

FUTURE PANDEMIC PREPAREDNESS, POWERED BY GENOMICS

Next-generation sequencing is helping to steer the global response to SARS-CoV-2, and will help the world to prepare for future pandemics.



For **illumina**[®]

by **nature**research
CUSTOM MEDIA



Research into novel pandemic viruses is based on knowledge of their genome.

FUTURE PANDEMIC PREPAREDNESS, POWERED BY GENOMICS

Using **NEXT-GENERATION SEQUENCING TOOLS**, scientists are exploring how SARS-CoV-2 interacts with the immune system to better understand the disease, identify those at higher risk, and minimize its impact. Next-generation sequencing is helping to steer the global response to SARS-CoV-2, and will help the world to prepare for future pandemics.

“Genomics provides us with a ‘molecular Esperanto’—a shared language that everyone understands,” says Vitali Sintchenko, a public health microbiologist at Sydney Medical School and Director of the Centre for Infectious Diseases and Microbiology (CIDM) Public Health.

SARS-CoV-2, the novel coronavirus that causes COVID-19, emerged in China

in December 2019. Since then, researchers around the world have used genomics to understand its spread and evolution, uncovering vital clues that are helping to inform public health decision-making and the development of new diagnostics, therapeutics and vaccines.

“Genomics creates unlimited opportunities for researchers to study important issues about changes to the virus in real-time,”

Sintchenko adds.

In the longer term, next-generation sequencing (NGS) technologies offer huge potential to improve the response to emerging infectious diseases.

“We’ve got a bit of a head start now, and we should really build on this,” says Martin Hibberd, an expert in emerging infectious diseases at the London School of Hygiene and Tropical Medicine. “The routine

use of genomics as part of global surveillance systems would help us to spot potential threats earlier and faster.”

A wealth of information

Hibberd was part of the team in 2003 that sequenced the original SARS-CoV coronavirus, the causative agent of severe acute respiratory syndrome (SARS). “It took us about three months, which we thought was

really fast," he recalls. "This time around we're talking about weeks, which just shows how much the technology has advanced."

Before the end of January 2020, NGS had revealed the full sequence of the SARS-CoV-2 genome.¹ This information enabled the design of the polymerase chain reaction (PCR) test, routinely used to detect the infection. Frontline PCR testing only determines the presence of the virus, but genomics data can be extracted from the same sample, opening a multitude of new opportunities to study COVID-19.

"The beauty of sequencing is that you can see everything that has DNA or RNA, so you can take a snapshot of the entire microbiome as well as the host genome," says Victoria Parikh, a cardiologist at Stanford University School of Medicine. "That gives you an incredibly rich dataset to better understand what's going on."

Parikh was part of a team that developed a scalable, high-throughput protocol for the rapid generation of whole host and viral genome, major histocompatibility (HLA) sequences, and host transcriptome, from a single nasopharyngeal swab (the study is yet to be peer-reviewed).²

"We've been amazed by the amount of data we can generate from this really small quantity of material," says Parikh. "You can quickly collect it and produce all this data. Then it's just a question of what you do with it."

Coordinating the collection and sharing of genomics data is accelerating research into SARS-CoV-2, and will also help in tackling future infectious disease outbreaks.

Tracking pathogen spread

Researchers are using NGS to examine pathogen transmission networks at a resolution not previously possible. "If all the

sequences in a local cluster look very similar to each other, then you know they are the result of community transmission," Hibberd explains.

Genomic surveillance is already used to monitor public health threats such as foodborne pathogens, influenza, or the tuberculosis bacterium. "When the pandemic hit, we were able to rapidly mobilize our established infrastructure and capacity to do large-scale sequencing of SARS-CoV-2," says Sintchenko, who was part of a collaboration that used near-real-time viral genome sequencing during the first wave of infections in Australia.³ "We helped resolve epidemiological uncertainties by linking ambiguous cases with established transmission pathways—we could also separate local transmission events from those that were imported."

"SUDDENLY, PATHOGEN GENOMICS MOVED FROM A NICHE APPLICATION TO THE CENTRE OF CONTROL EFFORTS."

Genomic surveillance is now an integral part of tracing SARS-CoV-2 transmission, supporting major decision-making such as lockdowns, travel restrictions and infection control in hospitals throughout Australia. "Suddenly, pathogen genomics moved from a niche application for a couple of communicable diseases to the centre of control efforts on a national scale," says Sintchenko.

A team of researchers in San Francisco is also realizing the benefits of additional genomic information. "We've sequenced thousands of viral genomes, which is helping to uncover new details about community transmission pathways," says Joe DeRisi, a molecular biologist and biochemist at UCSF, and

co-president of the Chan Zuckerberg Biohub. "And we've been involved with our county department of health in about a dozen events where that information was actionable."

For example, they found that in April 2020 SARS-CoV-2 was freely circulating among workers on low-incomes living in a densely populated district.⁴ This led to the introduction of city- and state-wide legislation that gives employees the right to isolate and quarantine without fear of losing their job.

Monitoring and diagnosing infections

GISAID, which was originally set up to share influenza data, opened its repositories to SARS-CoV-2. By December 2020, it contained more than 235,000 genome sequences from across the world. "It's amazing to see that happening in real time," says Hibberd. "The number of sequences continues to rise rapidly, which is enabling researchers to closely monitor evolution and spread."

An analysis of early genome sequences, which is yet to be peer-reviewed,⁵ revealed that SARS-CoV-2 was showing relatively few evolutionary selection pressures. "This virus seems fully adapted to humans," says Hibberd.

