protein made by humans, mice and 19 other widely studied organisms. Over the months that followed, the catalogue swelled to around 1 million structures.

"We're bracing ourselves for the release of this huge trove," says Christine Orengo, a computational biologist at University College London, who has used the AlphaFold database to identify new families of proteins. "Having all the data predicted for us is just fantastic."

High-quality structures

The release of AlphaFold last year made a splash in the life-sciences community, whose members have since been scrambling to use the tool. The network produces highly accurate predictions of many proteins' structures. It also provides information about the accuracy of its predictions, so researchers know whether they can be relied on. Conventionally, scientists have needed to use time-consuming and costly experimental methods such as X-ray crystallography and cryo-electron microscopy to solve protein structures.

According to EMBL–EBI, around 35% of the more than 214 million predictions are deemed to be highly accurate, which means they are as good as experimentally determined structures. Another 45% are considered to be accurate enough for many applications.

Many AlphaFold structures are good enough to replace experimental structures for some applications. In other cases, researchers use AlphaFold predictions to validate and make sense of experimental data. Poor predictions are often obvious, and some of them are caused by intrinsic disorder in the protein itself that means it has no defined shape — at least, not without other molecules present.

The 200 million predictions released last week are based on the sequences in another database, called UniProt. It's likely that scientists will have already had an idea about the shapes of some of these proteins, because they are included in databases of experimental structures or resemble other proteins in such repositories, says Eduard Porta Pardo, a computational biologist at Josep Carreras Leukaemia Research Institute (IJC) in Barcelona, Spain.

But such entries tend to be skewed towards human, mouse and other mammalian proteins, Porta says. It's likely that the AlphaFold dump will add significant knowledge, because it includes a diverse set of organisms. "It's going to be an awesome resource," says Porta.

Because AlphaFold's software has been available for a year, researchers have already had the capacity to predict the structure of any protein they wish. But many say that the availability of predictions in a single database will save researchers time, money — and faff. "It's another barrier of entry that you remove," says Porta. "I've used a lot of AlphaFold models. I have not ever run AlphaFold myself."

Jan Kosinski, a structural modeller at EMBL Hamburg in Germany who has been running the AlphaFold network over the past year, can't wait for the database expansion. His team once spent three weeks predicting the proteome — the set of all of an organism's proteins — of a pathogen. "Now we can just download all the models," he said at the briefing.

DeepMind chief executive Demis Hassabis.

JUNG YEON-JE/AFP/GETTY

Having almost every known protein in the database will also make new types of study possible. Orengo and her team have used the AlphaFold database to identify new protein families, and they will now do this on a much larger scale. They will also use the expanded repository to help them to understand the evolution of proteins with helpful properties — such as the ability to consume plastic — or worrying ones, like those that can drive cancer. The identification of distant relatives of these proteins in the database can pinpoint the basis for their properties.

Martin Steinegger, a computational biologist at Seoul National University who helped to develop a cloud-based version of AlphaFold, is excited about seeing the database expand. But he says that researchers are still likely to need to run the AI network themselves. Increasingly, people are using AlphaFold to determine how proteins interact, and such predictions are not in the database. Other predictions that are not there include microbial proteins identified by sequencing genetic material from soil, ocean water and other 'metagenomic' sources.

Some sophisticated applications of the expanded AlphaFold database might also depend on downloading its entire 23-terabyte contents, which won't be feasible for many teams, Steinegger says. Cloud-based storage could also prove costly. Steinegger has co-developed a software tool called FoldSeek that can quickly find structurally similar proteins and which should also be able to squash the AlphaFold data down.

Even with almost every known protein included, the AlphaFold database will need updating as new organisms are discovered. AlphaFold's predictions can also be improved as new structural information becomes available. Hassabis says DeepMind hopes to update the database annually. His hope is that the repository will have a lasting impact on the life sciences. "It's going to require quite a big change in thinking."

HOW LONG IS COVID INFECTIOUS? WHAT SCIENTISTS KNOW SO FAR

People with SARS-CoV-2 are told to isolate for a few days. But some can pass on the virus for much longer.

By David Adam

When the US Centers for Disease Control and Prevention (CDC) halved its recommended isolation time for people with COVID-19 to five days back in December, it said that the change was motivated by science. Specifically, the CDC said that most SARS-CoV-2 transmission occurs early in the course of the illness, in the one to two days before the onset of symptoms and for two to three days after.

Many scientists disputed that decision then and they continue to do so. Such dissent is bolstered by a series of studies confirming that many people with COVID-19 remain infectious well into the second week after they first experience symptoms. Reductions in the length of the recommended isolation period — now common worldwide — are driven by politics, they say, rather than any reassuring new data.

