 9 out of 15 patients experienced irAEs that led to withdrawal of treatment. After a formal safety review, 6 additional patients were recruited and treated at a reduced dose of 300 mg per day. Again, three out of six patients had irAEs that led to discontinuation of treatment. One patient experienced grade 4 colitis after completion of 24 daily doses of AMG319 and eventually required colectomy (Fig. [1a](#)). The most prevalent irAEs were skin rashes (29%; 25% in the treatment group and 4% in placebo group), diarrhea (29%; 28% in the treatment group and 1% in placebo group) and transaminitis (14% all in the treatment group), consistent with a treatment-mediated loss of T_{reg} cells or T_{reg} cell functionality in multiple tissues causing immunopathology. The onset of irAEs was surprisingly rapid (median time to onset of 9 days) and led to treatment discontinuation in 12 out of 21 AMG319-treated patients. Clinically, and most probably reflecting the brief treatment period, we did not observe any significant differences in the measured tumour volumes between the study arms in the 23 patients in whom this was evaluable. Two patients with partial responses and one with complete pathological response were among the AMG319-treated patients (Extended Data Fig. [1c](#)), all of whom also exhibited grade 3/4 irAEs.

Fig. 1: PI3K δ i drives anti-tumour immunity but causes significant irAEs.



a, Swimlane plot depicting treatment regimen, intervals and occurrence and grade of irAEs in PI3K δ i-treated (top) and placebo-treated (bottom) patients; patients with partial response or complete pathological response are highlighted in magenta. Vertical dashed lines show average duration of treatment. **b, d**, Volcano plots of whole-tumour RNA-seq analysis (**b**) or bulk RNA-seq analysis of purified tumour-infiltrating CD8 $^{+}$ T cells (**d**) comparing patients treated with AMG-319 to those treated with placebo. DEGs between pre- and post-treatment samples are highlighted in red and were called by DEseq2; adjusted P values were calculated with the Benjamini–Hochberg method. Depicted are transcripts that changed in expression by more than 0.75-fold and had an adjusted P value of ≤ 0.1 (**b**) or < 0.05 (**d**). **c**, Median cell count of FOXP3 $^{+}$ cells in pre- and post-treatment samples of placebo- or AMG319-treated patients. AMG-319-treated patients have been further stratified into patients for whom the interval between stopping of treatment and immunohistochemistry assessment was more than four days (long interval (LI)) or less than one day (short interval (SI)). $P = 0.015$ for SI. Data are mean \pm s.e.m.; two-tailed Wilcoxon matched-pairs signed rank test (**c**). Differential expression

analysis (**b**, **d**) was performed using DESeq2 (v1.24.0). AE, adverse event; NS, not significant.

[Source data](#)

PI3K δ i alter the tumour microenvironment

Whole-tumour RNA-sequencing (RNA-seq) analysis of pre- and post-treatment tumour samples revealed substantial differences in the AMG319 treatment group (93 differentially expressed genes (DEGs)), but not for the placebo group (3 DEGs) (Fig. [1b](#)). As PI3K δ inhibition led to a significant reduction in *FOXP3* transcript levels in the tumour samples (Fig. [1b](#)), we assessed T_{reg} cell levels in tumour tissue by immunohistochemistry, hypothesizing that the duration between cessation of treatment and tumour resection might be a critical factor influencing T_{reg} cell abundance, owing to the relatively short half-life of the compound. Indeed, we found significantly reduced intratumoural T_{reg} cells only in patients in which their abundance could be assessed directly after treatment (PI3K δ i short interval) (Fig. [1c](#)), implying that T_{reg} cell levels normalize quickly once treatment has been stopped.

Bulk RNA-seq analysis of sorted tumour-infiltrating CD8⁺ T cells revealed higher expression of *IFNG*, *GZMB* and *PRF1* in post-treatment samples, indicating enhanced cytotoxic potential of tumour-infiltrating CD8⁺ T cells following PI3K δ i treatment (Fig. [1d](#)). We corroborated these results by single-cell RNA-seq (scRNA-seq) analysis, which demonstrated that CD4⁺ and CD8⁺ T cell clusters showed a treatment-associated increase in expression of cytotoxicity genes (for example, *GZMB* and *PRF1*) (Extended Data Fig. [2a,b](#)). We also found a modest clonal expansion of CD4⁺ and CD8⁺ T cells after treatment (Extended Data Fig. [2c](#)). As low cell numbers of CD4⁺FOXP3⁺ T cells (0–27 cells per patient) precluded a more detailed analysis in our cohort, we next assessed circulating T_{reg} cells. PI3K δ inhibition led to a significant increase in activated circulating T_{reg} cells, while the proportion and activation status of T_{reg} cells in the placebo group remained stable (Extended Data Fig. [2d,e](#)). This implies that PI3K δ

inhibition either influences proliferation or displaces activated T_{reg} cells from tissues, presumably by altering the expression of tissue homing factors such as KLF2 and S1PR1 (direct targets of FOXO1 in line with previous studies^{5,6,7}), probably contributing to toxicity. Together, these data indicate that PI3K δ inhibition causes profound changes in the tumour microenvironment (TME), characterized by enhanced CD4 $^{+}$ and CD8 $^{+}$ T cell activation, oligoclonal T cell expansion and increased cytolytic activity, consistent with a decrease in intratumoural T_{reg} cells, enabling T cell activation and leading to a rapid onset of dose-limiting toxicity.

Systemic effects of PI3K δ i on T_{reg} cells

Next, to understand the mechanistic basis of PI3K δ i-induced toxicity and anti-tumour immune responses, we tested the effect of a PI3K δ inhibitor in a mouse solid tumour model. We inoculated wild-type C57BL/6 mice with B16F10-OVA melanoma cells and treated them with the PI3K δ i PI-3065⁷. Consistent with previous studies^{4,5}, we found a significant decrease in tumour volume (Extended Data Fig. 3a) and a significant increase in the number of intratumoural CD8 $^{+}$ T cells that expressed high levels of PD-1 and exhibited increased proliferative and cytotoxic capacity (Extended Data Fig. 3b–e). TOX, a transcription factor recently identified as critical for the adaptation and survival of CD8 $^{+}$ T cells in the TME⁹, was also increased after PI3K δ i treatment (Extended Data Fig. 3f). Notably, and contrary to previous reports^{10,11}, we found that the expression of both granzyme B and Ki67 was almost exclusively limited to TOX $^{+}$ CD8 $^{+}$ T cells (Extended Data Fig. 3g), demonstrating that these cells, despite showing high expression of PD-1 and TOX, are not functionally exhausted in this tumour model.

Given that PI3K inhibitors were initially considered to target mainly cancer cell-intrinsic PI3K activity, we used *Rag1* $^{-/-}$ and *Cd8* $^{-/-}$ mice to verify that the observed anti-tumour effects were dependent on immune cells and, more specifically, on CD8 $^{+}$ T cells (Extended Data Fig. 4h). As PI3K δ inhibition caused substantial irAEs in non-malignant organs (Fig. 1b), and given that T_{reg} cells have been shown to be susceptible to this form of treatment, we next assessed whether PI3K δ i act locally within the tumour tissue or

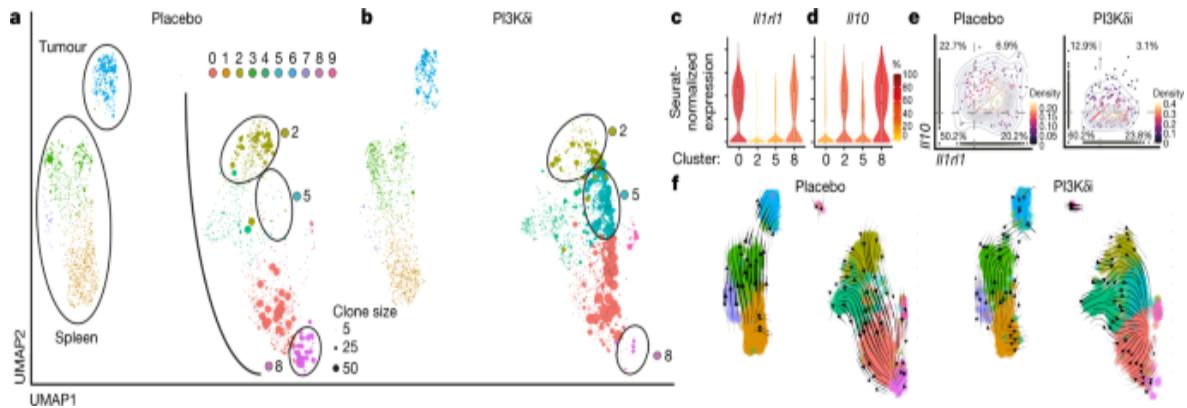
systemically. Of note, in PI3K δ i-treated mice, but not placebo-treated control mice, we found a significant decrease in T_{reg} cells in tumour, spleen and colon, indicative of systemic effects of PI3K δ i on T_{reg} maintenance or survival (Extended Data Fig. 3*i–k*).

PI3K δ i affect specific T_{reg} cell subsets

As gastrointestinal toxicity is one of the major irAEs in patients receiving PI3K δ i^{4,6,12} (Fig. 1*b*), we hypothesized that T_{reg} cells present in colonic tissue may be especially sensitive to PI3K δ i. To test this hypothesis in an unbiased manner, we performed scRNA-seq of T_{reg} cells isolated from tumour, spleen (lymphoid organ) and colonic tissue of PI3K δ i- and placebo-treated B16F10-OVA tumour-bearing mice. UMAP analysis identified 10 T_{reg} cell clusters, implying substantial T_{reg} cell heterogeneity and tissue-dependent adaptations (Fig. 2*a, b*); this supports the notion that several distinct T_{reg} subtypes exist in different locations (Extended Data Fig. 4*a, b*), in agreement with previous studies^{13,14}. Colonic T_{reg} cells exhibited the most pronounced differences between PI3K δ i and placebo treatment, with 869 DEGs, whereas splenic and tumour T_{reg} cells exhibited fewer differences (Extended Data Fig. 4*c–e*). Two of the colonic T_{reg} subsets (clusters 2 and 8) were depleted in PI3K δ i-treated mice (Fig. 2*a, b*, Extended Data Fig. 4*f*). Cluster 2 colonic T_{reg} cells were enriched for the expression of *Ctla4* and genes encoding chemokine receptors (*Ccr1*, *Ccr2* and *Ccr4*), which are critical for their suppressive^{15,16} and migratory¹⁷ capacities, respectively (Extended Data Fig. 4*g*). Cluster 8 colonic T_{reg} cells—which showed substantial clonal expansion in control-treated mice but were depleted in PI3K δ i-treated mice—resembled the recently described tissue-resident ST2 T_{reg} cells^{18,19,20}, which are critical for protection against chronic inflammation and facilitation of tissue repair (Fig. 2*a, b*). Accordingly, we found enrichment in the expression of the ST2 T_{reg} signature genes *Il1rl1* (which encodes IL1RL1 (also known as IL-33R or ST2)), *Gata3* and *Id2*, as well as several genes associated with highly suppressive effector T_{reg} cells (*Klrg1*, *Cd44*, *Cd69*, *Pdcld1*, *Areg*, *Nr4a1*, *Il10* and *Tgfb1*) (Extended Data Fig. 5*a*). We verified ST2 expression on T_{reg} cells at the protein level and

found that PI3K δ inhibition led to a substantially increased ratio of CD8 $^{+}$ T cells to ST2 T_{reg} cells (Extended Data Fig. 5b). Whereas colonic T_{reg} cells in cluster 0 and cluster 8 shared this ST2 signature (Fig. 2c), only cells in cluster 8 showed high transcript expression of the immunosuppressive cytokine IL-10 (Fig. 2d). These T_{reg} cell clusters (2 and 8) with highly suppressive properties were depleted in PI3K δ i-treated mice, whereas the clonally expanded cluster 5 T_{reg} cells were enriched in PI3K δ i-treated mice showed a lack of transcripts associated with suppression (Fig. 2e, Extended Data Fig. 5c) and higher expression of several interferon-related response genes^{21,22} (*Stat1*, *Stat3* and *Ifird1*), suggestive of a pro-inflammatory environment (Extended Data Fig. 5a). Accordingly, ST2 $^{+}$ IL10 $^{+}$ T_{reg} cells were substantially reduced in PI3K δ i-treated mice (Fig. 2e). Notably, RNA velocity analysis, a tool to assess the developmental stage of cells in scRNA-seq data^{23,24}, infers a developmental trajectory over several progenitor states (cluster 2,4 and then 0) culminating in clonally expanded ST2 T_{reg} cells (cluster 8) in placebo-treated mice (Fig. 2f). These data indicate that PI3K δ inhibition prevents the cellular differentiation into ST2 T_{reg} cells, and instead diverts development to cluster 5 T_{reg} cells that lack expression of transcripts associated with suppressive capacity, pointing to a possible mechanism for the onset of inflammation and colitis. We also observed a significant increase in CD8 $^{+}$ T cells in colonic but not splenic tissue (Extended Data Fig. 5d,e). Colonic CD8 $^{+}$ T cells expressed higher levels of PD-1 and ICOS upon PI3K δ inhibition (Extended Data Fig. 5f,g), implying treatment-related changes in cell activation. Together, these findings suggest a heightened sensitivity of certain colonic T_{reg} subsets to PI3K δ i, potentially related to the high incidence of colitis observed in patients treated with PI3K δ i.

