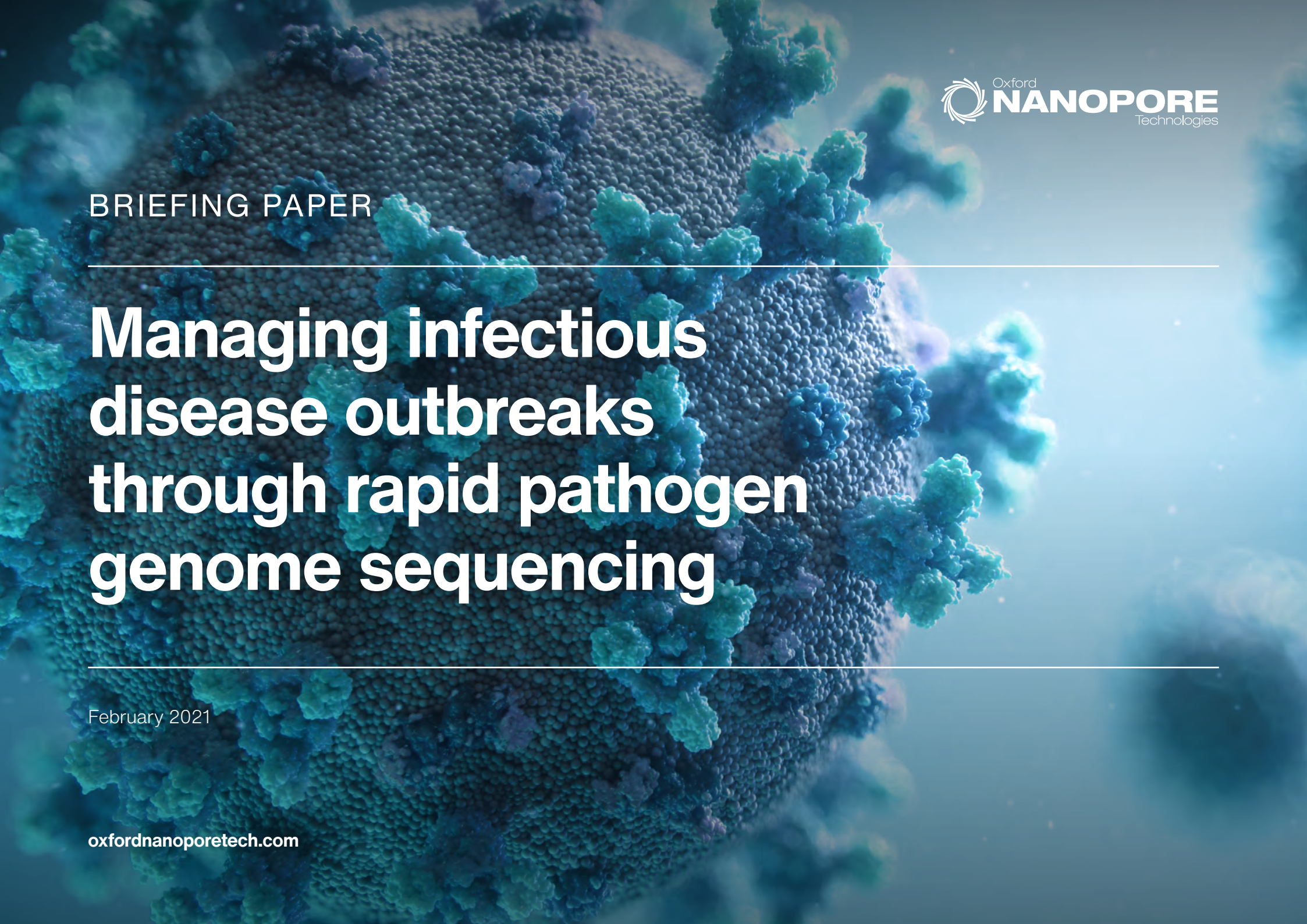


Oxford NANOPORE Technologies

BRIEFING PAPER

Managing infectious disease outbreaks through rapid pathogen genome sequencing

February 2021

oxfordnanoporetech.com

Introduction

With over two million attributed deaths to date and a projected economic cost of $28 trillion[1], the COVID-19 pandemic has refocused global attention on the acute, ever-present threat of infectious disease.

Executive summary

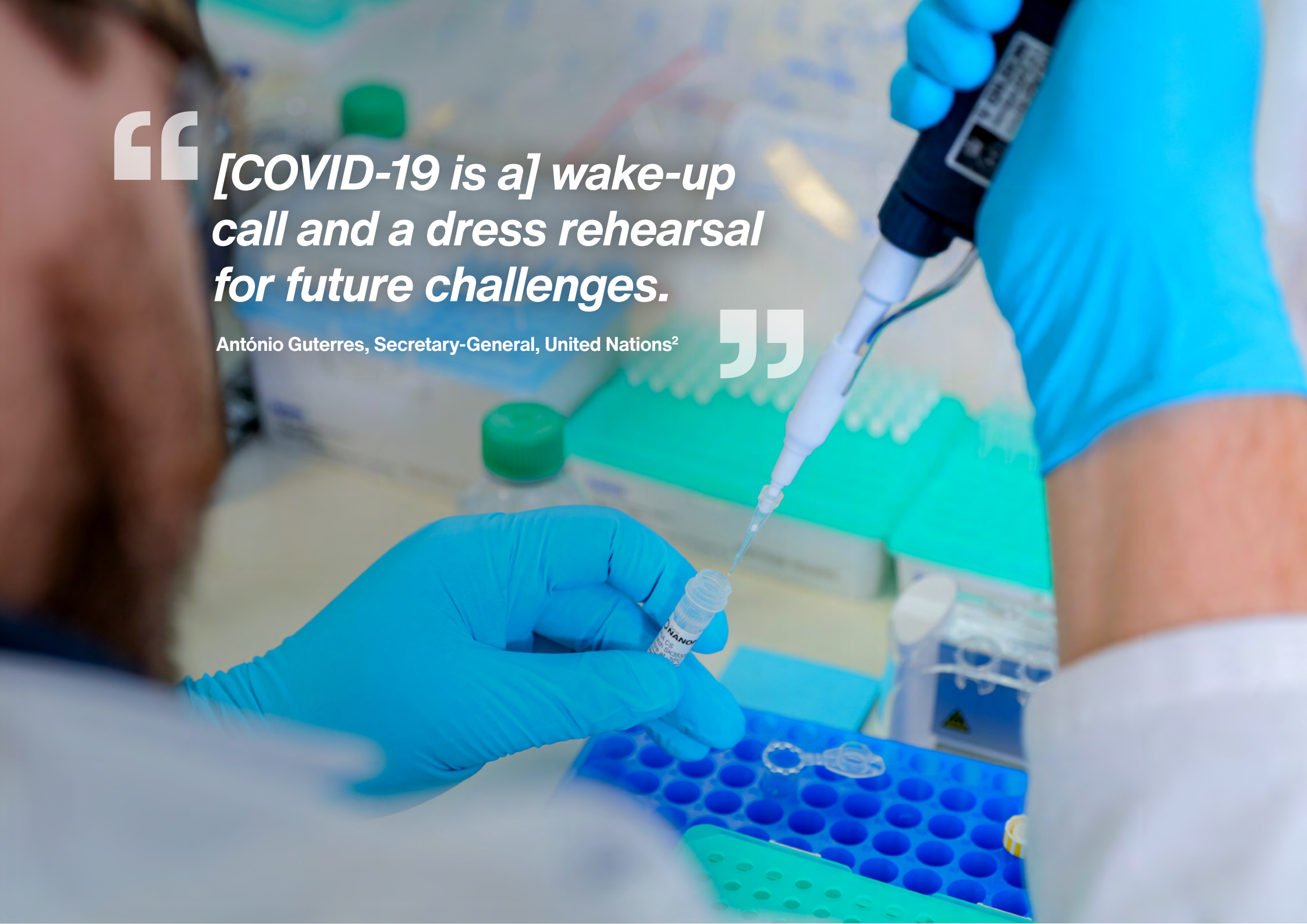
- The threat of infectious disease is ever present — caused by seasonal and novel viruses, bacteria, and fungi — and further compounded by growing antimicrobial resistance
- Significant outbreaks are an increasing risk, driven by globalisation, population growth, urbanisation, and climate change
- Preparedness is key to minimising the potentially devastating impacts of these outbreaks
- Genomic epidemiology, powered by **rapid, accessible sequencing** technology, provides vital information and time savings essential for rapid disease identification and control
- Investment in national and regional disease surveillance and monitoring infrastructure is critical

Global health security depends on the rapid recognition and containment of infectious diseases and no government can afford to be complacent about the risks posed to population health, economic, political, and social stability and wellbeing. It is possible to be prepared to prevent and control such threats by investing in intelligent and agile public health tools to monitor for potential risks, enabling responses at appropriate speed and scale to problems as they appear.

Genomic epidemiology is a crucial weapon in the public health fight against infectious diseases, providing rapid identification and complete characterisation of infectious disease pathogens. As evidenced by the COVID-19 pandemic, genomic epidemiology further supports precise tracking of pathogen evolution, providing detailed insights into sources of infection, routes of transmission, and the potential association of novel pathogen variants (such as the COVID-19 B1.1.7 and B1.351 variants originally identified in the UK and South Africa) with changes to disease severity, transmission, and diagnostic and therapeutic efficacy.

