BioTechniques

METAGENOMICS: PREVENTING FUTURE PANDEMICS

Metagenomic approaches have been key to successful tracing and outbreak management during the COVID-19 pandemic. How can we use this knowledge to better prepare, strategize and prevent future pandemics?

COVID-19 first emerged in Wuhan (Hubei, China) in December 2019 [1]. The novel coronavirus causing the disease, SARS-CoV-2, was first identified using metagenomic RNA sequencing [2] on 5 January 2020 by a team led by Zhang Yongzhen (Shanghai Public Health Clinical Center, China). On 10 January 2020, the novel coronavirus genome was posted publicly on virological.org and GenBank to assist investigations around the world [3].

COVID-19 spread rapidly amongst humans, and to other countries, leading to its classification as a pandemic by the WHO on 11 March 2020, just 2 months after the genomic information was shared. As a novel disease, there were no vaccine or targeted drugs to be used to treat or halt the spread of COVID-19, hence rapid diagnosis and isolation of patients became essential.

National lockdown, which included orders to stay and work at home, was the initial response by many countries to slow the rapid spread of the virus. Whilst this decision was met with a mixed response, it is hard to deny that this is an essential step to reduce mortality and prevent healthcare services from becoming overwhelmed.

This was supported by Flaxman *et al.*, who studied the effect of major interventions across 11 European countries for the period from the start of the COVID-19 epidemics in February 2020 until 4 May 2020, when lockdowns started to be lifted. The results indicate that major non-pharmaceutical interventions – and lockdowns in particular – have had a large effect on reducing transmission [4].

To the majority, lockdown and social distancing measures appear to be the only pandemic management strategy universally rolled out. As of 13 January 2021, there have been 92,148,761 confirmed cases of COVID-19, including 1,973,486 deaths, reported to the WHO [5]. What could have been done to keep these figures lower? Have we learned anything from previous pandemics? And, how can we prepare for future pandemics?

PROGRESS FROM PANDEMICS PAST

Few people reading this Technology News article will remember a pandemic on this scale, but history reveals that what we are currently experiencing is nothing uncommon. In the 20th century alone there have been three influenza pandemics – 1918, 1957 and 1968 (Figure 1). Promisingly, there was progression seen between these events – by 1957 there was a global network of laboratories linked to the World Influenza Research Centre in London (UK), which acted as a hub for research and virus tracking [6].

Interestingly, for all the advances made against infectious disease – health infrastructure and technology to name a few – our very growth in terms of population, migration, trade and urbanization has made us more vulnerable.

Ultimately, the goal is to detect, understand and contain infectious outbreaks at the earliest stage possible. This is fundamental in order to prevent and control outbreaks. A universal surveillance network has the potential to answer these needs.

Metagenomics is emerging as an important tool in biosurveillance, public health and clinical applications. Pandemic risk calculations employ technologies like metagenomics to trace the molecular changes in pathogens during their emergence, and mathematical models to assess risk. This combination of technologies enables us to predict an abundance of useful information – hot spots of emergence, populations at risk and the pathogens under genetic evolution.

The problem is that while the technology for surveillance is available, it is often restricted to the western hemisphere; with many diseases continuing to emerge in areas such as Southeast Asia, it is vital to ensure metagenomics and predictions of pandemic risk are shared.

Whilst this just skims the surface of previous pandemics and ideal responses, it is important to note that strides have ▶