Fig. 2: PI3K δ i affects distinct T_{reg} cell subtypes.



a, b, Uniform manifold approximation and projection (UMAP) plots single-cell transcriptomes and T cell receptor (TCR) sequence data of FOXP3⁺CD4⁺ T cells in placebo-treated control mice (**a**; $n = 3$ mice) and PI-3065-treated mice (**b**; $n = 3$ mice). Circle size indicates degree of clonal expansion. **c, d**, Violin plots showing Seurat-normalized expression levels (colour scale depicts percentage of expressing cells) of highlighted genes in the indicated clusters from **a, b**. The centre line depicts the median, edges delineate the 25th and 75th percentiles and whiskers extend to minimum and maximum values. **e**, Scatter plots showing Seurat-normalized expression levels of highlighted genes in colonic T_{reg} cells in placebo-treated and PI-3065-treated mice. The dashed line indicates the expression cut-off; numbers indicate the frequency in each quadrant. **f**, RNA velocity analysis visualized by UMAP, depicting likely developmental trajectories of T_{reg} cells from **a, b**. Arrows indicate velocity streamlines.

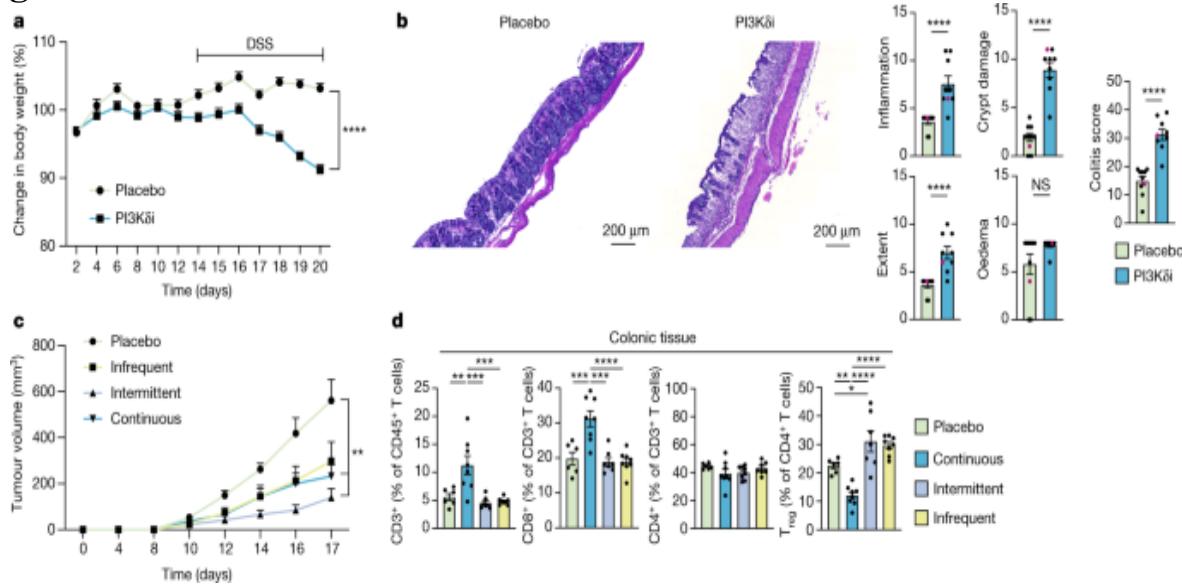
[Source data](#)

PI3K δ inhibition exacerbates colitis

To explore the connection between PI3K δ inhibition and gastrointestinal toxicity in more detail, we used a dextran sulfate sodium-induced acute colitis model. Crucially, when compared with placebo-treated mice, we found that mice treated with PI3K δ i showed an accelerated and exacerbated disease phenotype, with a swift reduction in body weight and a higher overall colitis score characterized by significantly higher inflammation, crypt damage and area of infiltration (extent) (Fig. 3a, b), which indicate treatment-mediated alterations in tissue homeostasis driving

immunopathology. To circumvent the emergence of these irAEs, we hypothesized that a transient depletion of T_{reg} cells might suffice to restrict the immunosuppressive milieu in the tumour and thus drive anti-tumour immunity without causing substantial toxicity in non-malignant organs. We tested this hypothesis by using distinct treatment regimens, on which mice would either be kept on PI3K δ i for the duration of the experiment (continuous dosing), be kept on PI3K δ i for 4 days followed by 3 days off drug (intermittent dosing) or be kept on PI3K δ i for 2 days followed by 5 days off drug (infrequent dosing) for a total of two treatment cycles (Fig. 3c). All treatment conditions led to a decrease in tumour growth, albeit not significantly for the infrequent dosing condition, suggesting that transient interruptions of the immunosuppressive TME drive anti-tumour immunity. Most importantly, only continuous dosing led to increased CD8 $^+$ T cell infiltration and decreased T_{reg} cell levels in colonic tissue (Fig. 3d), indicating that intermittent dosing regimens might also decrease irAEs in human.

Fig. 3: PI3K δ i exacerbates colitis.



a, Mice were fed either a control diet or a diet containing PI-3065 for the duration of the experiment and were additionally treated with 2.5% dextran sulfate sodium (DSS) from day 14 to day 20. Change in body weight is shown relative to body weight before treatment on day 0. $n = 10$ mice per group, $P < 0.0001$. **b**, Representative sections from haematoxylin and eosin (H&E) histology scans and colitis scoring from zinc-formalin-fixed colonic

tissue sections from mice treated with placebo or PI3K δ i in **a**. $n = 10$ mice (placebo) and $n = 9$ (PI3K δ i) (one mouse died before the experimental endpoint); $P < 0.0001$ for inflammation, extent, crypt damage and overall colitis scoring; representative samples from the H&E staining are highlighted in magenta. **c, d**, Mice were inoculated subcutaneously with B16F10-OVA cells and fed either a control diet or a diet containing PI-3065. Infrequent dosing, PI3K δ i for 2 days followed by 5 days off drug; intermittent dosing, PI3K δ i for 4 days followed by 3 days off drug; continuous dosing, PI3K δ i for the duration of the experiment. Tumour volume (**c**) and flow-cytometric analyses of cell frequencies (**d**) of mice treated as indicated. $n = 6$ mice (placebo), $n = 7$ mice (intermittent dosing), $n = 8$ mice (continuous dosing) and $n = 8$ mice (infrequent dosing). Placebo versus intermittent dosing (**c**), $P = 0.0023$; placebo versus continuous dosing (**c**), $P = 0.0059$; placebo versus continuous dosing, $P = 0.003$; continuous dosing versus intermittent dosing and infrequent dosing (left), $P = 0.0003$; placebo versus continuous dosing, $P = 0.0005$; for continuous dosing versus intermittent dosing, $P = 0.0001$; placebo versus infrequent dosing (**d**; third from left), $P < 0.0001$; placebo versus continuous dosing, $P = 0.0086$; placebo versus intermittent dosing, $P = 0.045$; continuous dosing versus intermittent dosing and infrequent dosing, $P < 0.0001$. Data are mean \pm s.e.m.; two-tailed Mann–Whitney test (**a–c**) or one-way ANOVA comparing the mean of each group with the mean of each other group followed by Dunnett’s test (**d**). Data are representative of at least two independent experiments.

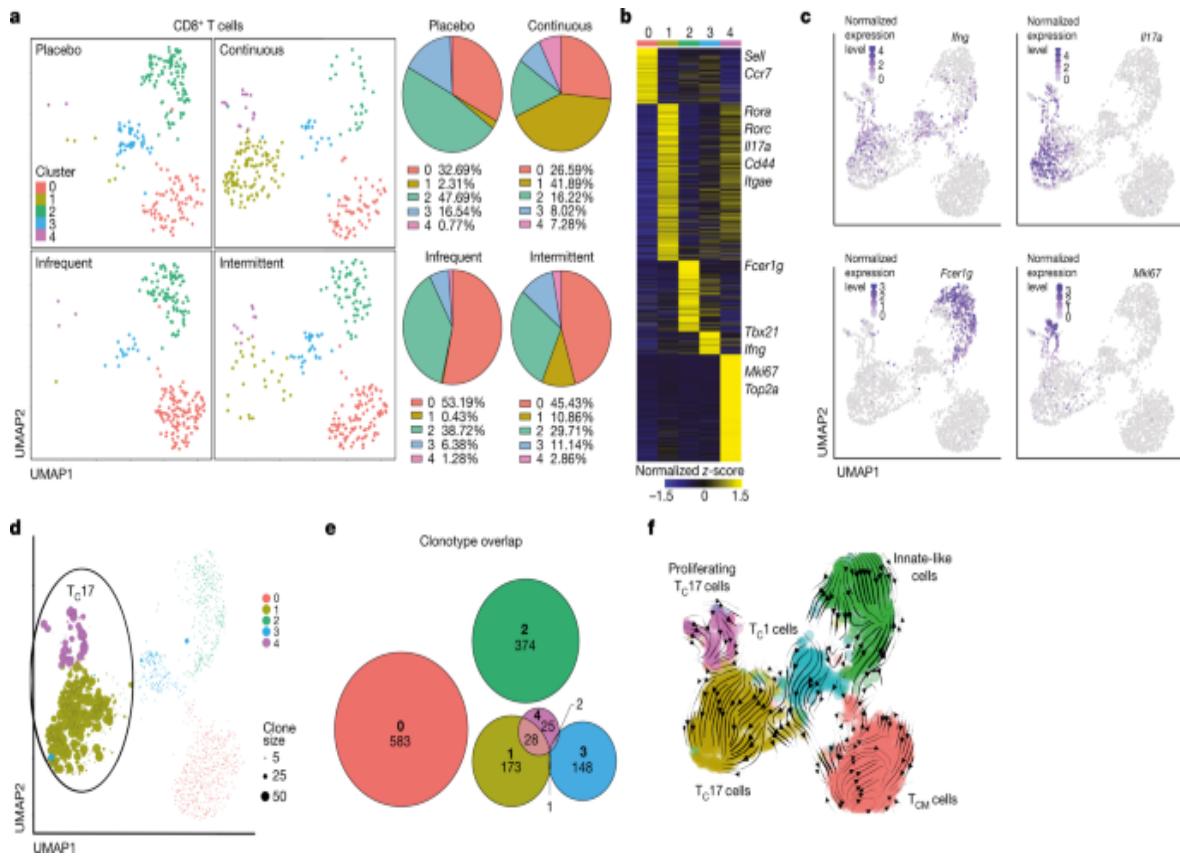
[Source data](#)