This briefing paper describes when, where, and how genomic epidemiology can offer critical and timely insights for infectious disease experts, public health professionals, and policy-makers to stay a step ahead of infectious disease threats, responding with maximal effect.

This briefing paper is published by Oxford Nanopore Technologies. Our low-cost, scalable genomic sequencing devices have been used extensively to rapidly characterise pathogens and track outbreaks, enabling informed and decisive action. *See page 32 for more information.*


[COVID-19 is a] wake-up call and a dress rehearsal for future challenges.
António Guterres, Secretary-General, United Nations[2]

The importance of managing infectious disease threats

KEY MESSAGES

- **Infectious diseases are an increasing threat**
- **Outbreaks can have devastating health and economic consequences**

Even a relatively small and well-controlled **outbreak** can exacerbate existing pressures on healthcare systems, whilst major outbreaks may overwhelm them. Serious **epidemics** also pose significant risks of wider social and economic destabilisation.

After an epidemic has subsided, there may be longer-term local, national, and global adverse effects on employment, health, education, trade, travel, and quality of life. According to the International Monetary Fund (IMF), the economic impact of the current COVID-19 pandemic could be in the region of US $28 trillion[1].

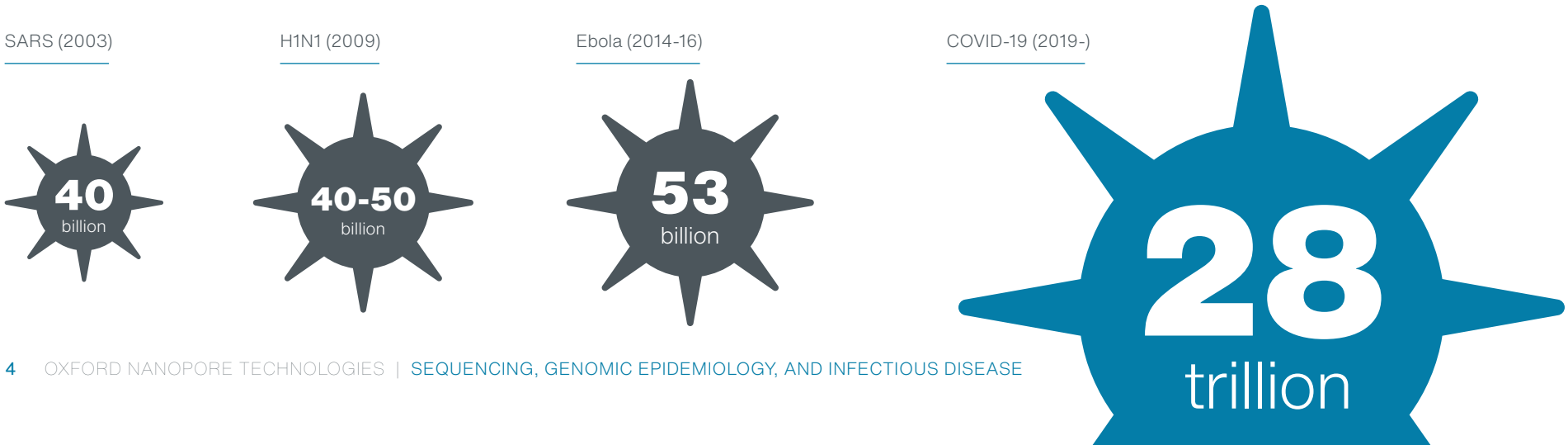
Significant new outbreaks are an ongoing risk, driven by increasing globalisation, population growth, urbanisation, and climate change. They may arise from the resurgence of historical infectious diseases and the spread of antibiotic resistance, or the emergence of infectious diseases caused by *novel* infectious agents (Figure 1). Novel pathogens are particularly dangerous, because of the time needed to understand the disease, and to develop control measures and treatments. Dr Michael Ryan, Executive Director of the World Health Organization (WHO) Emergencies Programme, recently described constant outbreaks of serious infectious diseases as *'a new normal'* for which countries need to be ready[4].

Although diseases passed between humans are the most urgent concern, infectious agents that cause disease in wild or domestic animals and plants can also have severe economic and environmental impacts, and indirectly or directly harm human health; for example, foodborne disease is a major burden in low and middle-income countries[5].

Preparedness is vital to minimise the damage infectious diseases can cause. Even small delays or mistakes in response can have significant implications for public health, and modest changes in outcome have much wider effects. For example, it has been suggested that the 2003 severe acute respiratory syndrome (SARS) outbreak would have tripled in size had there been a delay of just one week in applying control measures[6].

Examination of the relative success of New Zealand in control of the COVID-19 pandemic highlighted that the key elements were early, decisive reactions from health authorities, effective surveillance (including genomic surveillance systems), and targeted testing strategies[7].

Estimated economic impacts of recent infectious disease outbreaks.
Data from WHO[3]. (US dollars)



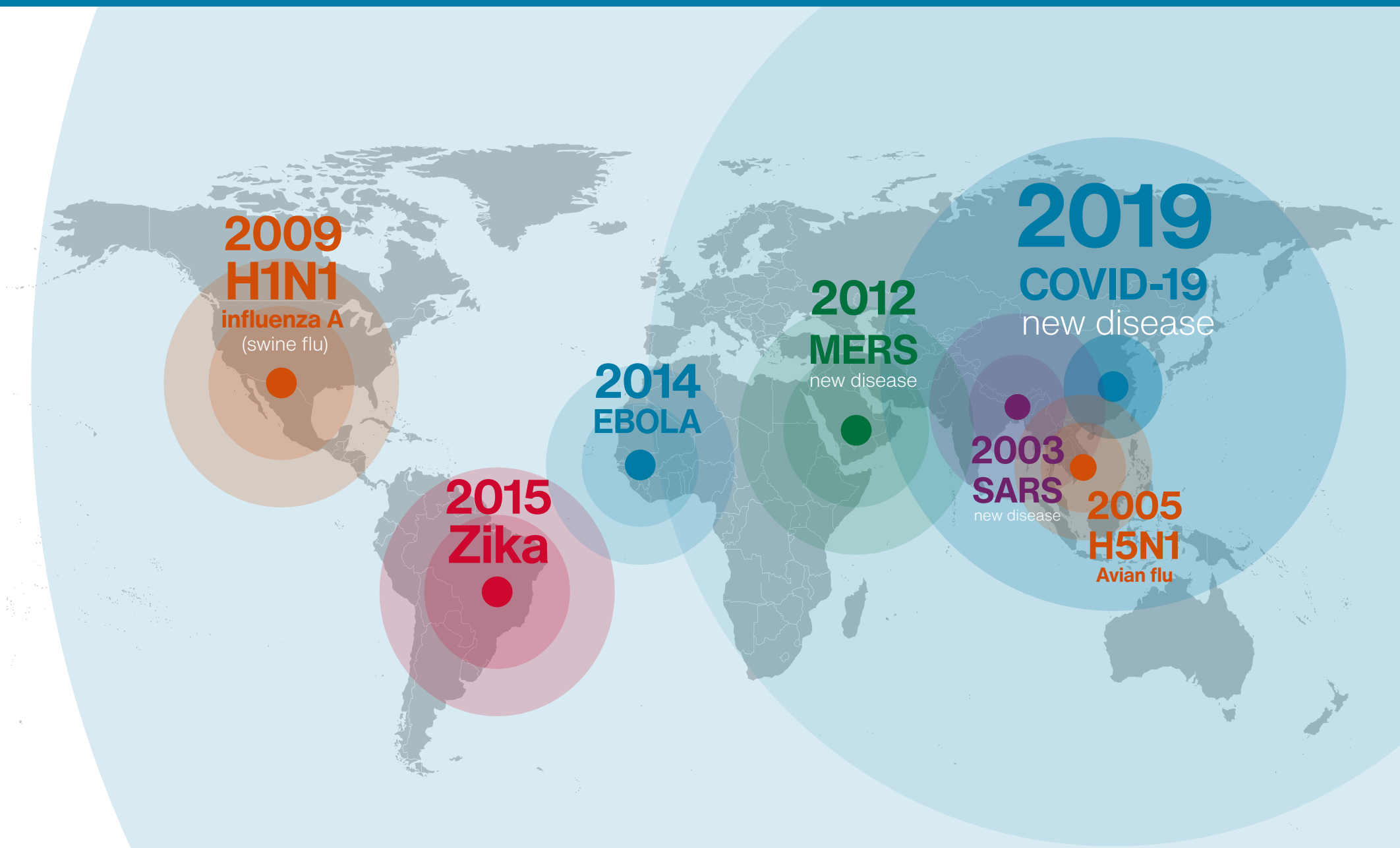


Figure 1: Selected infectious disease outbreaks in the 21st century
Infectious diseases are an increasing threat, for example in 2019 alone the WHO recorded over 100 outbreaks of 19 different infectious diseases, each posing a potential epidemic or pandemic threat[3].