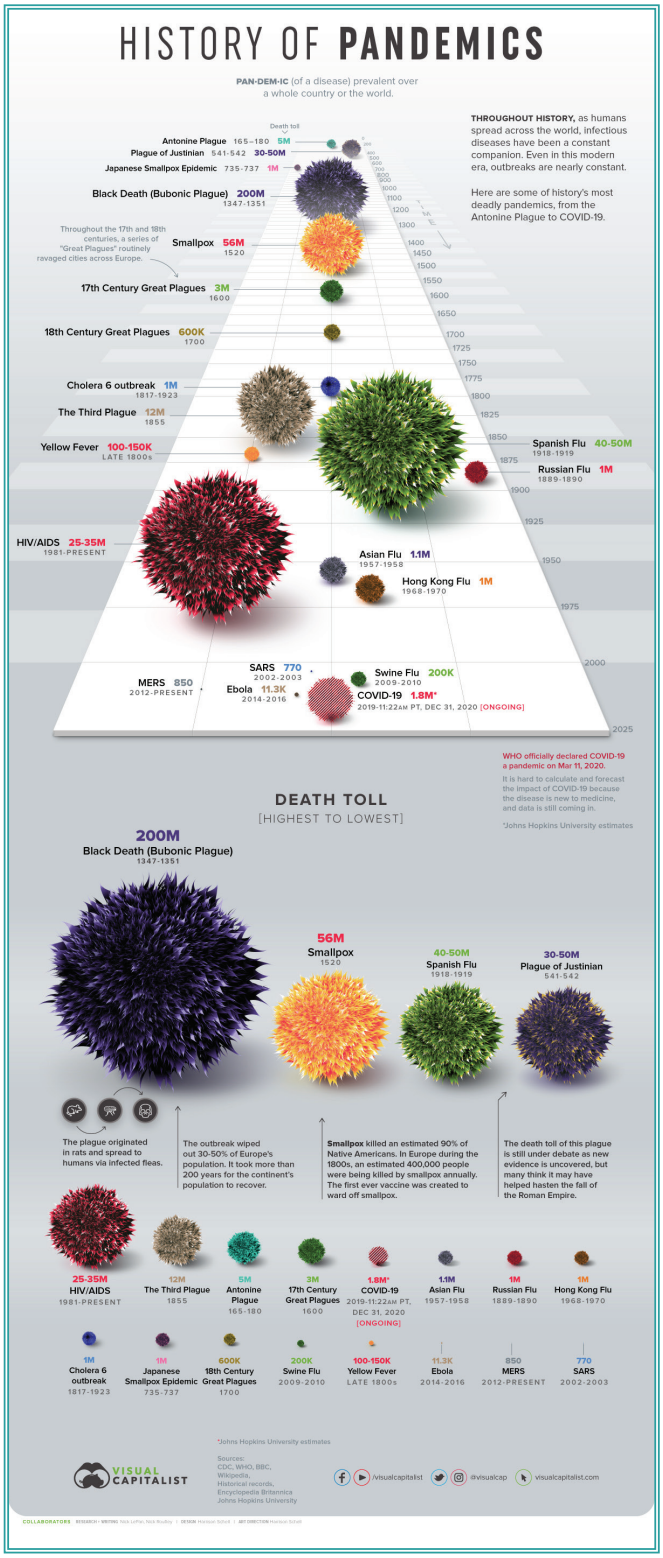

Figure 1. The history of pandemics.
Reprinted from [7].

been made in terms of pandemic response; however, this piece is to highlight the potential of metagenomics in preventing a future pandemic.

HOW CAN WE DETECT A VIRUS?

Traditionally, detection of a pathogen for public health infectious disease surveillance relies upon the identification of pre-established markers of a particular disease, through assays or screening. However, through the use of next-generation sequencing (NGS), metagenomics has the ability to detect all microorganisms in a sample, regardless of whether they are known or novel pathogens [8].

There are two defined strategies through which to detect disease outbreaks. The first is syndromic surveillance, which relies on health indicators such as symptoms and patterns, before a laboratory-confirmed diagnosis is made [9]. However, this is not always an accurate method to indicate infection spread, due to the lack of a definitive disease diagnosis.

The more reliable strategy is laboratory-based surveillance, in which a number of methods can be implemented to detect and confirm pathogen presence. Within the realm of traditional pathogen detection, samples can often be left undiagnosed when there is a failure to detect a causative agent. Instead of using assays that are targeted to search for a specific pathogen, metagenomics has the potential to entirely overhaul this process, detecting the presence of all microorganisms through a single sequencing technique, and without the need for culture.

Further to this, the genomic information that can be gathered by metagenomic NGS can be used in ways beyond informing outbreak investigations, including to identify virulence genes and predict potential antibiotic resistance [8].

HOW HAVE METAGENOMICS BEEN PROPOSED TO PREVENT PANDEMICS?