Intermittent dosing curbs toxicity

To discern whether specific T cell subsets drive immunopathology on PI3K δ inhibition, we performed scRNA-seq of colonic CD8 $^+$ and CD4 $^+$ T cells in the different treatment regimens. Unbiased clustering depicted by UMAP revealed five distinct CD8 $^+$ cell clusters and six distinct CD4 $^+$ T cell clusters (Fig. [4a](#), Extended Data Fig. [6a](#)). In both instances, we identified a central memory T (T_{CM}) cell subset (cluster 0, red) expressing high levels of *Ccr7* and *Cd62L*, a T_{C1} and T_{H1} subset expressing high levels of interferon- γ

transcripts (cluster 3, blue for T_C1 and cluster 1, ochre for T_H1), a T_C17 and T_H17 subset enriched for Il-17 transcripts (cluster 1, ochre for T_C17 and cluster 2, green for T_H17), and a proliferative subset that exhibited features of T_C17 (violet) or T_H17 (blue) cells, respectively (Fig. 4b,c, Extended Data Fig. 6b,c). Notably, we found dosing-dependent enrichment of the T_C17 and T_H17 subsets and pertaining proliferating clusters, making up approximately 50% of all cells in the continuous dosing regimen, whereas they were nearly completely absent in the other treatment conditions (Fig. 4a, Extended Data Fig. 6a). Of note, IL-17 producing cells have been shown to cause colitis^{25,26,27}. Moreover, cells in these Il-17⁺ clusters were heavily clonally expanded and exhibited substantial cellular and clonotypic overlap in both CD8⁺ and CD4⁺ T cells (Fig. 4d,e, Extended Data Fig. 6d,e), probably contributing to their rapid expansion. Conversely, we found a dosing-dependent decrease of innate-like CD8⁺ T cells, which have been implicated in controlling inflammation and the onset of colitis^{28,29} (Fig. 4a–c). Last, RNA velocity analyses imply that the pathogenic T_C17 and T_H17 subsets are derived from IFN- γ -expressing progenitor cells (Fig. 4c,f, Extended Data Fig. 6c,f). Accordingly, T_C17 and T_H17 cells maintained high transcript expression of *Ifng* (Fig. 4c, Extended Data Fig. 6c).

Fig. 4: Continuous dosing drives pathogenic T_C17 responses.



Mice were inoculated subcutaneously with B16F10-OVA cells and fed either a control diet or a diet containing PI-3065 inhibitor, with treatment conditions as in (Fig. 3c, d). **a**, Seurat clustering visualized by UMAP of CD8⁺ T cells in colonic tissue at day 18 after tumour inoculation of mice treated as indicated. Pie charts depict the percentage of each cluster under the different treatment conditions. **b**, Heat map comparing gene expression of cells in all clusters. Depicted are transcripts that change in expression by more than 0.5-fold with adjusted *P* values of ≤ 0.05 . DEGs were called by MAST analysis; adjusted *P* values were calculated with the Benjamini–Hochberg method. *Sell* is also known as *Cd62l*; *Itgae* is also known as *Cd103*. **c**, Seurat-normalized expression of indicated genes in the different clusters. **d**, Clone size of cells in indicated clusters in UMAP space. **e**, Euler diagrams show the clonal overlap between CD8⁺ T cells in the different clusters. **f**, RNA velocity analysis visualized by UMAP depicting likely developmental trajectories of CD8⁺ T cells. Arrows indicate velocity streamlines. T_C cells, cytotoxic T cells.

Source data

Given that IL-10⁺ ST2 T_{reg} cells have been implicated in controlling IL-17 responses that would otherwise cause colitis³⁰, our data provide an explanation for the ripple effects ensuing after PI3Kδ inhibition that eventually cause irAEs. Specifically, our data imply that *Il10*-expressing ST2 T_{reg} cells are highly susceptible to PI3Kδ inhibition, leading to a decrease in their abundance and thus to a disruption of gut homeostasis by causing a rapid expansion of pathogenic T_H17 and T_C17 cells that, together with a decrease in innate-like CD8⁺ T cells, cause colitis. Moreover, intermittent PI3Kδi dosing provides the means to uncouple the anti-tumour effects from irAEs, providing ample rationale to test this concept in a follow-up clinical trial.

Discussion

Here we find that in human and mouse tumour tissue, PI3Kδ inhibition leads to substantial changes in the cell composition of the TME by reducing the number of T_{reg} cells and activating intratumoural CD4⁺ and CD8⁺ T cells, which clonally expand and display heightened cytotoxic and cytolytic features. Notably, in mouse models, we find substantial changes in the transcriptional features and composition of colonic T_{reg} cell subsets, which indicate that PI3Kδ inhibition affects T_{reg} functionality, survival and tissue retention, thus altering T_{reg} cell frequencies or T_{reg} subtype compositions in both tumour and non-malignant tissues. These treatment-mediated changes, specifically the depletion of *Il10*-expressing ST2 T_{reg} cells, is associated with colitis and expansion of pathogenic T_C17 and T_H17 T cell subsets in colonic tissue. Notably, these findings might be more broadly applicable, as tissue-resident ST2 T_{reg} cells have been described in many non-malignant organs frequently affected by irAEs (for example, in skin) or might be affected by other T_{reg} cell-targeting immunotherapies (for example, anti-CTLA-4). We show in mouse models that intermittent dosing with PI3Kδi is a rational treatment strategy that combines sustained anti-tumour immunity with reduced toxicity.

Our data show that the immunomodulatory effects of PI3Kδi need to be evaluated judiciously in treatment-naive patients unaffected by multiple

lines of treatment and the immunosuppressive effect of haematological malignancies such as chronic lymphocytic leukaemia (CLL). It is clear that in the neoadjuvant setting in patients with HNSCC, at the evaluated doses and with daily scheduling, PI3K δ inhibition has an unfavourable safety profile, limiting its feasibility and clinical benefit by causing frequent and severe grade 3/4 irAEs, probably driven by modulation of T_{reg} cell behaviour in non-malignant tissues. On the basis of our findings, decreased dosages or an altered PI3K δ i treatment regimen will be required in solid tumours—especially in immune-competent patients—in order to be able to exploit the clear anti-tumour immune response induced by PI3K δ i while limiting the adverse effects associated with reduced T_{reg} function in healthy tissues. Finally, our data suggest that the unique cellular composition of effector versus regulatory cells in the TME of each patient might be an important determinant of the efficacy of PI3K δ inhibition. Thus, PI3K δ i might be especially useful in patients with high levels of intratumoural T_{reg} cells and an unfavourable ratio of T_{reg} versus CD8 $^{+}$ tumour-infiltrating lymphocytes (TILs) in pre-treatment samples. Our study sets the stage for further exploration of PI3K δ inhibitors as immunomodulatory agents in solid tumours.

Methods

Double-blind, randomized clinical trial and sample collection

To explore the immunomodulatory effects of PI3K δ inhibition in humans, we conducted a multicenter, placebo-controlled phase II neoadjuvant trial with the PI3K δ i AMG319 in resectable HNSCC (Extended Data Fig. 1a, b; <https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-004388-20/results>).

All patients provided written informed consent for participation in the clinical trial. We focused on human papilloma virus (HPV)-negative HNSCC, as this cancer type is more prevalent, and because patients with this cancer type have poorer outcomes when compared to HPV-positive HNSCC, probably due to overall lower TIL infiltration^{31,32,33}. The clinical trial was sponsored by Cancer Research UK Center for Drug Development (CRUKD/15/004) and approved by the Southampton and South West Hampshire Research Ethics Board; the trial EudraCT number is 2014-

004388-20. Detailed information about the trial design, randomization procedure, protocol amendments, recruitment data, patient characteristics and adverse events have been deposited at <https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-004388-20/results#moreInformationSection> and are in the CONSORT checklist. Patients were recruited after initial diagnosis and before definitive surgical treatment; drug treatment or placebo was given for up to 24 or 28 days respectively, prior to resection of tumour. In a previous phase I dose escalation study of heavily pretreated patients with either CLL or non-Hodgkin lymphoma, AMG319 doses of up to 400 mg were explored without reaching a maximally tolerated dose, and exhibited pharmacokinetic dynamics with a mean half-life of 3.8–6.6 h in plasma³⁴. In that phase 1 study, daily dosing with 400 mg AMG319 led to near complete target inhibition (BCR-induced pAKT in ex vivo IgD-stimulated CLL samples) and >50% nodal regression³⁴, while irAE at grade 3 or above according to the common toxicity criteria (CTC) occurred after days 40 and 60. We thus reasoned that high grade irAEs were unlikely to occur during the shorter treatment duration in the neoadjuvant setting, and therefore selected 400 mg d⁻¹ as the starting dose. The intended time from initiating treatment with AMG319 or placebo to surgical resection of tumour was up to four weeks, with weekly blood draws. The full evaluation of radiological measurements has previously been reported at <https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-004388-20/results> to the EU Clinical Trials Register in compliance with regulatory requirements. Primary endpoints were safety and assessment of CD8⁺ immune infiltrates, secondary endpoints tumour responses and AMG319 pharmacokinetic evaluation (<https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-004388-20/results#endPointsSection>). The sample size was calculated as follows: in a pilot cohort, the CD8 count in the biopsy taken at diagnosis, and in the resected tissue sample was quantified. The mean value at diagnosis was 25 cells per high power field (hpf), and this remained almost the same in the resected sample (26 cells per hpf). With an observed s.d. of 5 cells we posited we would observe a doubling to 50 cells per hpf following treatment with AMG 319, hence a difference between the two treatment groups of 25. To detect a standardized difference of 0.5 with 80% power and one-sided test of statistical significance of 20%, we required 36 patients to be randomized to AMG319 and 18 to placebo (54 in total). Randomization was

at the level of the individual patient, using block randomization with randomly varying block sizes. During the course of the clinical trial the randomization list was held by the unblinded trial statistician and within the IWRS. Patients and care providers were blinded to the treatment allocation, and all immunological evaluations were completed by a pathologist and researchers who were blinded to the patient allocation to treatment arms.

Patients were recruited from October 2015 to May 2018 in the UK (University Hospital Southampton NHS Foundation Trust, Poole Hospitals NHS Foundation Trust, Liverpool University Hospitals NHS Foundation Trust and Queen Elizabeth University Hospital Glasgow; two additional centres did not recruit patients); written informed consent was obtained from all subjects. Patients were eligible if they were \geq 18 years of age, with histologically proven HNSCC for whom surgery was the primary treatment option, with laboratory results within specified ranges. Patients had to be clinically eligible for tumour resection; patients who had undergone prior radiotherapy, immunotherapy, chemotherapy or other anti-cancer therapy for their current HNSCC were excluded. Clinical data were obtained for age, gender, tumour size (T stage), and nodal status (N stage) (summarized in Source Data, Patient characteristics). Adverse event reporting was according to the National Cancer Institute CTCAE Version 4.02. Performance status and overall survival was collected to death or censored at last clinical review; clinical data were anonymized once the data had been collated and verified by the sponsor. Drug dosing was at 400 mg of the oral PI3K δ inhibitor AMG319 (15 patients) and, after an independent safety review, dosing at 300 mg in 6 patients; all patients who had at least 4 doses of the drug were included in the final analyses. Radiological evaluation of change in tumour volume (Extended Data Fig. 1e) was undertaken by comparing baseline bi-dimensional measurements of tumour at baseline and before surgery. For response assessment, RECIST 1.1 was used. The full data on radiological measurements is available at

<https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-004388-20/results> in the EU Clinical Trials Register in compliance with regulatory requirements. The study was discontinued after 30 (of the target sample size of 54) patients had been dosed with AMG319 or placebo, thus limiting the clinical information on outcomes that can be gained from this trial. All patients had tissue collected as a dedicated research biopsy after consent and prior to randomization, with an additional sample collected during surgical

resection. Tumour tissue was obtained fresh on the day of biopsy or surgery and a sample was immediately snap frozen. A proportion of the tumour tissue was cryopreserved in freezing medium (90% FBS and 10% DMSO) for subsequent analyses or, alternatively, directly disaggregated using a combination of enzymatic and mechanical dissociation for immediate analysis by FACS or cryopreservation as a single-cell suspension, as previously described³⁵. Blood samples were collected during the course of the study from which plasma and peripheral blood mononuclear cells (PBMCs) were collected. PBMCs were isolated by centrifugation over lymphoprep (Axis-Shield PoC AS).