Understanding infectious diseases

KEY MESSAGES

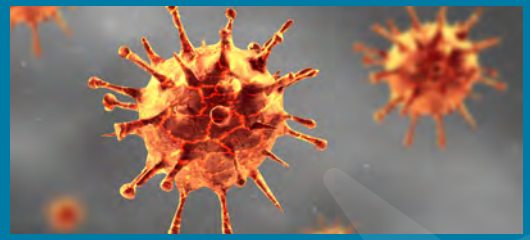
- **Effective treatments are not available for all pathogens and drug resistance is increasing**
- **Prevention of transmission is critical to managing infectious disease**
- **Pathogen genome sequencing is a key technique for identifying and understanding infectious organisms**

Infectious diseases are caused by **pathogenic** microorganisms **(microbes)**, including viruses, bacteria, fungi, and parasites. They can cause disease of varying type and severity, and many are easily spread through close contact between people. New human pathogens often come from microbes that normally infect animals, but have also acquired the ability to infect humans.

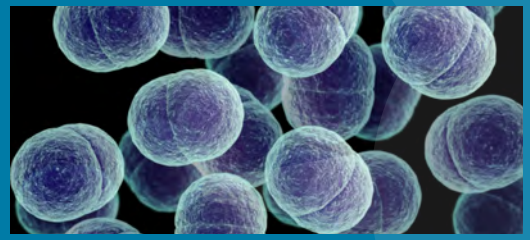
A wide range of anti-microbial drugs are currently available, though in recent years the spread of antimicrobial resistance, especially in bacteria, has reduced the effectiveness of some treatments. Detecting antimicrobial resistance in patients with bacterial infections is critical to ensuring that effective alternative treatments can be used[8] or for the identification of particularly dangerous cases where no current antibiotics will work, such as carbapenem-resistant infections[9]. Vaccination to prevent infection is possible for certain diseases, though even then some microbes such as influenza virus change so quickly that new vaccinations are needed every year; for many disease threats, no vaccines exist.

Given these limitations, control of all human infectious diseases relies on a combination of prevention wherever possible (including actions to reduce or block transmission between people) and rapid suppression of outbreaks where they do occur, alongside the use of effective treatment options, where available. To do this, insights into the biology and behaviour of the pathogens that cause disease are required, and as quickly as possible; DNA/RNA sequencing is an important tool for obtaining this information. Genomic epidemiology — powered by sequencing — is an essential element in the design of effective prevention and control measures.

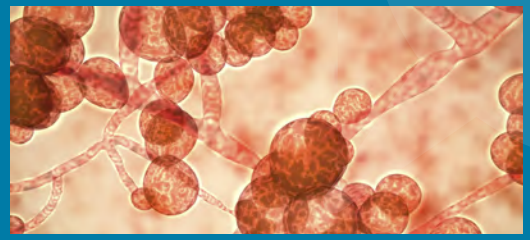
Types of pathogen and exemplar diseases



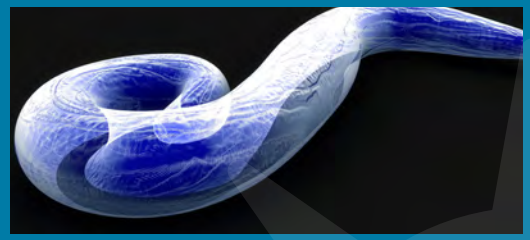
Viral diseases
Declared a global pandemic in March 2020, COVID-19, the respiratory disease caused by SARS-CoV-2, has claimed the lives of over 2 million people to date (February 2021).



Bacterial diseases
Responsible for over 10,000 deaths and health care costs of $2 bn per year in the US alone, methicillin-resistant Staphylococcus aureus (MRSA) is resistant to many first-line antibiotics.



Fungal diseases
First described in 2009 and commonly acquired in hospitals, Candida auris is associated with high mortality rate if detected late. Some strains are resistant to all three available classes of antifungals.



Parasitic diseases
Transmitted by infected mosquitoes, Plasmodium falciparum is the predominant cause of malaria, which results in over 400,000 deaths per year.

What is genomic epidemiology?

KEY MESSAGES

- **Genomic epidemiology provides the detailed insights required to prevent and manage infectious disease outbreaks**
- **Easily deployable genome sequencing technologies are integral to this approach, delivering rapid results in centralised or near-sample environments**

Epidemiology is the study of who develops disease, and why. Understanding how diseases occur and spread in different groups of people, including which people are at greater risk, helps plan strategies to prevent and manage them.

Genomics is the study of the genome — the full set of genetic information, providing coded instructions for living things to grow, function, and reproduce. Modern technologies make it possible to read (sequence) and understand (analyse) genomes.

Genome sequencing has already revealed a great deal about how genomes are involved in health and disease, and more is being discovered all the time.

The rapid development of genome sequencing technologies has enabled the use genomic epidemiology for infectious disease control — combining genomic data from pathogens with epidemiological investigation to understand how, where, and why an infectious disease emerges and spreads. This can inform decisions to deploy a variety of responsive control measures, for example quarantining cases and contacts, drug treatment strategies, enhanced hospital infection control procedures, strategic deployment of mosquito nets to relevant areas, curbs on travel or social contact, and so on.

> ***"All countries must develop a system for immediately sharing genome sequences of any new pathogen..."***
>
> **Global Preparedness Monitoring Board[3]**

Figure 2: In-field surveillance of Zika virus
The advent of highly portable sequencing devices has enabled low-cost disease surveillance and characterisation at point of infection, providing faster access to informative results. Device shown: MinION™ Mk1B from Oxford Nanopore Technologies. Image courtesy of Professor Nuno Faria, University of Oxford, UK.

How genomic epidemiology improves disease control

KEY MESSAGES

- **Genome sequencing enables pathogen identification, characterisation, and monitoring — vital for design of precise diagnostic tools and therapeutics**
- **Identification of pathogen source and transmission routes supports effective outbreak management**
- **Routine analysis of confirmed clinical cases allows identification of novel pathogen variants and tracking of associated pathogenic traits (e.g. transmissibility and virulence)**
- **Pathogen surveillance is key to early detection and successful disease control**

Traditional epidemiological techniques such as the generation of good clinical, environmental, and geographical data are vital for effective infectious disease management. The additional inclusion of genomic analysis provides a greater level of precision, resolving potential ambiguities in the data and supporting further actionable insights for disease control (Figure 3).

Pathogen identification

Genome sequencing provides a highly accurate means of precisely identifying and characterising the organism(s) causing infectious disease, which is especially important in outbreak situations, and most of all for new pathogens.

> ***[Whole-genome sequencing] will improve the accuracy and effectiveness of disease surveillance, outbreak investigation and evaluation of prevention policies by enhanced assessment of disease and drug resistance transmission dynamics.***
>
> **European Centres for Disease Control[10]**

For example, scientists examining the sudden outbreak of pneumonia in Wuhan, China, in late 2019 applied technique known as **metagenomic sequencing** — where nucleic acids from all organisms in a sample are analysed — to quickly identify the microbe behind the disease as a new coronavirus, SARS-CoV-2[11].

The technique of metagenomic sequencing, which requires no prior knowledge of the infectious organism, also allows the identification of co-infections, where multiple different pathogens may be present in a single sample. For example, deploying real-time metagenomic sequencing to COVID-19 patients in intensive care units has been shown to enable the identification of secondary infections, antimicrobial resistance, and hospital-associated transmission[12].

Once the pathogen responsible for a disease has been identified, it is possible to design sequencing assays that specifically detect and amplify only the genome of the pathogen of interest. Such targeted assays often support faster time to results, lower sequencing and analysis costs, and can be utilised for samples with lower levels of pathogen.

Understanding transmission

Rapidly replicating pathogens tend to accumulate mutations in their genetic code. By comparing pathogen genomes obtained through genome sequencing (**phylogenetic analysis**), these genetic differences can be used to identify how closely related each pathogen is to another (Figure 4). Combined with traditional epidemiological data (e.g. time, location, exposure), this enables rapid and highly accurate tracking of the sources and transmission routes of the disease.