The idea of using metagenomics to scan for and catch novel viruses as they arise – allowing researchers to stymie the spread of the virus early, thereby preventing the possibility of a pandemic before it has begun – is not a new one. In fact, a 2016 paper published in *PNAS* warned of a SARS-like virus *"poised for human"* emergence [10].

The paper highlighted WIV1-CoV, isolated from Chinese horseshoe bat populations in a previous metagenomic study (Figure 2) [11], as being of particularly high risk for emergence in humans after both chimeric and full-length zoonotic versions of the virus were shown to be capable of replicating *in vivo* in humans. Sound familiar?

The paper highlighted that – while metagenomic screens had identified that SARS-like viruses were circulating in these bat populations and that species with the potential to evolve into human infectious strains had been identified in previous studies [11] – these observations alone were not enough. Commenting on the *PNAS* paper in another article for the same issue, Vincent Racaniello (Columbia University Medical Center, NY, USA) stated that *"gazing at viral sequences has its limits; experiments need to be done"* [12].

The authors of the PNAS paper outlined an approach to take metagenomic data and examine them experimentally to determine the probability of a virus becoming infectious to humans, the likely severity of the resultant infection and our preparedness to deal with the emergent species if it does [10]. To explore the

severity and probability they conducted a series of mouse model experiments involving transgenic mice, manipulated to express the human ACE2 receptor – the receptor bound by coronavirus spike proteins.

To assess our preparedness for the emergence of WIV1-CoV into human populations, the authors tested the efficacy of antibodies generated to protect against SARS-CoV in blocking infection by WIV1-CoV, first in cell culture and then in the transgenic mouse model. This identified an antibody that did prevent replication of the virus and protect against sever disease; however, preliminary attempts to determine whether vaccines composed of inactivated SARS-CoV would be effective against WIV1-CoV proved negative [10].

Ultimately, this case study highlighted that there are viruses on the cusp of making the leap from animal to human pathogenicity, that we were not prepared for these viruses with an easily adaptable vaccine and that potential treatments for these viruses could include broadly neutralizing antibodies.

Racaniello, drawing next steps from these key points, laid out several recommendations for further study [12]. First, a study to identify a panel of antibodies effective in preventing the invasive action of the spike protein common to coronaviruses. Next, the identification of the genome alterations required for WIV1-CoV to become infectious to humans – an action that should be applied to all viruses identified as at risk of infecting humans. Finally, an examination of the mechanism by which pathogenicity of these viruses could increase.

WHAT WENT WRONG?

So, with these actions highlighted and the risk of future pandemics laid bare, what went wrong? To investigate many of these actions would have required gain-of-function experiments on these viruses. At the time of the paper's publication, there was heated debate about the risks of these experiments, with opposition occasionally citing apocalyptic scenarios worthy of Hollywood. These arguments, whilst founded on little actual evidence [13], were nonetheless captivating enough to provoke the US government to issue a moratorium severely restricting these experiments [14].

While this restriction was lifted in December 2017 [15], this period would have delayed these crucial studies and dampened the impact of this paper, perhaps preventing the recommendations of Racaniello being pursued.

Meanwhile, debate on the gain-of-function studies still rages and can perturb researchers from undertaking them, with many legitimate reservations needing to be thoroughly addressed [16] and more sensationalist articles fanning the flames of debate [17]. While the debate surrounding the topic is by no means entirely one-sided in the virology community, instead of quashing the practice, protocols need to be put in place to ensure that these studies are conducted in the safest possible manner, in an open and accountable setting – actions that may take the heat out of the debate and allow safe progress in research that that could prove invaluable. For metagenomic scanning to be worth its salt, it needs to be accompanied by the ability to follow-up the findings in an effective but safe manner.

Figure 2. A Chinese horseshoe bat. Shutterstock.com.

THE ICELANDIC APPROACH

PCR testing was quickly established as the 'gold standard' of diagnosis, allowing for scalable, rapid and comparatively cheap results, with a good level of reliability. However, there are clear benefits in combining traditional laboratory methods with metagenomic NGS to understand diseases and improve outbreak response, as has been highlighted by an open-source project, Nextstrain [18].