Histology and immunohistochemistry

Double immunostaining for CD8 and FOXP3 was performed on a Leica Bond RX platform, with antigen retrieval performed for 20 min at 97 °C Bond ER2 antigen retrieval solution. Primary antibodies were incubated for 30 min at room temperature (FoxP3 - Abcam: Clone 236A/E7 1:100 dilution; CD8 - DAKO: Clone C8/144B 1:50 dilution) and detected using the Leica Refine Polymer brown and red detection systems. Analysis was performed by two independent and blinded head and neck pathologists counting intratumoural CD8⁺ and FOXP3⁺ TIL in multiple random high-power fields at 200 \times magnification. Where possible, ten high-power fields were counted.

Pharmacokinetics of AMG319

Fifty microlitres of thawed plasma samples were mixed with 300 μ l of extraction solution (100 ng ml⁻¹ [2H³, N¹⁵]-AMG319 in methanol), centrifuged at 10,000g for 5 min to precipitate the plasma proteins. The supernatant was transferred to a UPLC vial and placed on the autosampler (maintained at 8 °C) for analysis. A freshly prepared calibration curve in the range 1–1,000 ng ml⁻¹ and frozen QC samples at 10, 100, 500 and 1,000 ng ml⁻¹ (K2 EDTA human plasma spiked with AMG319) were analysed alongside each batch of patient samples. Five microlitres of supernatant was injected into the UPLC-MS/MS system, configured with a Waters Acquity UPLC and Waters Quattro Premier XE mass spectrometer. Analytes were separated on an Acquity UPLC BEH C18 1.7 μ m (2.1 mm \times 100 mm)

column with a mobile phase flow rate of 0.3 ml min^{-1} . Mobile phase was composed of water, acetonitrile and formic acid. Analytes were detected using the multiple reaction monitoring (MRM) mode of the MS/MS system, operating in positive ion electrospray mode. MRMs were set up at m/z $386.4 \rightarrow 251.3$, $386.4 \rightarrow 236.6$, $251.3 \rightarrow 251.3$ and $251.3 \rightarrow 236.3$ for AMG319 and at m/z $390.5 \rightarrow 254.4$ for [2H^3 , N^{15}]-AMG319. MassLynx software (version 4.1, Waters Ltd.) was used to control the instrumentation and for analysis of the peaks of interest and processing of spectral data.

pAKT measurement

Whole blood samples (10 ml) were collected in sodium heparin tubes pre-dose and 4 h post dose on days 1 and 15 for the first 11 patients (day 8 and 15 for the remaining 19 patients). Blood was stimulated with double-diluted anti-IgD ($25\text{--}0.008 \mu\text{g ml}^{-1}$) in deep well plates for 5 min. Blood was then lysed and fixed with BD PhosFlow Lyse/Fix buffer. Cell pellets were washed and then stored at -80°C until all samples from the same patient were ready for further analysis. Upon thawing, cells pellets were incubated with anti-human CD3-FITC and CD14-FITC, washed in PBS + 1% FBS, permeabilized with 80% MeOH and washed again before intracellular staining with CD20-PE Cy7 and pAKT (S473). Stained cell pellets were washed again before staining with a secondary antibody (anti-rabbit Alexa 647). Events were subsequently acquired on a Canto II flow cytometer (BD Bioscience), and analysed using FACS Diva. MFI of pAKT in B cells was plotted against the anti-IgD concentration, which was used to activate the B cells. The area under the curve was calculated and a drop of 50% in area under the curve between pre- and post-dose was validated to be the result of drug inhibition.

Mice

C57BL/6J (JAX stock no. 000664), OT-I (JAX stock no. 003831), *Rag1*^{-/-} (JAX stock no. 002216) and *CD8*^{-/-} (JAX stock no. 002665) mice were obtained from Jackson labs. *Foxp3*^{RFP} mice (JAX stock no. 008374) were a gift from K. Ley. Age (6–12 weeks) and sex-matched mice were used for all experiments. The housing temperature was controlled, ranging from 20.5–24 °C, humidity was monitored but not controlled and ranges from 30–70%.

The 12 h daily light cycle was from 06:00 to 18:00. All animal work was approved by the relevant La Jolla institute for Immunology Institutional Animal Care and Use Committee.

Tumour experiments

Mice were inoculated with 1×10^5 to 1.5×10^5 B16F10-OVA cells subcutaneously into the right flank. Mice were put on either a control diet or a diet containing the PI3Kδi PI-3065 on day 1 or day 5 after tumour inoculation. Diets were prepared using powdered 2018 global rodent diet (Envigo) mixed with or without PI-3065 at 0.5 g kg^{-1} , which corresponds to a daily dose of 75 mg kg^{-1} as used in our previous study⁸. To pellet the food, 50% v/w water was added to the diet and dough thoroughly mixed, compressed, moulded and dried before use. Tumour size was monitored every other day, and tumour harvested at indicated time points for analysis of tumour-infiltrating lymphocytes. Tumour size limit of 15 mm in diameter was not exceeded and volume was calculated as $\frac{1}{2} \times D \times d^2$, where D is the major axis and d is the minor axis, as described³⁶.

Bulk transcriptome analyses

Cryosections (10 µm) were cut from snap frozen tumour and RNA was extracted using the Maxwell RSC instrument and Maxwell RSC SimplyRNA Tissue kit (Promega), according to the manufacturer's instructions. RNA was quantified using the Qubit fluorometer (ThermoFisher Scientific) and quality was assessed using the Agilent 2100 Bioanalyzer generating an RNA integrity number (RIN; Agilent Technologies). RNA sequencing was performed by Edinburgh Genomics; mRNA libraries were prepared using the TruSeq Stranded Total RNA Library Prep Kit (Illumina) and paired-end sequenced (100 bp) on the NovaSeq 6000 platform (Illumina) to yield an average read depth of 40×10^6 reads. Reads were mapped to hg19 reference genome using STAR with our in-house pipeline (https://github.com/ndu-UCSD/LJI_RNA_SEQ_PIPELINE_V2). A total of 22 paired (14 from treatment and 8 from placebo group) samples with at least 70% of mapping reads were selected. Differential expression analysis between the pre and post treatment, as well as between pre and post placebo, was performed

using DESeq2 (v1.24.0). The threshold for DEGs was determined with fold change of $>\log_2 0.75$ and an adjusted P value <0.1 . Between treatment pre and post, 93 genes were identified as significant, whereas 3 genes were significant between placebo pre and post. Cells were dispersed from fresh tumour tissue and used immediately for flow cytometric analysis and cell sorting. CD8⁺ T cells were bulk sorted into ice-cold TRIzol LS reagent³⁵ (Thermo Fisher Scientific) on a BD FACS Fusion (BD Bioscience). Reads from sorted CD8 RNA were mapped to hg19 reference genome using STAR with the same in-house pipeline as above. In total, we had 17 samples available, placebo (2 pre-treatment and 3 post-treatment) and treatment (6 pre-treatment and 6 post-treatment), out of which 3 were paired (1 placebo and 2 treatment). The DEGs between post treatment and remaining samples resulted in 455 significant genes (fold change of $>\log_2 0.75$ and an adjusted P value <0.05).

Flow cytometry

Cells dispersed from cryopreserved tumour tissue or PBMCs were prepared in staining buffer (PBS with 2% FBS and 2 mM EDTA), FcR blocked (clone 2.4G2, BD Biosciences) and stained with antibodies as indicated below for 30 min at 4 °C. Cell viability was determined using fixable viability dye (ThermoFisher).

Mouse lymphocytes were isolated from the spleen by mechanical dispersion through a 70-µm cell strainer (Miltenyi) to generate single-cell suspensions. RBC lysis (Biolegend) was performed to remove red blood cells. Tumour samples were harvested and lymphocytes were isolated by dispersing the tumour tissue in 2 ml of PBS, followed by incubation of samples at 37 °C for 15 min with DNase I (Sigma) and Liberase DL (Roche). The suspension was then diluted with MACS buffer and passed through a 70-µm cell strainer to generate a single-cell suspension. Colons were collected and rinsed in 1 mM dithiothreitol to remove faeces. Each colon was cut into 2–3 mm pieces and incubated 3 times in pre-digestion solution (HBSS containing 5% FBS and 2 mM EDTA) at 37 °C for 20 min under high rotation to remove epithelial cells. Then tissues were minced with scissors and incubated with digestion solution (HBSS containing 5% FBS, 100 µg ml⁻¹ DNase I (Sigma) and 1 mg ml⁻¹ collagenase (Sigma)) at 37 °C for 20 min under high rotation to get

single-cell solutions of lamina propria cells. Cells were prepared in staining buffer (PBS with 2% FBS and 2 mM EDTA), FcR blocked (clone 2.4G2, BD Biosciences) and stained with antibodies as indicated below for 30 min at 4 °C; secondary stains were done for selected markers. Samples were then sorted or fixed and intracellularly stained using a FoxP3 transcription factor kit according to manufacturer's instructions (eBioscience). Cell viability was determined using fixable viability dye (ThermoFisher). The following antibodies from BD Biosciences, Biolegend, Miltenyi or eBioscience were used: anti-human PD-1 (EH12.1, 1:30), CD4 (OKT4, 1:30), CD137 (4B4-1, 1:30), GITR (108-17, 1:30), ICOS (C398.4A, 1:50), CD8A (SK1, 1:30), CD25 (M-A251, 1:20), CD3 (SK7, 1:30), CD127 (eBioRDR5, 1:50), CD45 (HI30, 1:30), CD14 (HCD14, 1:50), CD20 (2H7, 1:50); anti-mouse CD3 (145-2C11, 1:100), CD4 (RM4-5, 1:100), CD8 (53-6.7, 1:100), PD-1 (29F1.A12, 1:100), ST2 (U29-93, 1:100) Ki67 (B56, 1:40), TOX (REA473, 1:40), CD19 (6D5, 1:100), CD45 (30-F11, 1:100), FOXP3 (FJK-16s, 1:100) and GZMB (QA16A02, 1:40). All samples were acquired on a BD FACS Fortessa or sorted on a BD FACS Fusion (both BD Biosciences) and analysed using FlowJo 10.4.1 for subsequent scRNA-seq.

Colitis experiments

DSS (molecular mass ≈ 40,000) (Alfa Aesar) 2.5% (w/v) was added to the drinking water of mice with ad libitum access. Body weight of the mice was monitored. Colon tissues were collected for histological analysis at the end point. Whole colons were harvested from mice between cecum and rectum. Stools were flushed out of lumen with PBS. Then colons were fixed with zinc formalin for 5 min. Fixed colons were opened longitudinally, flattened, cut into 3 fragments and further fixed in zinc formalin for 48 h in cassettes. After fixation, samples were transferred to 70% isopropanol for long term storage or H&E staining. Slides were scored blindly according to the following criteria: inflammation, area of infiltration (extent), crypt damage and oedema. The colon was divided into three equal parts and the middle section was utilized for scoring according to system shown in Extended Data Table 1. Four randomly selected areas were analysed and a histological score was determined.