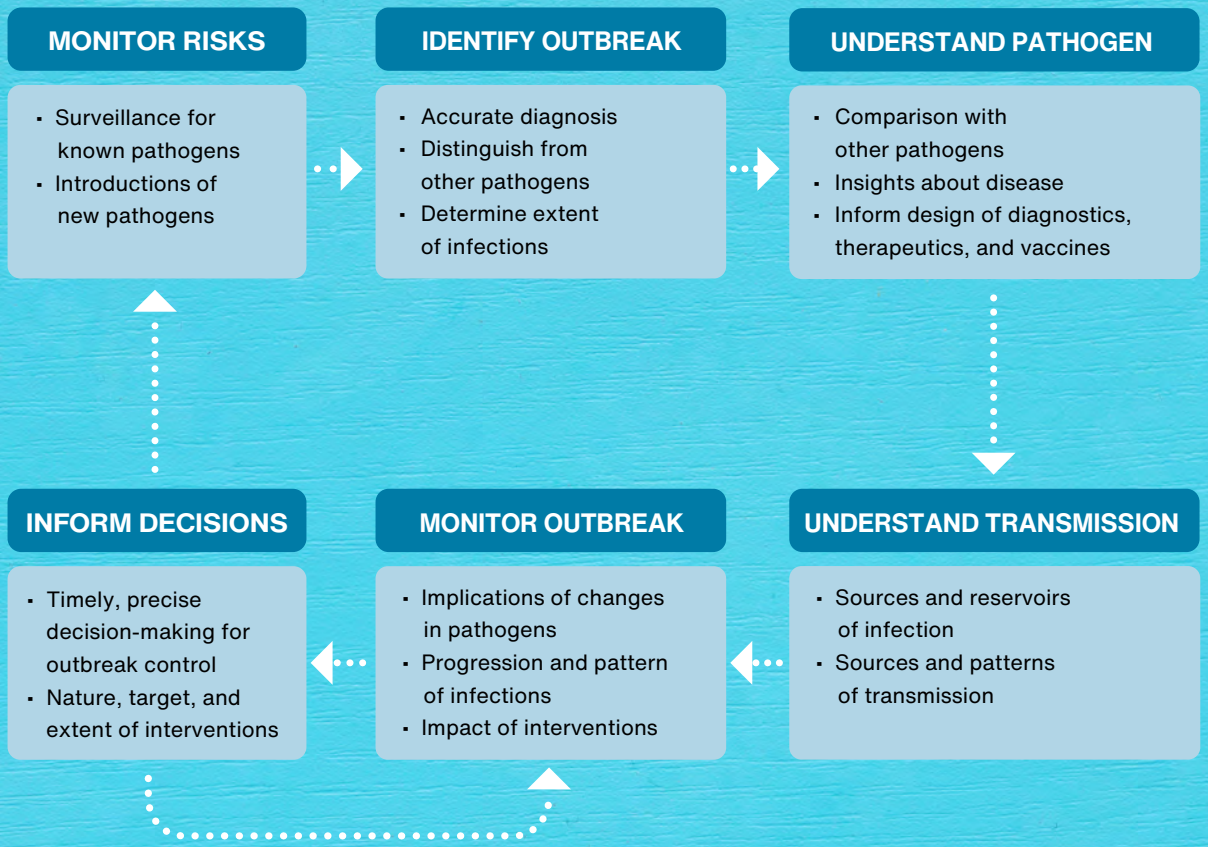
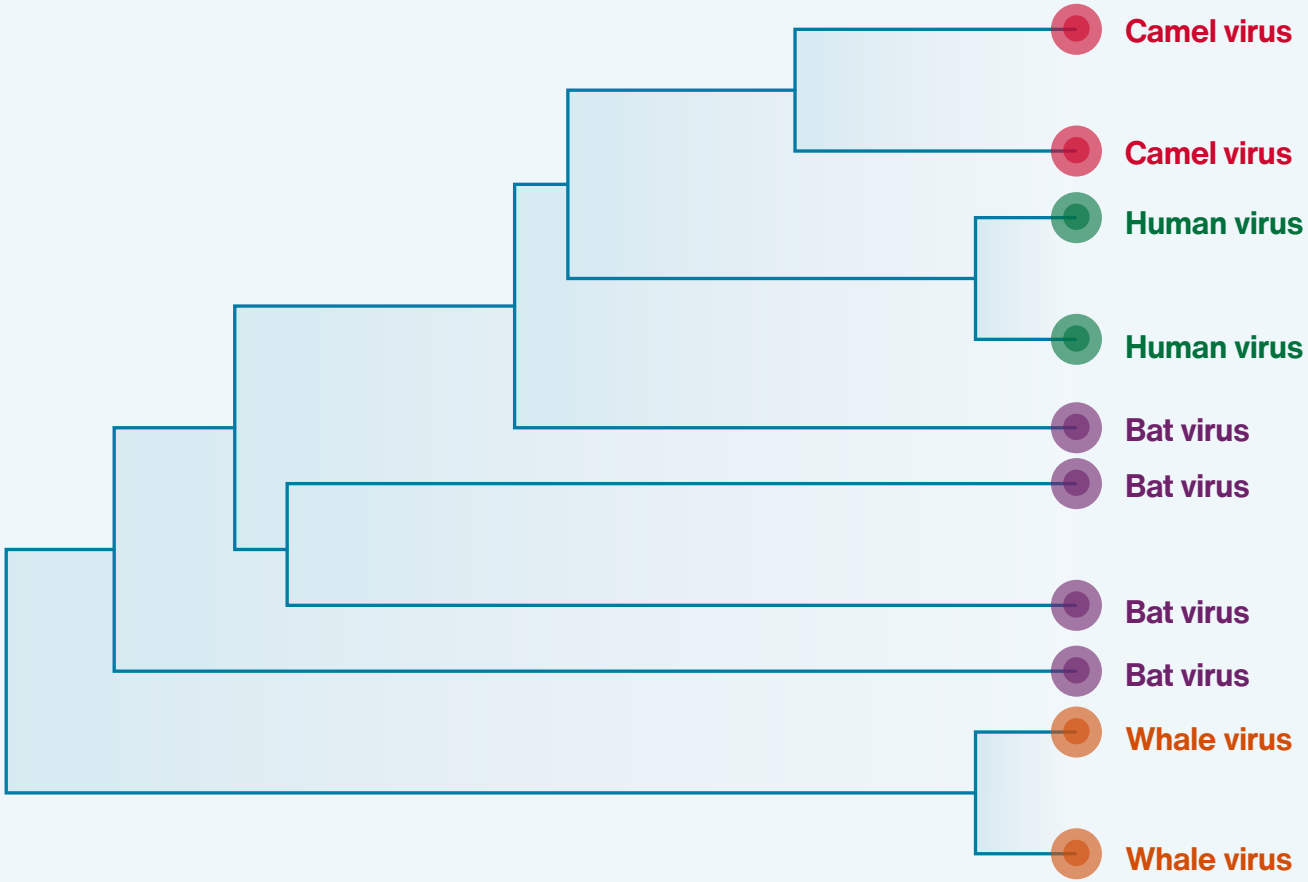


Figure 3:
How genomic data informs infectious disease control.

Figure 4: Phylogenetic trees show the evolutionary relationship between organisms or strains of the same organism
Hypothetical phylogeny showing inferred relationship of viruses isolated from different host organisms. The longer the horizontal line, the larger the amount of genetic change. Figure adapted from ARTIC Network[13].



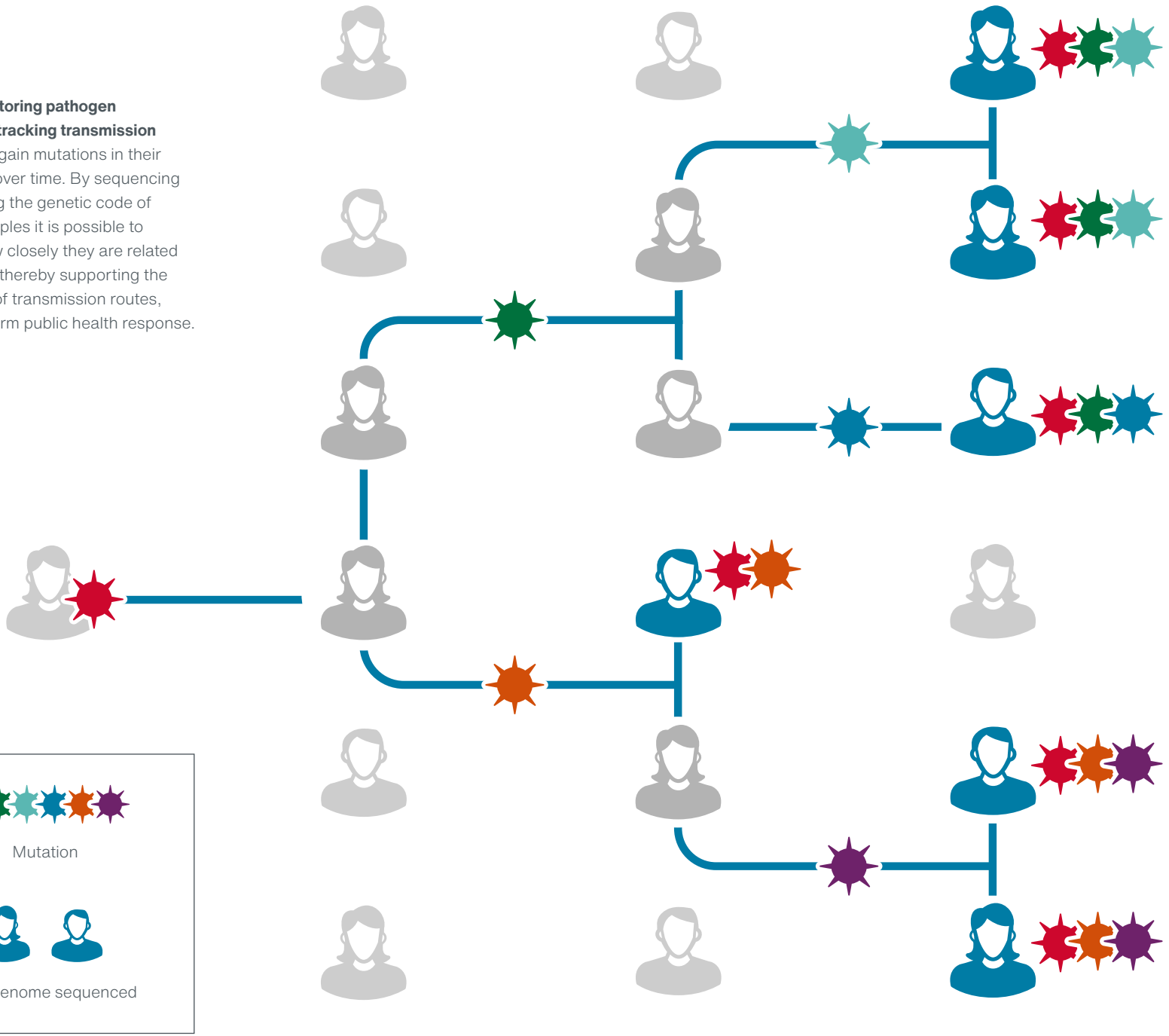
Identifying outbreak origin

By comparing the genome sequence of the novel coronavirus, now called SARS-CoV-2, the cause of the COVID-19 disease, against databases of other pathogen genomes, the virus was shown to be related to the SARS-CoV-1 virus that originally emerged in Guangdong, China, and caused the 2003 SARS outbreak. The genome sequence of subsequently termed SARS-CoV-2 virus also revealed a high degree of similarity to certain bat coronaviruses, suggesting a potential origin for the new disease. By identifying outbreak origin, containment measures can be put in place to prevent further spread.