Globally, the primary focus during the COVID-19 outbreak was on rapid testing, not whole-genome sequencing, with some countries completely polarized in their position on metagenomic analysis for tracing SARS-CoV-2 infection.

In May 2020, a team of researchers across countries in East Africa contributed to an article for BioTechniques questioning where the SARS-CoV-2 genomes from East Africa were [19]. At the time of writing the article, which was published on 15 May 2020, there were no publicly available SARS-CoV-2 genomes from East Africa – more than 10 weeks after the first reported COVID-19 case in Kenya.

As outlined in the article, a French study that was listed on bioRxiv – now published in Eurosurveillance [20] – demonstrated through NGS analysis that the first recorded cases of COVID-19 in France were not of Chinese origin, suggesting that SARS-CoV-2 was present in the country prior. This insight is crucial for biosurveillance purposes and outbreak management.

The Icelandic approach to COVID-19 focused heavily on scientific methods, encompassing metagenomic techniques, and has been hailed as *"pivotal"* in its contribution to understanding the pandemic [21]. deCODE (Reykjavik, Iceland), a human genomics company and subsidiary of Amgen (USA), offered their services to Iceland's Directorate of Health. Together, they tracked the health of every Icelander who tested positive for SARS-CoV-2, isolated and sequenced captured

COVID-positive samples and screened more than half of the 368,000 population for infection.

They both uncovered crucial insights about COVID-19 infection, including that almost half of infected people are asymptomatic, and prevented high numbers of deaths – reporting approximately 7 per 100,000 in comparison to the USA's approximate 80 per 100,00 – while still keeping their borders open to tourists.

METAGENOMIC PROTOCOLS IN THE COVID ERA

A study published in the Journal of Clinical Virology in October 2020 outlined the validation of a metagenomics protocol for coronavirus identification by simulating novel virus discovery and thus, providing a potential tool for pandemic preparedness [22].

The team used clinical samples containing the coronaviruses MERS-CoV, SARS-CoV and SARS-CoV-2 to perform their metagenomic protocol. To simulate novel virus discovery, the databases they used for classification contained only known viruses prior to the discovery of these three coronaviruses.

The resultant NGS reads enabled the identification of the coronaviruses as being novel, related to the coronaviruses whose genomic information was present in the databases, validating the protocol for novel coronavirus discovery.

In June 2020, a study published in Clinical Chemistry also outlined the detection of SARS-CoV-2 by metagenomic analysis. The researchers evaluated laboratory-confirmed COVID-positive and -negative samples with metagenomic NGS, comparing the readouts to a genomic database from 2019, which was created prior to SARS-CoV-2 discovery [23].

Through this method, they were able to identify the novel coronavirus within 36 hours. This study did have a poor sample size; however, the benefits of metagenomic sequencing were also highlighted. The team identified numerous other viruses present in the samples, providing a less targeted diagnosis route, identifying why the patient is ill and the best course for treatment.

METAGENOMIC PROTOCOLS POST-COVID

The extent and severity of the COVID-19 pandemic has some researchers looking to the future and questioning: where might the next pandemic come from and how can it be prevented?

An article published in *Nature* in March 2020 outlined the identification of two SARS-CoV-2-related coronaviruses in Malayan pangolins through metagenomic sequencing, one of which had a receptor-binding domain incredibly similar to that of SARS-CoV-2 [24].

The researchers suggest that since the SARS-CoV-2 outbreak has been associated with a seafood market, pangolins should be removed from markets as a preventative measure for a future coronavirus outbreak.

The potential of metagenomics in detecting and tracing novel coronaviruses is immense. Although metagenomics may be a relatively recent field, it is certainly not novel. Due to the immediate severity, rapid spread and relative lack of preparedness (in comparison to the vast knowledge that has been gained during the pandemic), there perhaps was not the time or resource to focus on metagenomic sequencing.

However, there is a successful example to be noted in the Icelandic approach. Through careful planning and strategy, combining traditional diagnostic tests with metagenomic approaches to trace the spread and find the root of COVID-19 infection, Iceland was able to contain and minimize the spread.