Single-cell transcriptome analysis

Human

scRNA-seq was performed by Smart-seq2 as described³⁷. Reads were mapped with our in-house pipeline as above. Good quality cells were defined as those with at least 200 genes, at least 60 percent of mapping reads, mitochondrial counts of at most 20%, at least 50,000 total counts (reported by STAR excluding tRNA and rRNA), and a 5' to 3' bias of at most 2. Filtered cells were analysed using the package Seurat (v3.1.5). In order to separate CD4 and CD8 more effectively, we performed differential gene expression analysis between single-positive cells using *CD4* and *CD8B* genes. Cells were clustered using 178 significant genes (adjusted *P* value < 0.05).

Mouse T_{reg} cells

scRNA-seq was performed using the 10x platform (10x Genomics) according to the manufacturer's instructions. Reads were mapped with Cell Ranger followed by our in-house QC pipeline (https://github.com/vijaybioinfo/quality_control) and demultiplexed with bcl2fastq using default parameters. The Cell Ranger aggr routine was used and CITE-seq data was processed using our custom pipeline (https://github.com/vijaybioinfo/ab_capture). In brief, raw output from Cell Ranger was taken and cell barcodes with less than 100 unique molecular identifier (UMI) counts as their top feature were discarded and the remaining barcodes were classified by MULTIseqDemux from Seurat. Finally, cell barcodes where the assigned feature did not have the highest UMI count were fixed, and cells with a fold change of less than 3 between the top two features were reclassified as doublets. Before clustering, cells were filtered for at least 300 and at most 5,000 genes, at least 500 and at most 10,000 UMI counts, and at most 5% of mitochondrial counts. Cell types were identified using Seurat's FindAllMarkers function. Differential expression was calculated with MAST³⁸ (v1.10.0) DESeq2 (v1.24.0) as previously described³⁷ and genes with an adjusted *P* value < 0.05 and a fold change of $>\log_2 0.5$ were defined as significant. P-values were corrected for multiple comparisons using the Benjamini–Hochberg method. Gene Set Enrichment Analysis (GSEA) scores were estimated with fgsea (v1.10.1) in

R using signal-to-noise ratio as the metric (`minSize = 3` and `maxSize = 500`). Enrichment scores were shown as GSEA plots. Signature scores were computed using Seurat's `AddModuleScore` function with default parameters. In short, the score is defined for each cell by subtracting the mean expression of an aggregate of control gene lists from the mean of the signature gene list. Control gene lists were randomly selected (same size as the signature list) from bins delimited based on the level of expression of the signature list.

Mouse colonic CD4⁺ and CD8⁺ T cells

scRNA-seq was performed by using the 10x platform. Mapping, aggregation, and QC was carried out as described above with the following thresholds: genes per cells range of [300; 4,500], UMI content per cell was [500; 20,000], percent of mitochondrial counts of $\leq 10\%$, and a doublet score of ≤ 0.3 . Clusters of contaminant cells expressing epithelial, monocyte, and fibroblast markers were eliminated after the first round of clustering. The final number of cells comprised $n = 6,415$ CD4⁺ T cells and $n = 2,715$ CD8⁺ T cells.

T cell receptor analysis

TCRs were reconstructed from scRNA-seq reads using MiXCR with default parameters. Then, shared TCRs were defined by having the same CDR3 sequence in both the alpha and beta chains and coming from the same donor. Enriched TCR were defined as those with a frequency higher or equal to two. Finally, TCR network plots were generated using the Python package graphviz.

Quantification and statistical analysis

The number of subjects, samples or mice per group, replication in independent experiments and statistical tests can be found in the figure legends. Details on quality control, sample elimination and displayed data are stated in the methods and figure legends. Sample sizes were chosen based on published studies to ensure sufficient numbers of mice in each group enabling reliable statistical testing and accounting for variability.

RNA-seq samples that did not pass the QC check were excluded from downstream analyses. Experimental results were reliably reproduced in at least two independent experiments. Animals of same sex and age were randomly assigned to experimental groups, and blinding was not performed. Extended Data Figs. [1a](#), [3a](#) were created with BioRender.com, the statistical analyses were performed with Graph Pad Prism 9 and statistical tests used are indicated in the figure legends.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Sequencing data has been uploaded onto the Gene Expression Omnibus (accession code [GSE166150](#)). [Source data](#) are provided with this paper.

Code availability

Scripts used for this paper and explanations thereof are available at our GitHub repository (https://github.com/vijaybioinfo/PI3Kd_2022).

References

1. Furman, R. R. et al. Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **370**, 997–1007 (2014).
2. Gopal, A. K. et al. PI3K δ inhibition by idelalisib in patients with relapsed indolent lymphoma. *N. Engl. J. Med.* **370**, 997–1007 (2014).
3. Vanhaesebroeck, B., Perry, M. W. D., Brown, J. R., André, F. & Okkenhaug, K. PI3K inhibitors are finally coming of age. *Nat. Rev. Drug Discov.* **20**, 741–769 (2021).

4. Ali, K. et al. Inactivation of PI(3)K p110 δ breaks regulatory T-cell-mediated immune tolerance to cancer. *Nature* **510**, 407–411 (2014).
5. Lim, E. L. et al. Phosphoinositide 3-kinase δ inhibition promotes antitumor responses but antagonizes checkpoint inhibitors. *JCI Insight* **3**, e120626 (2018).
6. Lampson, B. L. et al. Idelalisib given front-line for treatment of chronic lymphocytic leukemia causes frequent immune-mediated hepatotoxicity. *Blood* **128**, 195–203 (2016).
7. Lim, E. L. & Okkenhaug, K. Phosphoinositide 3-kinase δ is a regulatory T-cell target in cancer immunotherapy. *Immunology* **157**, 210–218 (2019).
8. Chellappa, S. et al. The PI3K p110 δ isoform inhibitor idelalisib preferentially inhibits human regulatory T cell function. *J. Immunol.* **202**, 1397–1405 (2019).
9. Scott, A. C. et al. TOX is a critical regulator of tumour-specific T cell differentiation. *Nature* **571**, 270–274 (2019).
10. Khan, O. et al. TOX transcriptionally and epigenetically programs CD8 $^{+}$ T cell exhaustion. *Nature* **571**, 211–218 (2019).
11. Alfei, F. et al. TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection. *Nature* **571**, 265–269 (2019).
12. Lampson, B. L. & Brown, J. R. PI3K δ -selective and PI3K α/δ -combinatorial inhibitors in clinical development for B-cell non-Hodgkin lymphoma. *Expert Opin. Investig. Drugs* **26**, 1267–1279 (2017).
13. DiSpirito, J. R. et al. Molecular diversification of regulatory T cells in nonlymphoid tissues. *Sci. Immunol.* **3**, eaat5861 (2018).
14. Miragaia, R. J. et al. Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. *Immunity* **50**, 493–504.e7

(2019).

15. Wing, J. B., Ise, W., Kurosaki, T. & Sakaguchi, S. Regulatory T cells control antigen-specific expansion of Tfh cell number and humoral immune responses via the coreceptor CTLA-4. *Immunity* **41**, 1013–1025 (2014).
16. Ovcinnikovs, V. et al. CTLA-4-mediated transendocytosis of costimulatory molecules primarily targets migratory dendritic cells. *Sci. Immunol.* **4**, eeaw0902 (2019).
17. Zhang, N. et al. Regulatory T cells sequentially migrate from inflamed tissues to draining lymph nodes to suppress the alloimmune response. *Immunity* **30**, 458–469 (2009).
18. Delacher, M. et al. Precursors for nonlymphoid-tissue T_{reg} cells reside in secondary lymphoid organs and are programmed by the transcription factor BATF. *Immunity* **52**, 295–312 (2020).
19. Whibley, N., Tucci, A. & Powrie, F. Regulatory T cell adaptation in the intestine and skin. *Nat. Immunol.* **20**, 386–396 (2019).
20. Arpaia, N. et al. A distinct function of regulatory T cells in tissue protection. *Cell* **162**, 1078–1089 (2015).
21. Aaronson, D. S. & Horvath, C. M. A road map for those who don't know JAK–STAT. *Science* **296**, 1653–1655 (2002).
22. Hu, X. & Ivashkiv, L. B. Cross-regulation of signaling pathways by interferon- γ : implications for immune responses and autoimmune diseases. *Immunity* **31**, 539–550 (2009).
23. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
24. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).

25. L, R. et al. Differential regulation of interleukin 17 and interferon gamma production in inflammatory bowel disease. *Gut* **58**, 1629–1636 (2009).
26. Harbour, S. N., Maynard, C. L., Zindl, C. L., Schoeb, T. R. & Weaver, C. T. Th17 cells give rise to Th1 cells that are required for the pathogenesis of colitis. *Proc. Natl Acad. Sci. USA* **112**, 7061–7066 (2015).
27. Feng, T. et al. Th17 cells induce colitis and promote Th1 cell responses through IL-17 induction of innate IL-12 and IL-23 production. *J. Immunol.* **186**, 6313–6318 (2011).
28. Denning, T. L. et al. Mouse TCR $\alpha\beta^+$ CD8 $\alpha\alpha$ intraepithelial lymphocytes express genes that down-regulate their antigen reactivity and suppress immune responses. *J. Immunol.* **178**, 4230–4239 (2007).
29. Poussier, P., Ning, T., Banerjee, D. & Julius, M. A unique subset of self-specific intraintestinal T cells maintains gut integrity. *J. Exp. Med.* **195**, 1491–1497 (2002).
30. Chaudhry, A. et al. CD4 $^+$ regulatory T cells control T_H17 responses in a STAT3-dependent manner. *Science* **326**, 986–991 (2009).
31. Ward, M. J. et al. Staging and treatment of oropharyngeal cancer in the human papillomavirus era. *Head Neck* **37**, 1002–1013 (2015).
32. Wood, O. et al. Gene expression analysis of TIL rich HPV-driven head and neck tumors reveals a distinct B-cell signature when compared to HPV independent tumors. *Oncotarget* **7**, 56781–56797 (2016).
33. Ottensmeier, C. H. et al. Upregulated glucose metabolism correlates inversely with CD8 $^+$ T-cell infiltration and survival in squamous cell carcinoma. *Cancer Res.* **76**, 4136–4148 (2016).
34. Lanasa, M. C. et al. First-in-human study of AMG 319, a highly selective, small molecule inhibitor of PI3K δ , in adult patients with relapsed or refractory lymphoid malignancies. *Blood* **122**, 678 (2013).

35. Vijayanand, P. et al. Invariant natural killer T cells in asthma and chronic obstructive pulmonary disease. *N. Engl. J. Med.* **356**, 1410–1422 (2007).
36. Juneja, V. R. et al. PD-L1 on tumor cells is sufficient for immune evasion in immunogenic tumors and inhibits CD8 T cell cytotoxicity. *J. Exp. Med.* **214**, 895–904 (2017).
37. Patil, V. S. et al. Precursors of human CD4⁺ cytotoxic T lymphocytes identified by single-cell transcriptome analysis. *Sci. Immunol.* **3**, 8664 (2018).
38. Finak, G. et al. MAST: A flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).

Acknowledgements

We thank C. Rommel, K. Ali, P. Johnson and D. Scott for support and help during the initial phases of setting up the CRUK/Amgen collaboration; F. Martin for project support over the years; M. Lopez and the research pathology core at University Hospital Southampton for performing the immunohistochemistry on the paraffin embedded material; I. Henderson for help with the tetanus ELISAs; the La Jolla Institute (LJI) Flow Cytometry Core for assisting with cell sorting; the LJI sequencing core for the bulk and single-cell sequencing; the LJI histology core for processing the mouse histology samples; G. Means for method and training for the phospho-AKT assays; P. Friedmann for constructive critique of the paper; and D. Singh for advice on and help with experimental work. This clinical trial work was funded by a CDD trial Grant CRUKD/15/004 (C.H.O.), a Cancer Research UK Centres Network Accelerator Award Grant (A21998) (O.W., K.M. and J.T.), the CRUK and NIHR Experimental Cancer Medicine Center (ECMC) Southampton (A15581), the CRUK and NIHR ECMC Liverpool (A25153), Cancer Research UK Programme Grant (C23338/A25722 (E.L.-G. and B.V.)); the UK NIHR UCLH Biomedical Research Centre (B.V.), S10OD025052 (Illumina Novaseq6000), S10RR027366 (FACSaria II cell sorter), NIH grant P01 DK46763 (M.K.), the William K. Bowes Jr

Foundation (P.V.), Whittaker iCure Foundation (P.V., L.C. and C.H.O.), the Deutsche Forschungsgemeinschaft DFG research fellowship no. WI 5255/1-1:1 (A.v.W.) and Erwin Schrödinger Fellowship (M.D.). The clinical delivery of this work was supported by the Wessex Clinical Research Network and National Institute of Health Research UK. We further acknowledge Cancer Research UK (Centre for Drug Development) as the clinical trial Sponsor and for funding and management of the Phase II clinical trial, as well as Amgen for supply of the PI3K δ i AMG319.