Patterns of transmission

Routinely comparing the genetic code of pathogens from infectious disease samples during clinical outbreaks can provide significant insights into sources of infection and routes of transmission. Thanks to new ultra-rapid sequencing technology, these analyses are now possible in near real-time (as opposed to retrospectively over weeks and months) and at scale, and can thereby actively inform ongoing outbreak management (Figure 5).

Figure 5: Monitoring pathogen evolution and tracking transmission
All organisms gain mutations in their genetic code over time. By sequencing and comparing the genetic code of pathogen samples it is possible to determine how closely they are related to each other, thereby supporting the identification of transmission routes, which can inform public health response.



Key:

These advances allow infectious disease experts to monitor the outbreak and determine whether different strains of the pathogen are emerging over time, what their implications may be for the likely progression of the outbreak, and what changes in prevention or treatment measures may be needed. For example, during the COVID-19 pandemic, genomic epidemiology has been used to differentiate international introductions from domestic spread of the virus, and provide evidence to support a range of public health interventions — from limiting international travel and implementing robust healthcare infection control procedures (see Case study 1) to closing shops, restaurants, and schools, on both a national and regional level (see page 18).

Genomic surveillance can also uncover 'cryptic' or hidden transmission pathways in instances where traditional epidemiological observations may not be suggestive of an outbreak. This is particularly useful for diseases with mild or non-specific clinical symptoms. For example, genomic epidemiological analysis of Zika virus samples during the 2015–16 outbreak in North and South America not only traced its origin to Brazil, it also revealed that the virus was circulating there for at least a year before the first reported cases[14]. Had routine genomic epidemiological surveillance been in place, the outbreak could have been detected much sooner and steps taken to prevent spread to other countries, thereby vastly reducing the devastating number of affected births.

> ***Member States should urgently increase [SARS-CoV-2] genome sequencing to at least 5% and preferably 10% of positive test results ... to identify the progression of the variants or detect any new ones.***
>
> **European Commission[16]**

Pathogen characterisation

In addition to enabling precise pathogen identification, genome sequencing data can be used to infer significant information about the biology of the pathogen. This can be achieved through existing knowledge on organisms with similar genetic composition and through lab experiments where different genomic variants — either naturally occurring or experimentally introduced are linked to the ability of the organism to cause disease.

Understanding the genetic code of a given pathogen allows researchers to develop molecular assays that provide a rapid, accurate, and cost-effective means of diagnosing the pathogen in clinical samples. Such assays can be applied at a scale and speed far above that of whole-genome sequencing; however, for optimal outbreak response, once identified, positive samples should still undergo whole-genome analysis. Design and research for novel vaccines and therapeutics is also supported through whole-genome analysis, and again speed may be critical — for example, in developing novel vaccines such as those for SARS-CoV-2, or in modifying existing vaccines to be effective against dominant new strains of a pathogen (see Case study 4).

Outbreak monitoring

Monitoring genomic evolution of a pathogen during an outbreak is of vital importance. Novel variants can alter the ability of a pathogen to cause disease, as has been evidenced by the increased transmissibility of the SARS-CoV-2 variants B.1.17, B.1.351, and P1, first detected in the UK, South Africa, and Brazil respectively. They may also reduce the effectiveness of diagnostic assays, compromising disease control initiatives. Critically, they may reduce the efficacy of disease therapeutics such as vaccines, so called **vaccine escape**.

Outbreak surveillance

New outbreaks caused by previously unknown or substantially changed pathogens can cause considerably more ill health and social and economic damage, due to a lack of immunity in human populations. For example, the constant gradual genetic changes in influenza viruses (combined with seasonal weather) drive regular epidemics, but occasional big shifts (such as when an animal influenza virus gains the ability to infect humans) can cause pandemics. A pathogen that is easily passed between people and to which they have little or no natural resistance has potentially devastating capacity to spread around the world.

The faster information can be learnt about the pathogen, the faster appropriate steps to manage the outbreak can be taken. This is why effective surveillance to monitor potential new disease threats is so important.

That Zika was circulating silently across most of the Americas highlights the need for a better surveillance system for Zika virus across the globe. There is also a need to extend genomic surveillance to other pathogens.

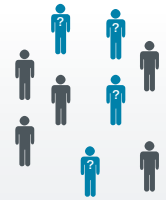
Professor Nuno Rodrigues Faria, University of Oxford

Precise, adaptable, rapid management of a novel virus with genomic epidemiology (COVID-19 example)

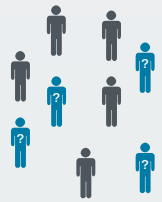
DETECTION OF NEW CASES

New COVID-19 cases are confirmed via clinical diagnostic test, revealing new clusters.

SCENARIO A



SCENARIO B



Where were these individuals infected?

Do these represent a single cluster or many?

How many more people are at risk of infection?

How can further cases be prevented effectively?

SEQUENCING

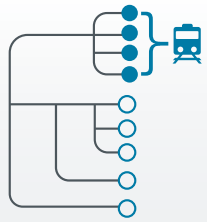
A sample is taken from every new case and the SARS-CoV-2 viral RNA extracted.

The whole viral genome is sequenced for each sample.

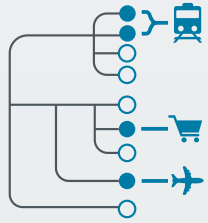
ANALYSIS

Each viral genome is analysed alongside publicly available genome sequences, to determine which other cases they are most closely related to.

SCENARIO A



SCENARIO B



Identification of small differences in the viral genomes, combined with data from individuals on their movements, reveals potential sources of transmission.

TARGETED ACTION

Clusters of transmission and individuals who may have been exposed and identified.

With this information, precise public health measures can now be implemented to limit further spread — without disruption to unaffected areas.

SCENARIO A

Cases are revealed to belong to the same cluster, originating from one location.

SCENARIO B

Cases are revealed to result from multiple, separate transmission events, with the potential for more widespread exposure.

Genomic epidemiology allows assessment of further risk of transmission and suitable prevention action.

SUPPORTING PUBLIC HEALTH EFFORTS

In the event that a new variant or strain is identified, this information is rapidly fed back into testing and vaccine development.

ACCELERATING COVID-19 RESEARCH

The new data is shared across the scientific community, enabling further outbreak tracking and providing crucial information for research.

Implementing genome sequencing for outbreak control

KEY MESSAGES

- **Global sharing of genomic epidemiology data is vital for disease management**
- **Centralised laboratory facilities for genome sequencing offer large-scale, high-throughput capacity**
- **Localised, near-sample deployment of genome sequencing offers additional capacity for rapid and responsive analysis at point of need**
- **Low-cost, portable sequencing technology enables genomic epidemiology in remote and resource-limited locations**

As we have seen, genomic epidemiology offers a critical tool to support rapid and effective assessment of infectious disease threats and targeted, cost-effective control of outbreaks.

This is a vital capability for local and national governments and health systems around the world. The ability to share genomic epidemiological and related data nationally and internationally is highly valuable, to enable the design and roll-out of rapid and effective at-scale testing and inform interventions against local, epidemic, and pandemic outbreaks.

To achieve this, a balance between centralised and localised sequencing capacity may be needed. Central coordination of regional or national responses is obviously desirable and may include larger laboratories with trained staff able to routinely receive and process large numbers of samples, and to analyse, store, and share large-scale data. However, the availability of such facilities will naturally vary between regions and countries, and for a range of practical reasons, such as sample transport logistics, availability of skilled staff, capital investment resources, and lines of communication with public health outposts.

Distributed sequencing centres that operate at or near the site of sample collection are an important addition to genomic surveillance systems; they can significantly increase the speed with which results can be returned to inform local or remote data analysis and genomic epidemiological surveillance, thus enabling swifter and more precisely targeted local interventions to control outbreaks. In more remote locations, it is also possible to use very small mobile sequencing units, such as the 'suitcase laboratory' used to support the 2015 West African Ebola[18] and 2016 South American Zika outbreaks[19], and rabies elimination initiatives[20,21].