"The facts of how long people are infectious for have not really changed," says Amy Barczak, an infectious-disease specialist at Massachusetts General Hospital in Boston. "There is not data to support five days or anything

shorter than ten days" of isolation. Barczak's own research, published on the medRxiv preprint server, suggests that one-quarter of people who have caught the Omicron variant of SARS-CoV-2 could still be infectious after eight days (J. Boucau *et al*. Preprint at medRxiv https://doi.org/gp3xcd; 2022).

A numbers game

Although the question is simple — for how long is someone with COVID-19 contagious? — the answer is complicated. "We always think of it as a black-and-white thing … if somebody's infectious or not infectious — but in reality, it's a numbers game and a probability," says Benjamin Meyer, a virologist at the University of Geneva in Switzerland.

And that numbers game has shifting rules and baselines. Emerging variants, vaccinations and varying levels of natural immunity elicited by previous infection can all influence how quickly someone can clear the virus from their system, Meyer says, and this ultimately dictates when they stop being infectious. Behavioural factors matter as well. People who feel unwell tend to socialize less, he adds, so the severity of someone's symptoms can influence how likely they are to infect others.

Something most scientists are confident about is that PCR tests can return a positive result even after someone is no longer infectious. This probably occurs when the tests, which detect viral RNA, pick up non-infectious remnants left behind after most of the live virus has been eliminated.

By contrast, lateral flow (or 'rapid antigen') tests offer a better guide to infectiousness, because they detect proteins produced by actively replicating virus.

"There's still all of these things that we're not exactly sure about, but if I had to sum it up in one very concise message, it would be that if you're antigen positive, you shouldn't go out and interact closely with people who you don't want to be infected," says Emily Bruce, a microbiologist and molecular geneticist at the University of Vermont in Burlington.

What about somebody who has tested negative on a lateral flow test for a few days but still has a fever and a hacking cough? Bruce says it's important to remember that although lingering symptoms might look and sound serious, they do not indicate continued infectiousness.

"You can definitely have symptoms for longer than you test positive on lateral flow," she says. "And I think that's because many of the symptoms are caused by the immune system and not directly by the virus itself."

Transmission tests

In countries such as the United Kingdom, the relaxation of the isolation guidelines coincided with the withdrawal of free lateral flow tests. So, assuming that many of the people who follow the new recommendations are going to stop isolating after five days, without testing, scientists have been investigating how many people with COVID-19 are likely to remain infectious after this point.

AGUSTIN MARCARIAN/REUTERS/ALAMY

It is difficult to measure how long a person with COVID-19 will remain infectious.

It's not practical to track direct onward transmission of the virus from large numbers of people and to measure how it reduces over time, so researchers instead rely on proxy measurements to determine the point at which they expect people to stop being contagious.

Researchers with access to a high-security biosafety level 3 laboratory — as Barczak has — can do this by running experiments to test whether live SARS-CoV-2 can be cultured from samples taken from patients over several days.

"We think of it as a black-and-white thing if somebody's infectious or not … but in reality, it's a numbers game."

"If you're still shedding virus that we can culture out of your nose, there's at least a good chance you're still infectious to other people," she says. As different variants have emerged and various research groups have done these experiments, Barczak says, a consensus has emerged that it's very unusual for people to shed culturable virus after ten days. "So, it's very unusual for people to stay infectious after ten days," she says.

Other studies use levels of viral RNA measured by PCR tests to infer whether someone is infectious. This makes it easier to work with large sample sizes. For example, a project run by the Crick Institute and University College Hospital, both in London, can draw on PCR test results for more than 700 participants, obtained from when symptoms developed.

A study based on this group suggests that significant numbers of people retain viral loads high enough to trigger onward transmission at days seven to ten. The study was published on the medRxiv preprint server on 10 July (H. Townsley *et al*. Preprint at medRxiv https://doi.org/h6fc; 2022).

"We're not measuring live virus, but there is now a huge amount of work in the literature that provides a pretty good mapping of what constitutes a viral load likely to yield infectious virus," says David LV Bauer, a virologist at the Crick Institute who is a co-investigator on that study. "So while it's not a perfect picture, it's a reasonable one."

Rebound phenomenon

Yonatan Grad, an infectious-disease specialist at the Harvard T.H. Chan School of Public Health in Boston, Massachusetts, who has worked on similar PCR-based studies, agrees that ten days is a useful rule of thumb for when people should no longer be contagious. But he cautions that a small number of people could still be infectious beyond that point.

Some such cases have been linked to the antiviral drug Paxlovid (nirmatrelvir–ritonavir), he says. "There's a rebound phenomenon where people will see that their symptoms seem to resolve and they may even test negative on a rapid test, but then a few days later symptoms and the virus come back."

Barczak says this is one of the key questions that researchers are now studying. "Antivirals change the dynamics of symptoms, change the dynamics of the immune response and change the dynamics of how you shed," she says. "I think this is really important, because people are out in the world thinking they're not infectious after ten days. But, if they have Paxlovid, rebound they might be."