Written by Abigail Sawyer, Tristan Free & Joseph Martin

REFERENCES

1. Wang H, Li X, Zhang S, Wang L, Wu X, Liu J. The genetic sequence, origin, and diagnosis of SARS-CoV-2. *Eur. J. Clin. Microbiol. Infect. Dis.* 39, 1629–1635 (2020).
2. Wu F, Zhao S, Yu B *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269 (2020).
3. Virological.org. Novel 2019 Coronavirus Genome (Accessed 9 December 2020). https://virological.org/t/novel-2019-coronavirus-genome/319
4. Seth Flaxman S, Mishra S, Gandy A *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 584(7820), 257–261 (2020).
5. World Health Organisation. WHO Coronavirus Disease (COVID-19) Dashboard (Accessed 9 December 2020). https://covid19.who.int/
6. Jackson C. History lessons: the Asian flu pandemic. *Br. J. Gen. Pract.* 59(565), 622–623 (2009).
7. Visual Capitalist. Visualizing the history of pandemics (Accessed 9 December 2020). https://www.visualcapitalist.com/history-of-pandemics-deadliest/
8. Miller R, Montoya V, Gardy J, Patrick D, Tang P. Metagenomics for pathogen detection in public health. *Genome Med.* 5, 81 (2013).
9. Mandl K, Overhage J, Wagner M *et al.* Implementing syndromic surveillance: a practical guide informed by the early experience. *J. Am. Med. Inform. Assoc.* 11(2), 141–150 (2004).
10. Menachery V, Yount B Jr, Sims A *et al.* SARS-like WIV1-CoV poised for human emergence. *Proc. Natl Acad. Sci. USA* 113(11), 3048–3053 (2016).
11. Ge X-Y, Li J-L, Yang X-L *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 508, 535–538 (2013).
12. Racaniello V. Moving beyond metagenomics to find the next pandemic virus. *Proc. Natl Acad. Sci. USA* 113 (11), 2812–2814 (2016).
13. The apocalypse as a rhetorical device in the influenza virus gain-of-function debate. *Mbio* 5(5), e02062–14 (2014).
14. The White House President Barack Obama. Doing diligence to assess the risks and benefits of life sciences gain-of-function research (Accessed 9 December 2020). https://obamawhitehouse.archives.gov/blog/2014/10/17/doing-diligence-assess-risks-and-benefits-life-sciences-gain-function-research
15. Burki T. Ban on gain-of-function studies ends. *Lancet Infect. Dis.* 18(2), 148–149 (2018).
16. Imperiale M, Casadevall A. Rethinking gain-of-function experiments in the context of the COVID-19 pandemic. *Mbio.* 11(4), e01868–20 (2020).
17. Vox. Why some labs work on making viruses deadlier — and why they should stop (Accessed 9 December 2020). https://www.vox.com/2020/5/1/21243148/why-some-labs-work-on-making-viruses-deadlier-and-why-they-should-stop
18. Nextstrain. Nextstrain SARS-CoV-2 resources (Accessed 9 December 2020). https://nextstrain.org/sars-cov-2
19. *BioTechniques*. Where are the SARS-CoV-2 genomes from East Africa? (Accessed 9 December 2020). https://www.biotechniques.com/covid-19/opinion_where-are-the-sars-cov-2-genomes-from-east-africa/
20. Gambaro F, Behillil S, Baidaliuk A. Introductions and early spread of SARS-CoV-2 in France, 24 January to 23 March 2020. *Euro. Surveill.* 25(26), 2001200 (2020).
21. *Nature*. How Iceland hammered COVID with science (Accessed 9 December 2020). https://www.nature.com/articles/d41586–020–03284–3
22. Carbo E, Sidorov L, Zevenhoven-Dobbe J *et al.* Coronavirus discovery by metagenomic sequencing: a tool for pandemic preparedness. *J. Clin. Virol.* 131, 104594 (2020).
23. Peddu V, Shean R, Xie H *et al.* Metagenomic analysis reveals clinical SARS-CoV-2 infection and bacterial or viral superinfection and colonization. *Clin. Chem.* 66(7), 966–972 (2020).
24. Lam T, Jia N, Zhang Y-W *et al.* Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* 583, 282–285 (2020).