Establishing and maintaining local capacity for rapid deployment of pathogen genomic sequencing can provide a critical time advantage in containing outbreaks. The recent development of low-cost, portable sequencing technology along with streamlined, real-time systems for data interpretation now makes this feasible.

Building a strong and resilient global sequencing network can maximize the public health impact of sequencing, not only for SARS-CoV-2 but also for future emerging pathogens.

World Health Organization
Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health[17]

Studying the virus's genome helps to highlight cryptic or hidden transmission. That's the real power of it — you can detect outbreaks and act while they're happening.

Dr. Estee Török, University of Cambridge / Addenbrookes Hospital[24]

Credit: Anastasia Taylor Lind

CASE STUDY

Local control of COVID-19 community and hospital outbreaks

UK, 2020

As a new pandemic took hold across the UK, hospital admissions of patients with COVID-19 surged, including at the Cambridge University Hospital. It was very important to understand where and how these infections were spreading. The recently established COVID-19 Genomics Consortium UK[22] swung into action; a world-leading distributed network of hospitals, public health agencies, and academic partners, with a mission to provide rapid, high-throughput whole-genome sequencing of SARS-CoV-2 samples.

The local Public Health England clinical microbiology laboratory and the University of Cambridge Department of Pathology worked together to sequence pathogen genomes from healthcare-associated COVID-19 infections within 24 hours of receiving samples from patients. They analysed combined genomic and epidemiological data, revealing a level of detail not detectable from the epidemiological data alone. In six weeks, they found over thirty different clusters of related cases from 1,000 viral genomes — including small outbreaks that would otherwise have been hidden amid the bigger picture of infections. This made it possible to identify distinct outbreaks linked to specific locations in hospital and community settings, information that was immediately fed back to doctors, managers, and infection control experts, allowing them to act quickly to prevent further infections from the sources identified[23].

For example, results revealed a cluster of COVID-19 patients in different parts of the hospital with near-identical sub-types of the SARS-CoV-2 virus, strongly suggestive of a shared origin for these cases. It was found that these patients all had kidney disease and had visited the hospital outpatient dialysis clinic on the same day, most travelling in the same patient transport vehicle. Closing the dialysis clinic was not an option, so instead both clinic and vehicle were intensively cleaned, and new personal protective equipment and social distancing measures swiftly put in place. There were no further cases amongst the dialysis patients.

> ***We're able to combine genomic data with patients' medical records to provide real time information to help the hospital review its infection control on a weekly basis. It's also highlighted possible transmission networks less well documented, such as care homes, outpatient units and ambulance services.***
>
> **Professor Ian Goodfellow, University of Cambridge[25]**

Examination of the viral genome sequences helped to reveal outbreaks and sources that would otherwise have been lost to view.

The genomic data also supported wider efforts, feeding into a central database and enabling a national genomic epidemiology overview of the spread of SARS-CoV-2 across the UK — for example, determining that there had been over 1,300 separate early introductions of the virus to the country, mostly from other countries in Europe such as Spain and Italy, rather than from China where the pandemic began.

Real-time sequencing of viruses can reveal crucial details that contribute to the understanding and the control of infectious disease outbreaks. It is now possible to resolve chains of transmission to a level of detail otherwise unachievable using traditional methods.

Nigerian Government press release[27]

CASE STUDY

Tracking transmission in a Lassa fever outbreak

Nigeria, 2018

In 2018, a 10-fold rise in the number of Lassa fever cases in Nigeria raised concern that the virus causing the disease had evolved to allow easier transmission between humans — potentially the beginning of a major outbreak. It was important to understand as quickly as possible whether or not this was the case.

The Nigerian Centre of Disease Control made an urgent call for genomic sequence data on Lassa fever virus strains involved in the outbreak, which was in a geographically remote setting with poor transport accessibility. This would normally have meant that samples would need to be shipped over a long distance for sequencing in the national reference laboratory, an extremely slow process. Instead, an international team of scientists used a portable nanopore-based sequencing device to provide real-time genomic analysis of 36 viral samples, followed by a further 85 samples from other cases to confirm their findings.

The genomic analysis results were communicated immediately to the Nigerian government and to the World Health Organization. Combined with conventional epidemiological investigations, they showed that there were no signs of a new and more dangerous strain of the virus emerging. Instead, a wide variety of strains were involved, suggesting that cases were mostly caused by independent transmission events of the virus to humans from rats, allaying fears that a major outbreak driven by human-to-human transmission was imminent[26].

This knowledge also meant that limited public health resources could be focused most effectively on community engagement about rodent control, environmental sanitation, and safe food storage.

A great deal [of] vigilance and monitoring of animal populations will be needed to understand genetic mutations and implications this could have for human vaccines.

Professor Marion Koopmans, Department of Viroscience, Erasmus MC, Netherlands[33]

Dr. Bas Oude Munnink, Erasmus MC, Netherlands, using the portable MinION sequencer from Oxford Nanopore Technologies.

CASE STUDY

Evidence-based public health decision-making in the COVID-19 pandemic

The Netherlands, 2020

The World Health Organization received notification of a new and infectious respiratory disease in Wuhan, China at the end of December 2019, and issued the first outbreak report about the 'pneumonia of unknown cause' in January 2020[28]. These were subsequently identified as COVID-19 cases arising from infection with the novel SARS-CoV-2 virus. With the pandemic potential of the new disease quickly becoming clear, countries around the world moved to plan their own control measures against this new threat to public health.

Drawing upon world-leading expertise in virology, exemplified by the Department of Viroscience at Erasmus MC in Rotterdam, the Netherlands, opted to use genomic epidemiology from the onset — introducing rapid sequencing and analysis of all patients with suspected diseases. The first cases were identified in late February 2020 as introductions of the virus from Italy. Cases increased rapidly among non-travellers, including healthcare workers; it was not clear how big a problem disease transmission in healthcare settings was. This was already a sensitive political issue, given an expected shortage of personal protective equipment to reduce risks to health professionals.

Rapid sequencing continued for all suspected cases in travellers and healthcare workers, plus local surveillance, and by mid-March, the genomes of 189 SARS-CoV-2 viruses from the Netherlands had been sequenced — over a quarter of the sequences available worldwide at that point. The data revealed that multiple virus sub-types were circulating in the population; ongoing local transmission was creating the local and regional clusters of disease, but there was little infection of healthcare workers taking place in medical settings. This meant that community transmission was the main mode of spread within the Netherlands at this point; there were also indications that mass gatherings such as festivals could be acting as super-spreading events.

> ***The combination of real-time whole genome sequencing with the data from the National Public Health response team has provided information that helped decide on the next steps in the decision-making.***
>
> **Oude Munnink *et al*. Nature Medicine[29]**

This information prompted an immediate change in approach to targeted public health interventions to control the outbreaks, including the introduction of restrictions on movement and the closure of schools, catering, and sports clubs, as well as physical distancing measures at regional (and subsequently national) levels[30].

The Netherlands continued to use genomic epidemiological surveillance and were able to spot transmission of SARS-CoV-2 variants between humans and minks on fur farms as early as April 2020, warning correctly that this could pose a further public health risk[31]; in November 2020, Danish authorities enacted a cull of all mink after a new viral variant transmissible to humans and with the potential to reduce the effectiveness of new vaccines emerged[32].

Had this virus caused a severe outbreak or pandemic, our proactive surveillance efforts and vaccine derivation would have provided an approximate 8-week time advantage for vaccine manufacturing.

Rambo-Martin and Keller *et al.* mSphere[34]