The coronavirus genome is unlikely to remain stable, however. "We expect to see resistance mutations as soon as we have effective therapies and vaccines," says Hibberd. "We will need to monitor the virus for any genetic changes that could impact their long-term effectiveness."

In August 2020, researchers in Hong Kong revealed the first confirmed case of SARS-CoV-2 reinfection, determining that the patient's two episodes were caused by strains with different genomes.⁶ "This added further complexity to the pandemic,"

says Sintchenko. "It's stimulated discussions into herd immunity, the duration of protection, and the role of vaccines in managing the disease."

"THIS VIRUS SEEMS FULLY ADAPTED TO HUMANS."

Identifying specific mutations in the SARS-CoV-2 genome may also have implications for diagnostic testing. "We identified⁷ a change that ablates one of the sites that's commonly used for PCR tests," says DeRisi. "This could lead to false negatives for tests that rely on a single site: the virus won't be detected in cases where it's really there."

In this regard, NGS protocols could provide a backup for diagnostics laboratories to ensure that PCR testing is working well. In June, the US Food and Drug Administration granted Emergency Use Authorization for Illumina's NGS-based COVID-19 diagnostic test; other tests have followed. NGS protocols have the potential to be scaled up to enable more high-volume screening to support a return to work and school.

Coupling a diagnosis of SARS-CoV-2 infection with its genome sequence opens further opportunities to understand its spread and evolution. Linking these data with electronic health records can also enable other types of investigation: such as into the genetic factors underlying the huge variability in clinical symptoms among infected individuals.

In the clinic, NGS protocols can help doctors diagnose COVID-19 patients who might also have other infections. Clinical symptoms of many respiratory pathogens are very similar, making it hard to distinguish the causal agent(s). If co-infection testing identifies SARS-CoV-2, the patient's treatment regimen



Lab tests rely on genomic surveillance work to detect and trace SARS-CoV-2.

will change. For example, those with severe COVID-19 are at risk of pneumonia if they are put on a ventilator.⁸

Unbiased sequence analysis

Metagenomics NGS is the comprehensive examination of all organisms in a sample, allowing researchers to work “free from any preconceived notions of what might be there”, says DeRisi.

The approach can be used to identify both emerging and existing pathogens—and was how SARS-CoV-2 was first identified. This technique was also used to discover the culprit behind an outbreak of meningitis in children in Bangladesh.⁹ “It turned out to be the mosquito-borne Chikungunya virus,” says DeRisi, who led the study. “This wasn’t even on the doctors’ radar, and they’d been treating it with antibiotics.”

Establishing a global

surveillance network will be crucial if novel pathogens with pandemic potential are to be identified swiftly in the future. “Ideally, you’d have a global network of hundreds of sites that are feeding in a constant stream of molecular real-time data from humans, and potentially animals too,” says DeRisi. “As conditions change, or a new set of patients shows up in a particular region, then you can immediately react.”

“THE SPEED OF KNOWLEDGE DISCOVERY FOR THIS DISEASE IS UNPRECEDENTED.”

Such a network would rely on leveraging clinical, public health and microbiological expertise. But metagenomics data analysis typically requires access to local IT infrastructure

and bioinformatics proficiency, which can present obstacles, especially in places where resources are limited.

To help overcome this challenge, DeRisi’s team has created IDseq, an automated cloud-based metagenomics platform that can quickly assemble NGS data into genomes.¹⁰ “It does the heavy lifting for you,” DeRisi explains. “And then it puts it into a graphical interface that anybody can use, making it fast and accessible around the world.”

Towards the end of January 2020, a facility in Cambodia used the system to characterize the country’s first case of COVID-19 (the study is yet to be peer-reviewed).¹¹ “They had a positive signal from PCR, so they sequenced and analysed the sample using IDseq, which we had set them up with a few weeks earlier,” recalls DeRisi,

who was involved in the study. “They captured the entire genome, one of the first outside China, and placed it on the family tree. It was a great validation for the technology.”

Now and in the future

Studying the genetic makeup of SARS-CoV-2 has been a crucial driver in shaping the response to the current pandemic, from understanding viral transmission pathways to tracking its evolution.

“The speed of knowledge discovery for this disease is unprecedented,” says Sintchenko. “The attention is huge, and data are freely available for everyone, with the skill and tools, to generate and test different hypotheses.”

Given that globalization has increased the threat from infectious diseases, the onus is on the world to learn from COVID-19. “We need to be ready for the next pandemic,” says Sintchenko. “We need to have the global infrastructure in place to manage things better, more efficiently, with less damage to our economies—and genomics can help us to do that.” ■

REFERENCES

1. Zhu, N., et al. *N Engl J Med* **382**: 727-733 (2020).
2. Gorzynski, J. E., et al. *medRxiv* 2020.07.27.20163147 [preprint].
3. Rockett, R. J., et al. *Nat Med* **26**: 1398-1404 (2020).
4. Chamie, G., et al. *Clin Infect Dis*. [published online ahead of print, 2020 Aug 21]
5. Phelan, J., et al. *bioRxiv* 2020.04.28.066977 [preprint].
6. To, K. K., et al. *Clin Infect Dis* [published online ahead of print, 2020 Aug 25].
7. Vanaerschot, M. et al. *J Clin Microbiol* [published online ahead of print, 2020 Oct 16]
8. Luyt, C-E, et al. *Ann Intensive Care* **10**: 158 (2020).
9. Saha, S. et al. *mBio* **10** (6) e02877-19 (2019).
10. Kalantar, K. L., et al. *Gigascience* **9** (10):giaa111 (2020)
11. Manning, J. E., et al. *bioRxiv* 2020.03.02.968818 [preprint].

illumina®