Author information

Author notes

1. These authors jointly supervised this work: Pandurangan Vijayanand, Christian H. Ottensmeier

Authors and Affiliations

1. La Jolla Institute for Immunology, La Jolla, CA, USA

Simon Eschweiler, Ciro Ramírez-Suástegui, Yingcong Li, Hayley Simon, Monalisa Mondal, Alice Wang, Martina Dicker, Ting-Fang Chou, Gregory Seumois, Ferhat Ay, Mitchell Kronenberg, Pandurangan Vijayanand & Christian H. Ottensmeier

2. Division of Biological Sciences, University of California San Diego, La Jolla, CA, USA

Yingcong Li & Mitchell Kronenberg

3. CRUK and NIHR Experimental Cancer Medicine Center, University of Southampton, Southampton, UK

Emma King, Jaya Thomas, Oliver Wood, Adrian von Witzleben, Danielle Jeffrey, Katy McCann, Gareth J. Thomas & Christian H. Ottensmeier

4. Dorset Cancer Centre, Poole Hospital NHS Foundation Trust, Poole, UK

Emma King & Joseph H. Davies

5. Liverpool Head and Neck Center and Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

Lindsey Chudley, Joseph J. Sacco, Andrew Graeme Schache, Richard Shaw, Terry M. Jones, Pandurangan Vijayanand & Christian H. Ottensmeier

6. Department of Otorhinolaryngology, Head and Neck Surgery, Ulm University Medical Center, Ulm, Germany

Adrian von Witzleben

7. UCL Cancer Institute, University College London, London, UK

Elena Lopez-Guadamillas & Bart Vanhaesebroeck

8. Centre for Drug Development, Cancer Research UK, London, UK

Nicola A. Dobbs, Louisa Essame, Gary Acton & Fiona Kelly

9. Cancer Research UK Formulation Unit, University of Strathclyde, Glasgow, UK

Gavin Halbert

10. Clatterbridge Cancer Centre NHS Foundation Trust and Liverpool Cancer Research UK Experimental Cancer Medicine Center Liverpool, Liverpool, UK

Joseph J. Sacco & Christian H. Ottensmeier

11. Liverpool University Hospitals NHS Foundation Trust, Liverpool, UK

Andrew Graeme Schache, Richard Shaw & Terry M. Jones

12. Queen Elizabeth University Hospital, Glasgow, UK

James Anthony McCaul

13. Beatson West of Scotland Cancer Centre, Glasgow, UK

Claire Paterson

14. Queen Alexandra Hospital, Portsmouth, UK

Peter A. Brennan

15. Southampton University Hospitals NHS Foundation Trust,
Southampton, UK

Rabindra P. Singh

16. University of Bradford, Institute of Cancer Therapeutics, Bradford, UK

Paul M. Loadman

17. Cancer Research UK and UCL Cancer Trials Centre, London, UK

William Wilson & Allan Hackshaw

18. Department of Pathology, University of Cambridge, Cambridge, UK

Klaus Okkenhaug

19. Amgen, Thousand Oaks, CA, USA

Greg Friberg

20. Department of Medicine, University of California San Diego, La Jolla,
CA, USA

Pandurangan Vijayanand

Contributions

S.E.: study design, experimental work, data interpretation and writing. C.R.-S.: bioinformatic evaluation, data interpretation and paper review. Y.L.: experimental work and paper review. E.K.: study design, patient recruitment and paper review. L.C.: data collation, data interpretation, paper writing and review. J.T.: bioinformatic evaluation, data interpretation and paper review. O.W.: experimental work, data interpretation and paper review. A.v.W.: experimental work, data interpretation and paper review. D.J.: experimental work, data interpretation and paper review. K.M.: experimental work, data interpretation and paper review. H.S., M.M. and A.W.: experimental work and paper review. E.L.-G.: experimental work, provision of study materials, data interpretation and paper review. T.-F.C.: experimental work. M.D.: experimental work. N.A.D., L.E. and F.K.: study conduct, safety data review and monitoring, data review and verification for sponsor. G.A.: study design, safety review and data review. G.H.: generation and provision of placebo and IMP. J.J.S., A.G.S., R.S., J.A.M., C.P., J.H.D., P.A.B. and R.P.S.: patient recruitment and paper review. P.L.: data generation and interpretation and paper review. W.W.: study design and statistical review for sponsor. A.H.: study design and statistical review. G.J.T.: histopathological evaluation, data generation and interpretation and paper review. T.M.J.: study development, patient recruitment and paper review. F.A.: bioinformatic analyses and supervision, data review and paper review. G.S.: RNA sequencing and quality control supervision. K.O.: study design and paper writing and review. G.F.: study development. M.K.: study design, paper writing and review. B.V.: study design, paper writing and review. P.V.: study design, data generation and review, paper writing and review. C.H.O.: study design, patient recruitment, data generation and review, paper writing and review. P.V. and C.H.O. conceived, supervised and led the work.

Corresponding authors

Correspondence to [Pandurangan Vijayanand](#) or [Christian H. Ottensmeier](#).

Ethics declarations

Competing interests

G.F. is an employee of Amgen Inc. B.V. is a consultant for iOnctura (Geneva, Switzerland), Venthera (Palo Alto, US) and Olema Pharmaceuticals (San Francisco, US) and has received speaker fees from Gilead (Foster City, US). K.O. has received consultancy fees from iOnctura, Macomics, Gilead Sciences and Karus Therapeutics and has received research funding from GSK. M.K. is on the scientific advisory board of Prometheus. C.H.O. led the clinical trial of AMG319 with funding by Cancer Research UK, Amgen provided clinical grade compound free of charge for this trial. All other authors declare no conflicts of interest.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work. [Peer review reports](#) are available.

Additional information

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

Extended data figures and tables

[Extended Data Fig. 1 Trial schematic, pharmacokinetic & pharmacodynamic assessments and tumor response evaluation.](#)

a, Trial schematic of the placebo-controlled randomized phase II study. **b**, Consort workflow, 36 patients were screened, of which 33 were recruited and randomly allocated to the placebo control arm or AMG319 drug-treatment arm; 30 patients ultimately received at least one dose of either AMG319 or placebo. Of the 21 patients that were treated with AMG319, 6 patients received daily doses of 300 mg and 15 patients (2 patients withdrew consent prior to receiving the first dose) received daily doses of 400 mg. An initial biopsy was taken before trial initiation and surgical resection of

tumors was performed 4–6 weeks after the first dose of treatment. Pre- on- and post-treatment blood samples were collected for further analysis. **c**, **d**, Assessment of the level of AKT phosphorylation in B cells at indicated time points pre-dose and 4h after treatment with AMG319, data from one representative patient are shown. **d**, Plasma concentrations of AMG319 in placebo-controlled and drug-treated patients at indicated time points, n = 9 patients for the Placebo group and n = 18 patients for the AMG319 group. Highlighted in red are patients who were either on-treatment or had only recently (2 days prior to analysis) or briefly discontinued treatment. **e**, Waterfall plot depicting the change in tumor volume from screening to pre-surgery measured by MRI scan in patients treated with AMG319 (blue bars) or placebo (green bars), shown are all patients for which MRI scans have been performed. Data in (**d**) are mean +/- S.E.M.

Extended Data Fig. 2 Single-cell RNA-seq analysis reveals substantial oligoclonal expansion of tumor-infiltrating CD8⁺ T cells post-treatment.

a, Analysis of smart-seq2 single-cell RNA-seq data of sorted tumor-infiltrating CD3⁺ T cells from patients 20, 30, 32, 33, 34 and 35 displayed by UMAP analysis. **b**, Violin plots depicting the Seurat normalized expression of differentially expressed highlighted genes in CD4⁺ T cells (left) or CD8⁺ T cells (right) of the 6 patients pertaining to **a**, the center line depicts the median, edges delineate the 25th and 75th percentiles and whiskers depict minimum and maximum values. **c**, Percentages of non-expanded and expanded CD8⁺ and CD4⁺ T cell clones in pre- *versus* post-treatment samples. **d**, **e**, flow-cytometric analyses of the frequency of and expression of activation markers in circulating T_{REG} cells in placebo-treated (d) and AMG319-treated patients (e), P = 0.0098 for the frequency of circulating T_{REG} cells (baseline vs d8 and baseline vs d22), P = 0.0234 for ICOS⁺ T_{REG} cells, at indicated time points; the blue lines depict patients with grade 3/4 irAEs and the red dots indicate the patients with CR/PR. Data are mean +/- S.E.M. Significance for comparisons were computed using two-tailed Wilcoxon matched-pairs signed rank test between baseline and d8 or d22 respectively.

[Source data](#)

[Extended Data Fig. 3 PI3K \$\delta\$ -inhibition induces a pro-inflammatory tumor microenvironment.](#)

Mice were inoculated *s.c.* with B16F10-OVA cells and fed either a control diet or a diet containing the PI-3065 PI3K δ inhibitor for the indicated treatment period. Tumor volume (**a**) and flow-cytometric analyses of cell frequencies (**b–g**) of mice treated as indicated; P = 0.003 (a), P = 0.0076 (b), P = 0.0279 (c), P = 0.0172 (d), P = 0.0013 (e), P = 0.0435 (f), n = 9 mice for the Placebo group and n = 10 mice for the PI3K δ group for a-f g, shown are representative contour plots of intratumoral CD8 $^{+}$ T cells depicting the indicated markers. **h**, Tumor volume of Rag1 $^{-/-}$ or CD8 $^{-/-}$ mice treated as indicated, n = 6 mice/group for Rag1 $^{-/-}$ and n = 5 mice/group for CD8 $^{-/-}$ mice. **i–k**, flow-cytometric analyses of T_{REG} cell frequencies (**b**) in indicated organs of mice treated as indicated, n = 6 mice for the Placebo group and n = 5 mice for the PI3K δ group. Shown are representative contour plots of FoxP3-expressing (RFP $^{+}$) CD4 $^{+}$ T cells in indicated organs; P = 0.0043 (spleen), P = 0.0043 (tumor), P = 0.0043 (colon). Not significant, P = 0.1234; *P = 0.0332; **P = 0.0002; and ***P < 0.0001. Data are mean +/- S.E.M and statistical significance for comparisons was computed using two-tailed Mann-Whitney test; data are representative of at least two independent experiments. DEGs in (**b–e**) were called using MAST and adjusted p-values were calculated with the Benjamini-Hochberg method.