CASE STUDY

Speeding up delivery of an effective vaccine for a new swine flu

USA, 2018

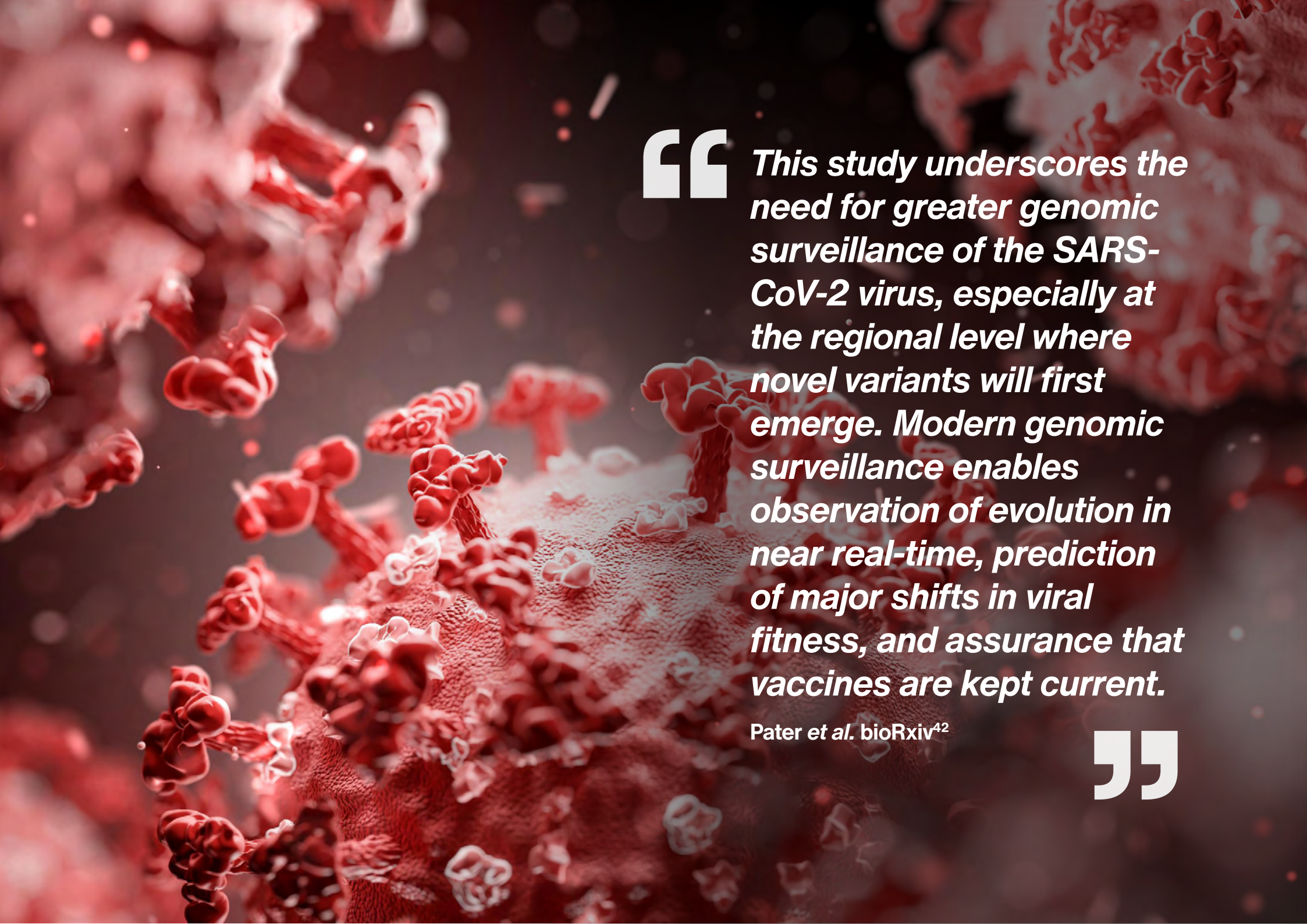
Influenza A viruses infect a wide range of animals, and can sometimes move from these animal hosts to infect humans, with potentially serious effects; the 2009 H1N1 flu pandemic arose from a swine-human virus transmission event. Since then, there have been hundreds of known human cases of swine-origin influenza in the United States, most commonly from exposure at agricultural fairs.

Part of routine work to limit the risks of a future flu pandemic is genomic surveillance of circulating viruses in swine and other animals. Genomic data can be used to develop new candidate vaccine viruses to form the basis for a new vaccine *before* an outbreak occurs in humans; however, collecting data typically takes weeks or even months — precious response time lost, should an unexpected new strain with pandemic potential emerge.

In 2018, a mobile genome sequencing station was set up at a large US swine exhibition. Scientists analysed samples from animals that had tested positive for influenza A virus (and neighbouring pigs) with a commercially available rapid testing kit. Using the rapid, portable genome sequencing platform and an automated, real-time data analysis pipeline — requiring neither a laboratory facility nor an internet connection — they were able to assess the viral genome sequences and immediately compare them with known swine flu viruses, including current candidate vaccine viruses.

Three different swine influenza A viruses were identified. The most common, an H1N2 viral subtype, was found to be genetically distinct to the most similar World Health Organization candidate vaccine virus — including differences in parts of the genome known to determine vaccine efficacy. That meant that if this virus subtype did start to infect humans, there was no rapid path to producing an effective vaccine against it.

As an exercise in pandemic preparedness, all genomic sequence data were sent to the US Centers for Disease Control and Prevention (CDC), who used it to develop a synthetic candidate vaccine virus. Importantly, the actionable genomic analysis results were obtained within 18 hours of setting up the mobile testing facility[34]. This unprecedented speed could offer a critical time advantage in mobilising design and production of a new vaccine to protect against a new influenza virus threat.



This study underscores the need for greater genomic surveillance of the SARS-CoV-2 virus, especially at the regional level where novel variants will first emerge. Modern genomic surveillance enables observation of evolution in near real-time, prediction of major shifts in viral fitness, and assurance that vaccines are kept current.

Pater *et al.* bioRxiv[42]

CASE STUDY

Identification of variants of concern

UK, 2020-21

The recent identification of a number of SARS-CoV-2 variants of concern (VOC) has demonstrated the importance of ongoing genomic surveillance to monitor changes in the viral genome sequence and their impact on virus biology. In the UK, the world-leading genomic surveillance infrastructure established by the COG-UK consortium, funded by the UK government and the Wellcome Trust, had sequenced over 200,000 SARS-CoV-2 genomes by late January 2021, around half of the global total at that point [35,36].

This infrastructure enabled COG-UK to monitor different mutations emerging in the viral genome as the virus circulated in the population, including a new lineage of SARS-CoV-2 called B.1.1.7 (VOC202012/01), often referred to as the 'UK variant'[37]. B.1.1.7 has a number of changes, particularly in the Spike protein, and increased in frequency in late 2020, even with lockdown measures in place, compared to other variants. These data showed that the genetic changes in B.1.1.7 had made the virus more transmissible[38]. Changes to the Spike protein are of particular interest and concern because it is this protein that enables the virus to enter human cells; it is also the target for many current vaccines and diagnostics[39].

The identification of B.1.1.7 alerted the global community to the emergence of a more transmissible variant. Subsequent international sequencing efforts identified other lineages — B1.351 and P1, first identified in South Africa[40] and Brazil/Japan[41] respectively – also associated with increased transmissibility. Further variants are also being identified in the United States[42].

The identification of these variants led to widespread global public health action, including renewed national lockdowns and a decrease in or total ban on travel between specific areas or countries. It also precipitated intensive study of these variants in the laboratory, to better understand any potential impact on the effectiveness of vaccines and diagnostics developed from genome sequences produced early in the pandemic.

Future-proofing for infectious disease control

Making the most of globally competitive technologies for genomic sequencing is the key to global, national, and local preparedness to deal with new and evolving disease threats. Investing in these genomic tools — including those that can be deployed quickly and flexibly where need arises and without major infrastructure or training requirements — is an essential defence against the very considerable health, economic, and social risks of infectious disease outbreaks.

In many countries, this distributed, real-time genomic analysis approach is already enabling management of the current COVID-19 pandemic, and it will undoubtedly also ensure more timely intervention and positive outcomes for future outbreaks of all kinds.

Learn more about the application of portable, real-time sequencing to the genomic epidemiology of COVID-19 at www.nanoporetech.com/covid-19.

The accelerated integration of genome sequencing into the practices of the global health community is a must if we want to be better prepared for the future threats.

Sylvie Briand, Director World Health Emergencies Programme, World Health Organization[17]

About Oxford Nanopore Technologies

Oxford Nanopore has developed ground-breaking genome sequencing technology that has been utilised extensively across the world to deliver fast and comprehensive insights into infectious disease outbreaks, including COVID-19, Ebola, Zika, and many more.

Benefits of nanopore sequencing technology include:

- Real-time data streaming — faster access to actionable results
- Portable pocket-sized devices — ideal for decentralised, near-sample analyses
- High-throughput benchtop platforms — enabling high-volume sample testing
- Streamlined, automatable sample preparation — reliable results with minimal hands-on time
- Low cost — starting at just $1,000 and capital-free purchase plans

In addition to informing genomic epidemiology, nanopore technology has been used to develop and run precise diagnostic assays, as evidenced by the recent LamPORE™ COVID-19 test*, which offers low-cost, scalable access to highly accurate diagnosis of SARS-CoV-2.

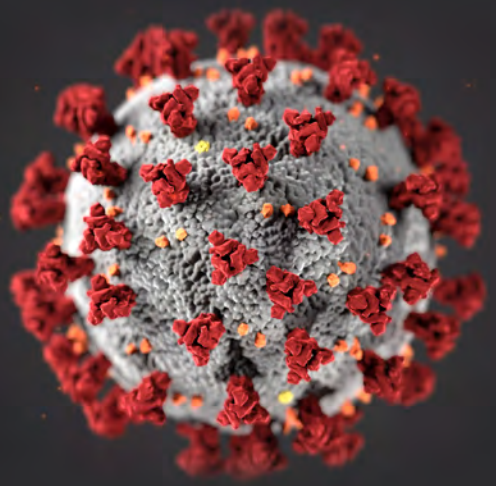
* LamPORE™ COVID-19 is CE marked for *in vitro* diagnostic use. Please check with your local sales representative for availability in specific markets.

Figure 6:
A range of nanopore sequencing devices are available — from portable Flongle™ and MinION devices, to compact, high-throughput, GridION™ and PromethION™ devices. VolTRAX™ offers streamlined and portable automated sample preparation.