[Source data](#)

[Extended Data Fig. 4 T_{REG} cells in different tissues exhibit unique transcriptomic signatures.](#)

a, Analysis of 10x single-cell RNA-seq data displayed by UMAP analysis. Seurat clustering of FoxP3 $^{+}$ CD4 $^{+}$ T cells in spleen (left), tumor (middle) and colon (right). **b**, Heatmap comparing gene expression of cells in all clusters. Depicted are transcripts that change in expression more than 0.5-fold and adjusted P value of ≤ 0.05 . **c–e**, Volcano plots of single-cell RNA-seq analysis of placebo-treated control mice and PI-3065-treated mice in spleen

(c) tumor **(d)** and colon **(e)**. Highlighted are transcripts with a >0.5 log₂ fold change. **f**, Bar charts depicting the proportion of cells in each cluster. Bars are colorized based on cells in indicated treatments making up the cluster. **g**, Violin plots showing normalized expression levels ($\log_2(\text{CPM}+1)$) of highlighted genes in cluster the colonic clusters pertaining to Fig. [2a, b](#), the center line depicts the median, edges delineate the 25th and 75th percentiles and whiskers depict minimum and maximum values.

Extended Data Fig. 5 Colonic ST2 T_{REG} cells exhibit features of superior suppressive capacity.

a, Curtain plot highlighting selected genes in each cluster with average transcript expression (color scale) and percent of expressing cells (size scale). **b**, Flow-cytometric analyses depicting the expression of ST2 in CD8⁺ T cells or CD4⁺ T_{REG} cells in representative zebra plots (left), the frequency of ST2⁺ T_{REG} cells (n.s., P = 0.0549), FOXP3⁺ cells (P = 0.0007), and the ratio of CD8⁺ T cells to ST2 T_{REG} cells (P = 0.0047) in placebo-treated (n = 6 mice) and PI3Kδi-treated mice (n = 8 mice) **c**, Violin plots showing normalized expression levels ($\log_2(\text{CPM}+1)$) of highlighted genes in indicated clusters pertaining to Fig. [2a, b](#), the center line depicts the median, edges delineate the 25th and 75th percentiles and whiskers depict minimum and maximum values. **d–g**, Flow-cytometric analyses of CD8⁺ T cell frequencies in colon **(d)** P = 0.0025, spleen **(e)** P = 0.0013 and of the expression of PD-1 **(f)** P = 0.0025 and ICOS **(g)** P = 0.0025 on colonic CD8⁺ T cells, n = 7 mice for the Placebo group and n = 5 mice for the PI3Kδ group. Not significant, P = 0.1234; *P = 0.0332; **P = 0.0002; and ****P < 0.0001. Data **(b, d–g)** are mean +/- S.E.M and statistical significance for comparisons was computed using two-tailed Mann-Whitney test; data are representative of at least 2 independent experiments.

Source data

Extended Data Fig. 6 Continuous dosing drives pathogenic T_H17 responses.

Mice were inoculated *s.c.* with B16F10-OVA cells and fed either a control diet or a diet containing the PI-3065 PI3K δ inhibitor, treatment conditions as in (Fig. 3c,d). **a**, depicted is Seurat clustering, visualized by UAMP, of CD4 $^{+}$ T cells in colonic tissue at day 18 after tumor inoculation of mice treated as indicated, pie charts depict the percentage of each cluster in the different treatment conditions. **b**, Heatmap comparing gene expression of cells in all clusters. Depicted are transcripts that change in expression more than 0.5-fold and adjusted *P* value of ≤ 0.05 , DEGs were called by MAST analysis, adjusted p-values were calculated with the Benjamini-Hochberg method. **c**, Seurat-normalized expression of *Ifng* (top left), *Il17a* (top right), *Foxp3* (bottom left) and *Il4* (bottom right) in the different clusters. **d**, Clone size of cells in indicated clusters in UMAP space. **e**, Euler diagrams show the clonal overlap between the CD4 $^{+}$ T cells in the different clusters. **f**, RNA velocity analysis visualized by UMAP depicting likely developmental trajectories of CD4 $^{+}$ T cells, arrows indicate velocity streamlines.

Extended Data Table 1 Colitis scoring system

Supplementary information

[Reporting Summary](#)

[Peer Review File](#)

[Supplementary Data](#)

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 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

Cite this article

Eschweiler, S., Ramírez-Suástegui, C., Li, Y. *et al.* Intermittent PI3K δ inhibition sustains anti-tumour immunity and curbs irAEs. *Nature* **605**, 741–746 (2022). <https://doi.org/10.1038/s41586-022-04685-2>

- Received: 17 March 2021
- Accepted: 24 March 2022
- Published: 04 May 2022

- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04685-2>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04685-2>

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

Amendments & Corrections

- [**Author Correction: Fluctuating interaction network and time-varying stability of a natural fish community**](#) [06 May 2022]
Author Correction •
- [**Author Correction: The loss of ions from Venus through the plasma wake**](#) [09 May 2022]
Author Correction •
- [**Publisher Correction: Reproducible brain-wide association studies require thousands of individuals**](#) [09 May 2022]
Publisher Correction •

[Download PDF](#)

- Author Correction
- [Published: 06 May 2022](#)

Author Correction: Fluctuating interaction network and time-varying stability of a natural fish community

- [Masayuki Ushio](#)^{1,2,3,4},
- [Chih-hao Hsieh](#)^{5,6,7},
- [Reiji Masuda](#)⁸,
- [Ethan R. Deyle](#)⁹,
- [Hao Ye](#)^{9,10},
- [Chun-Wei Chang](#)⁶,
- [George Sugihara](#)⁹ &
- [Michio Kondoh](#)¹

[Nature](#) volume **605**, page E9 (2022)

- 483 Accesses
- 2 Altmetric
- [Metrics details](#)

The [Original Article](#) was published on 07 February 2018

Correction to: *Nature* <https://doi.org/10.1038/nature25504> Published online
7 February 2018

In this Letter, we analysed time series of a natural fish community. Using empirical dynamic modelling (EDM), we quantified time-varying interspecific interactions in the community and reconstructed fluctuating interaction networks. Then, time-varying community stability, named dynamic stability, was evaluated by computing the dominant eigenvalue of the time-varying interaction matrix, but an amendment is needed in the calculation of dynamic stability. In the original Letter, dynamic stability, or the local Lyapunov stability, was calculated as ‘the absolute value of the real part of the dominant eigenvalue of the interaction matrix A ,’ as described in the Methods, but this should have been ‘the absolute value of the dominant eigenvalue of the interaction matrix A .’ In addition, ‘the absolute value of the real part of the dominant eigenvalue of the interaction matrix’ in the Fig. 3 legend should have read ‘the absolute value of the dominant eigenvalue of the interaction matrix.’ The original version of dynamic stability covers a point equilibrium with no imaginary part, whereas the corrected version covers a broader situation with a non-zero imaginary part.

We have re-calculated the dynamic stability and re-run all analyses that involved the dynamic stability. The reanalysis affects Figs 3a, 4 and Extended Data Figs 4, 5b and 6g–j of the original Article. We have found almost identical results, as shown in the corrected figures (see Supplementary Information to this Amendment). The insignificant results in the original Extended Data Fig. 6e, f are now significant, but these do not influence our discussion or conclusions. The original Letter has not been corrected online

Author information

Authors and Affiliations

1. Department of Environmental Solution Technology, Faculty of Science and Technology, Ryukoku University, Otsu, Japan

Masayuki Ushio & Michio Kondoh

2. Joint Research Center for Science and Technology, Ryukoku University, Otsu, Japan

Masayuki Ushio

3. Center for Ecological Research, Kyoto University, Otsu, Japan

Masayuki Ushio

4. PRESTO, Japan Science and Technology Agency, Kawaguchi, Japan

Masayuki Ushio

5. Institute of Oceanography, Institute of Ecology and Evolutionary Biology, and Department of Life Science, National Taiwan University, Taipei, Taiwan

Chih-hao Hsieh

6. Taiwan International Graduate Program (TIGP)–Earth System Science Program, Academia Sinica and National Central University, Taipei, Taiwan

Chih-hao Hsieh & Chun-Wei Chang

7. National Center for Theoretical Science, Taipei, Taiwan

Chih-hao Hsieh

8. Maizuru Fisheries Research Station, Field Science Education and Research Center, Kyoto University, Maizuru, Kyoto, Japan

Reiji Masuda

9. Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California, USA

Ethan R. Deyle, Hao Ye & George Sugihara

10. Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, Florida, USA

Hao Ye

Corresponding authors

Correspondence to [Masayuki Ushio](#) or [Michio Kondoh](#).

Additional information

Supplementary information is available in the online version of this Amendment.

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Figures 1–5, showing the original and corrected versions of Fig. 3a, Fig. 4, Extended Data Fig. 4, Extended Data Fig. 5b, and Extended Data Fig. 6d–j.

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Ushio, M., Hsieh, Ch., Masuda, R. *et al.* Author Correction: Fluctuating interaction network and time-varying stability of a natural fish community. *Nature* **605**, E9 (2022). <https://doi.org/10.1038/s41586-022-04810-1>

- Published: 06 May 2022
- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04810-1>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04810-1>

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

[Download PDF](#)

- Author Correction
- [Published: 09 May 2022](#)

Author Correction: The loss of ions from Venus through the plasma wake

- [S. Barabash](#)¹,
- [A. Fedorov](#)²,
- [J. J. Sauvaud](#)²,
- [R. Lundin](#)¹,
- [C. T. Russell](#)³,
- [Y. Futaana](#)¹,
- [T. L. Zhang](#)⁴,
- [H. Andersson](#)¹,
- [K. Brinkfeldt](#)¹,
- [A. Grigoriev](#)¹,
- [M. Holmström](#)¹,
- [M. Yamauchi](#)¹,
- [K. Asamura](#)⁵,
- [W. Baumjohann](#)⁴,
- [H. Lammer](#)⁴,
- [A. J. Coates](#)⁶,
- [D. O. Kataria](#)⁶,
- [D. R. Linder](#)⁶,
- [C. C. Curtis](#)⁷,
- [K. C. Hsieh](#)⁷,
- [B. R. Sandel](#)⁷,

- M. Grande⁸,
- H. Gunell⁹,
- H. E. J. Koskinen^{10,11},
- E. Kallio¹¹,
- P. Riihelä¹¹,
- T. Säles¹¹,
- W. Schmidt¹¹,
- J. Kozyra¹²,
- N. Krupp¹³,
- M. Fränz¹³,
- J. Woch¹³,
- J. Luhmann¹⁴,
- S. McKenna-Lawlor¹⁵,
- C. Mazelle²,
- J.-J. Thocaven²,
- S. Orsini¹⁶,
- R. Cerulli-Irelli¹⁶,
- A. Mura¹⁶,
- A. Milillo¹⁶,
- M. Maggi¹⁶,
- E. Roelof¹⁷,
- P. Brandt¹⁷,
- K. Szego¹⁸,
- J. D. Winningham¹⁹,
- R. A. Frahm¹⁹,
- J. Scherrer¹⁹,
- J. R. Sharber¹⁹,
- P. Wurz²⁰ &
- P. Bochsler²⁰

Nature volume **605**, page E10 (2022)

- 310 Accesses
- 1 Altmetric

- [Metrics details](#)

The [Original Article](#) was published on 29 November 2007

Correction to: *Nature* <https://doi.org/10.1038/nature06434> Published online 29 November 2007

In the version of this article initially published, there were errors in two authors' names. A. Mura and A. Milillo (Istituto di Fisica dello Spazio Interplanetari, Rome, Italy) originally were listed as M. Mura and M. Milillo.

Author information

Authors and Affiliations

1. Swedish Institute of Space Physics, Kiruna, Sweden

S. Barabash, R. Lundin, Y. Futaana, H. Andersson, K. Brinkfeldt, A. Grigoriev, M. Holmström & M. Yamauchi

2. Centre d'Étude Spatiale des Rayonnements, BP-44346, Toulouse, France

A. Fedorov, J. J. Sauvaud, C. Mazelle & J.-J. Thocaven

3. IGPP, University of California, Los Angeles, California, USA

C. T. Russell

4. Space Research Institute, Austrian Academy of Science, Graz, Austria

T. L. Zhang, W. Baumjohann & H. Lammer

5. Institute of Space and Astronautical Science, Sagamihara, Japan

K. Asamura

6. Mullard Space Science Laboratory, University College London,
Holmbury St Mary, Dorking, Surrey, UK

A. J. Coates, D. O. Kataria & D. R. Linder

7. University of Arizona, Tucson, Arizona, USA

C. C. Curtis, K. C. Hsieh & B. R. Sandel

8. University of Wales Aberystwyth, Penglais, Aberystwyth, Ceredigion,
UK

M. Grande

9. Department of Physics, West Virginia University, Morgantown, West
Virginia, USA

H. Gunell

10. Department of Physical Sciences, University of Helsinki, Box 64,
Helsinki, Finland

H. E. J. Koskinen

11. Finnish Meteorological Institute, Box 503, Helsinki, Finland

H. E. J. Koskinen, E. Kallio, P. Riihelä, T. Säles & W. Schmidt

12. Space Physics Research Laboratory, University of Michigan, Ann
Arbor, Michigan, USA

J. Kozyra

13. Max-Planck-Institut für Sonnensystemforschung, Katlenburg-Lindau,
Germany

N. Krupp, M. Fränz & J. Woch

14. Space Science Laboratory, University of California, Berkeley,
California, USA

J. Luhmann

15. Space Technology Ireland, National University of Ireland, Maynooth,
Co. Kildare, Ireland

S. McKenna-Lawlor

16. Instituto di Fisica dello Spazio Interplanetari, Rome, Italy

S. Orsini, R. Cerulli-Irelli, A. Mura, A. Milillo & M. Maggi

17. Applied Physics Laboratory, Johns Hopkins University, Laurel,
Maryland, USA

E. Roelof & P. Brandt

18. KFKI Research Institute for Particle and Nuclear Physics, PO Box 49,
H-1525, Budapest, 114, Hungary

K. Szego

19. Southwest Research Institute, San Antonio, Texas, USA

J. D. Winnigham, R. A. Frahm, J. Scherrer & J. R. Sharber

20. University of Bern, Physikalisches Institut, CH-3012, Bern,
Switzerland

P. Wurz & P. Bochsler

Corresponding author

Correspondence to [S. Barabash](#).

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Barabash, S., Fedorov, A., Sauvaud, J.J. *et al.* Author Correction: The loss of ions from Venus through the plasma wake. *Nature* **605**, E10 (2022). <https://doi.org/10.1038/s41586-022-04621-4>

- Published: 09 May 2022
- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04621-4>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04621-4>

[Download PDF](#)

- Publisher Correction
- Open Access
- [Published: 09 May 2022](#)

Publisher Correction: Reproducible brain-wide association studies require thousands of individuals

- [Scott Marek](#) ORCID: [orcid.org/0000-0003-0032-9052¹](https://orcid.org/0000-0003-0032-9052)^{na1},
- [Brenden Tervo-Clemmens](#) ORCID: [orcid.org/0000-0002-9557-1126^{2,3}](https://orcid.org/0000-0002-9557-1126)^{na1},
- [Finnegan J. Calabro](#) ORCID: [orcid.org/0000-0002-8092-3942^{4,5}](https://orcid.org/0000-0002-8092-3942),
- [David F. Montez](#)⁶,
- [Benjamin P. Kay](#)⁶,
- [Alexander S. Hatoum](#)¹,
- [Meghan Rose Donohue](#)¹,
- [William Foran](#)⁴,
- [Ryland L. Miller](#)^{1,6},
- [Timothy J. Hendrickson](#)⁷,
- [Stephen M. Malone](#)⁸,
- [Sridhar Kandala](#)¹,
- [Eric Feczkó](#)^{9,10},
- [Oscar Miranda-Dominguez](#)^{9,10},
- [Alice M. Graham](#)¹¹,
- [Eric A. Earl](#) ORCID: [orcid.org/0000-0001-5512-0083^{9,11}](https://orcid.org/0000-0001-5512-0083),
- [Anders J. Perrone](#)^{9,11},
- [Michaela Cordova](#)¹¹,

- [Olivia Doyle](#) ORCID: [orcid.org/0000-0003-3957-2949¹¹](https://orcid.org/0000-0003-3957-2949),
- [Lucille A. Moore](#) ORCID: [orcid.org/0000-0002-7665-2713¹¹](https://orcid.org/0000-0002-7665-2713),
- [Gregory M. Conan](#)^{9,11},
- [Johnny Uriarte](#)¹¹,
- [Kathy Snider](#)¹¹,
- [Benjamin J. Lynch](#) ORCID: [orcid.org/0000-0001-7310-2022^{9,12}](https://orcid.org/0000-0001-7310-2022),
- [James C. Wilgenbusch](#)^{9,12},
- [Thomas Pengo](#)⁷,
- [Angela Tam](#) ORCID: [orcid.org/0000-0001-6752-5707^{13,14,15,16}](https://orcid.org/0000-0001-6752-5707),
- [Jianzhong Chen](#)^{13,14,15,16},
- [Dillan J. Newbold](#)⁶,
- [Annie Zheng](#)⁶,
- [Nicole A. Seider](#) ORCID: [orcid.org/0000-0003-1511-1515⁶](https://orcid.org/0000-0003-1511-1515),
- [Andrew N. Van](#)^{6,17},
- [Athanasia Metoki](#)⁶,
- [Roselyne J. Chauvin](#) ORCID: [orcid.org/0000-0001-5786-4705⁶](https://orcid.org/0000-0001-5786-4705),
- [Timothy O. Laumann](#)¹,
- [Deanna J. Greene](#) ORCID: [orcid.org/0000-0002-0739-7771¹⁸](https://orcid.org/0000-0002-0739-7771),
- [Steven E. Petersen](#)^{6,17,19,20,21},
- [Hugh Garavan](#)²²,
- [Wesley K. Thompson](#)²³,
- [Thomas E. Nichols](#) ORCID: [orcid.org/0000-0002-4516-5103²⁴](https://orcid.org/0000-0002-4516-5103),
- [B. T. Thomas Yeo](#)^{13,14,15,16,25,26},
- [Deanna M. Barch](#) ORCID: [orcid.org/0000-0003-1693-8506^{1,21}](https://orcid.org/0000-0003-1693-8506),
- [Beatriz Luna](#)^{3,4},
- [Damien A. Fair](#) ORCID: [orcid.org/0000-0001-8602-393X^{9,10,27 na2}](https://orcid.org/0000-0001-8602-393X)
&
- [Nico U. F. Dosenbach](#) ORCID: [orcid.org/0000-0002-6876-7078^{6,17,19,28,29 na2}](https://orcid.org/0000-0002-6876-7078)

Nature volume **605**, page E11 (2022)

- 508 Accesses
- 2 Altmetric

- [Metrics details](#)

Subjects

- [Cognitive neuroscience](#)
- [Psychology](#)

The [Original Article](#) was published on 16 March 2022

Correction to: *Nature* <https://doi.org/10.1038/s41586-022-04492-9>

Published online 16 March 2022

Due to a tracking error during the peer review process, this article was published with the wrong received date. This article was received 19 May 2021, not 20 August 2020. Further, this article was originally published under the standard Springer Nature license (© The Author(s), under exclusive licence to Springer Nature Limited). It is now available as an open-access paper under a Creative Commons Attribution 4.0 International license, © The Author(s). The error has been corrected in the online version of the article

Author information

Author notes

1. These authors contributed equally: Scott Marek, Brenden Tervo-Clemmens
2. These authors jointly supervised this work: Damien A. Fair, Nico U. F. Dosenbach

Authors and Affiliations

1. Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA

Scott Marek, Alexander S. Hatoum, Meghan Rose Donohue, Ryland L. Miller, Sridhar Kandala, Timothy O. Laumann & Deanna M. Barch

2. Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Brenden Tervo-Clemmens

3. Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA

Brenden Tervo-Clemmens & Beatriz Luna

4. Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA

Finnegan J. Calabro, William Foran & Beatriz Luna

5. Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA, USA

Finnegan J. Calabro

6. Department of Neurology, Washington University School of Medicine, St Louis, MO, USA

David F. Montez, Benjamin P. Kay, Ryland L. Miller, Dillan J. Newbold, Annie Zheng, Nicole A. Seider, Andrew N. Van, Athanasia Metoki, Roselyne J. Chauvin, Steven E. Petersen & Nico U. F. Dosenbach

7. University of Minnesota Informatics Institute, University of Minnesota, Minneapolis, MN, USA

Timothy J. Hendrickson & Thomas Pengo

8. Department of Psychology, University of Minnesota, Minneapolis, MN, USA

Stephen M. Malone

9. Masonic Institute for the Developing Brain, University of Minnesota Medical School, Minneapolis, MN, USA

Eric Feczko, Oscar Miranda-Dominguez, Eric A. Earl, Anders J. Perrone, Gregory M. Conan, Benjamin J. Lynch, James C. Wilgenbusch & Damien A. Fair

10. Department of Pediatrics, University of Minnesota Medical School, Minneapolis, MN, USA

Eric Feczko, Oscar Miranda-Dominguez & Damien A. Fair

11. Department of Psychiatry, Oregon Health and Science University, Portland, OR, USA

Alice M. Graham, Eric A. Earl, Anders J. Perrone, Michaela Cordova, Olivia Doyle, Lucille A. Moore, Gregory M. Conan, Johnny Uriarte & Kathy Snider

12. Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN, USA

Benjamin J. Lynch & James C. Wilgenbusch

13. Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Angela Tam, Jianzhong Chen & B. T. Thomas Yeo

14. Centre for Sleep and Cognition, National University of Singapore, Singapore, Singapore

Angela Tam, Jianzhong Chen & B. T. Thomas Yeo

15. Centre for Translational MR Research, National University of Singapore, Singapore, Singapore

Angela Tam, Jianzhong Chen & B. T. Thomas Yeo

16. N.1 Institute for Health, Institute for Digital Medicine, National University of Singapore, Singapore, Singapore

Angela Tam, Jianzhong Chen & B. T. Thomas Yeo

17. Department of Biomedical Engineering, Washington University in St Louis, St Louis, MO, USA

Andrew N. Van, Steven E. Petersen & Nico U. F. Dosenbach

18. Department of Cognitive Science, University of California San Diego, La Jolla, CA, USA

Deanna J. Greene

19. Department of Radiology, Washington University School of Medicine, St Louis, MO, USA

Steven E. Petersen & Nico U. F. Dosenbach

20. Department of Neurological Surgery, Washington University School of Medicine, St Louis, MO, USA

Steven E. Petersen

21. Department of Psychological and Brain Sciences, Washington University in St Louis, St Louis, MO, USA

Steven E. Petersen & Deanna M. Barch

22. Department of Psychiatry, University of Vermont, Burlington, VT, USA

Hugh Garavan

23. Division of Biostatistics, University of California San Diego, La Jolla, CA, USA

Wesley K. Thompson

24. Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Thomas E. Nichols

25. Integrative Sciences and Engineering Programme, National University of Singapore, Singapore, Singapore

B. T. Thomas Yeo

26. Martins Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA

B. T. Thomas Yeo

27. Institute of Child Development, University of Minnesota Medical School, Minneapolis, MN, USA

Damien A. Fair

28. Program in Occupational Therapy, Washington University School of Medicine, St Louis, MO, USA

Nico U. F. Dosenbach

29. Department of Pediatrics, Washington University School of Medicine, St Louis, MO, USA

Nico U. F. Dosenbach

Corresponding authors

Correspondence to [Scott Marek](#), [Brenden Tervo-Clemmens](#), [Damien A. Fair](#) or [Nico U. F. Dosenbach](#).

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[Reprints and Permissions](#)

About this article

Cite this article

Marek, S., Tervo-Clemmens, B., Calabro, F.J. *et al.* Publisher Correction: Reproducible brain-wide association studies require thousands of individuals. *Nature* **605**, E11 (2022). <https://doi.org/10.1038/s41586-022-04692-3>

- Published: 09 May 2022
- Issue Date: 26 May 2022
- DOI: <https://doi.org/10.1038/s41586-022-04692-3>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04692-3>

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