Sequencing for COVID-19

Throughout the COVID-19 pandemic, researchers in the public health and scientific community have demonstrated the utility of rapid, nanopore sequencing to enable quick responses to evolving challenges. Nanopore sequencing delivers:

- **Full virus genome sequencing in <8 hours**
- **Scalable analysis of 1–480 samples on a single device**
- **Detailed protocols and comprehensive support**
- **Data analysis tools and resources**



Discover more at www.nanoporetech.com/covid-19

Related resources

(Clickable links)

Video: How nanopore sequencing works

Getting started guide: metagenomic sequencing

White papers: Clinical research; TB; Microbiology

Nanopore sequencing for COVID-19

COVID-19 infographics

Customer community COVID-19 timeline

LamPORE COVID-19 assay*

Visit www.nanoporetech.com

Figure 7:
The versatile GridION device from Oxford Nanopore provides highly scalable, real-time sequencing capability.

Glossary

epidemic: widespread occurrence of an infectious disease in a regional or national community at a specific time above normal or expected levels

epidemiology: study of how often diseases occur in different groups of people and why, typically examining the distribution of disease in specific populations and the associated causes or risk factors

DNA: deoxyribonucleic acid, the hereditary material in humans and almost all other organisms that contains genetic information and passes it between generations. DNA contains the instructions needed for an organism to develop, survive, and reproduce

DNA / RNA sequencing: process to determine the precise sequence (order) of the four chemical 'bases' that make up the DNA or RNA molecule, in order to understand what it does

genome: complete set of all the genetic material of an organism, usually DNA (some viruses have RNA genomes)

genomics: study of the genome

genomic epidemiology: study of how variations in genomes of pathogens or their hosts influence health and disease, including how common specific variations are, how they interact with environmental factors, and how they contribute to disease risk

genome sequencing: determining the genetic sequence of the complete genome

lineage: sequence of species each of which is considered to have evolved from its predecessor; in the case of viruses, may refer to subtly different versions of a given viral species that have evolved over time from an ancestor virus

microbe: microscopic organism, for example viruses and bacteria

mutation: alteration in the genome of an organism; change in DNA sequence

outbreak: sudden occurrence of disease cases in excess of normal or expected levels, usually (but not necessarily) caused by an infectious agent

pandemic: epidemic occurring worldwide, or over a very wide area, crossing international boundaries and usually affecting a large number of people; usually taken to exclude expected seasonal epidemics

pathogen: microbe or microorganism that can cause disease

phylogenetic analysis: examination of the evolutionary development of an organism or a particular characteristic of an organism

RNA: ribonucleic acid, a molecule that transfers information from the genome into proteins. Some viruses (including coronaviruses) use RNA, as opposed to DNA, as their genetic material

real-time sequencing: scientific technique that reads and reports DNA or RNA sequences simultaneously, as opposed to older methods that first capture and then report the sequence information in bulk after the sequencing run

vaccine escape: when the circulating strain of an infectious disease mutates (changes) such that the protective immune response elicited by vaccination is ineffective and does not prevent infection and disease; see also viral escape

variant: specific region of the genome that differs between two genomes; sometimes used to refer to a common point of variation where mutation is used to refer to a rare change; sometimes used to refer to a different viral strain with a changed genome

variant of concern (VOC): a pathogen strain with a genomic change or changes that may increase its ability to infect people, to be transmitted between people, or to cause more serious disease in people. Variants that are more easily spread between people in an infectious disease outbreak have an advantage over less easily spread versions of the pathogen

viral escape: when the circulating strain of an infectious disease mutates (changes) such that existing, protective immune responses become ineffective and do not prevent disease; see also vaccine escape

strain: genetic variant or subtype within a species, particularly important in microbes that reproduce and develop genetic changes very rapidly; a new strain is usually considered to have emerged if the microbe has genetic variants that affect how easily it can infect people, spread between people, or cause more or less serious disease in infected people

References

[1] International Monetary Fund (IMF). World economic outlook, October 2020: a long and difficult ascent. 2020. Available at: https://www.imf.org/en/Publications/WEO/Issues/2020/09/30/world-economic-outlook-october-2020 [Accessed 01 February 2020]

[2] General assembly of the United Nations, Plenary. Seventy-fifth session, 4th and 5th meetings. United Nations. 2020. Available at: https://www.un.org/press/en/2020/ga12268.doc.htm [Accessed 01 February 2020]

[3] World Health Organization. A world at risk: Annual report on global preparedness for health emergencies. 2019. Available at: https://apps.who.int/gpmb/assets/annual_report/GPMB_annualreport_2019.pdf [Accessed 01 February 2020]

[4] BBC. Large Ebola outbreaks new normal, says WHO. 2019. Available at: https://www.bbc.co.uk/news/health-48547983 [Accessed 01 February 2020]

[5] World Health Organization. Foodborne diseases. 2021. Available at: https://www.who.int/health-topics/foodborne-diseases. [Accessed 01 February 2020]

[6] Wallinga, J. and Teunis, P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol. 160(6): 509-16 (2004).

[7] Robert, A. Lessons from New Zealand's COVID-19 outbreak response. Lancet Public Health. 5(11): 30237-1 (2020).

[8] Niu, H., Zhang, W., Wei, L., et al. Rapid nanopore assay for carbapenem-resistant *Klebsiella pneumoniae*. Front Microbiol. 2019. 10:1672 (2019).

[9] Roberts, L. W., Harris, P. N. A., Forde, B. M., et al. Integrating multiple genomic technologies to investigate an outbreak of carbapenemase-producing *Enterobacter hormaechei*. Nat Commun. 11(1):466 (2020).

[10] European Centre for Disease Prevention and Control. Expert opinion on whole genome sequencing for public health surveillance. 2016; Available at: https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/whole-genome-sequencing-for-public-health-surveillance.pdf [Accessed 01 February 2020]

[11] Lu, R., Zhao, X., Li, J., et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet. 395(10224):565-574 (2020).

[12] Charalampous, T., Alcolea-Medina, A., Snell, L.B., et al. Application of respiratory metagenomics for COVID-19 patients on the intensive care unit to inform appropriate initial antimicrobial treatment and rapid detection of nosocomial transmission. medRxiv. 20229989 (2020).

[13] ARTIC Network. How to read a phylogenetic tree. Available at: https://artic.network/how-to-read-a-tree.html [Accessed: 05 February 2021]

[14] Worobey, M. Epidemiology: Molecular mapping of Zika spread. Nature. 546(7658):355-357 (2017).

[15] Faria, N. R. A sequencing road trip at the heartland of Zika virus in the Americas. Nature Microbiology. 2017.Available at: https://naturemicrobiologycommunity.nature.com/posts/17178-a-sequencing-road-trip-at-the-heartland-of-zika-virus-in-the-americas [Accessed 01 February 2020]

[16] European Commission. A united front to beat COVID-19. Available at: https://ec.europa.eu/info/files/communication-united-front-beat-covid-19_en [Accessed: 01 February 2021]

[17] World Health Organization. Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health. 2021. Available at: https://www.who.int/publications/i/item/9789240018440 [Accessed: 01 February 2021]

[18] Quick, J., Loman, N. J., Duraffour, S., et al. Real-time, portable genome sequencing for Ebola surveillance. Nature. 530(7589):228-232 (2016).

[19] Faria, N. R., Sabino, E. C., Nunes, M. R., et al. Mobile real-time surveillance of Zika virus in Brazil. Genome Med. 8(1):97 (2016).

[20] Brunker, K., Jaswant, G., Thumbi, S. M. et al. Rapid in-country sequencing of whole virus genomes to inform rabies elimination programmes. Wellcome Open Res. 5:3 (2020).

[21] Gigante, C. M., Yale, G., Condori, R.E. et al. Portable rabies virus sequencing in canine rabies endemic countries using the Oxford Nanopore MinION. Viruses 12:1255 (2020).

[22] COVID-19 Genomics UK Consortium. 2020. Available at: https://www.cogconsortium.uk [Accessed: 01 February 2021]

[23] Meredith, L. W., Hamilton, W. L., Warne, B., et al. Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic surveillance study. Lancet Infect Dis. 20(11):1263-1272 (2020).

[24] The Washington Post. Scientists have a powerful new tool for controlling the coronavirus: Its own genetic code. 2020; Available at: https://www.washingtonpost.com/graphics/2020/health/coronavirus-genetic-code [Accessed: 01 February 2021]

[25] University of Cambridge. Rapid genome sequencing and screening help hospital manage COVID-19 outbreaks. 2020. Available at: https://www.cam.ac.uk/research/news/rapid-genome-sequencing-and-screening-help-hospital-manage-covid-19-outbreaks [Accessed: 01 February 2021]

[26] Kafetzopoulou, L. E., Pullan, S. T., Lemey, P., et al. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. Science. 363(6422):74-77 (2019).

[27] Nigerian government. Early results of Lassa virus sequencing & implications for current outbreak response in Nigeria. reliefweb. 2018. Available from: https://reliefweb.int/report/nigeria/early-results-lassa-virus-sequencing-implications-current-outbreak-response-nigeria [Accessed: 01 February 2021]

[28] World Health Organization. Pneumonia of unknown cause — China. 2020. Available at: https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en [Accessed: 01 February 2021]

[29] Oude Munnink, B. B., Nieuwenhuijse, D. F., Stein, M., et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. Nat Med. 26(9):1405-1410 (2020).

[30] Sikkema, R. S., Pas, S. D., Nieuwenhuijse, D. F., et al. COVID-19 in health-care workers in three hospitals in the south of the Netherlands: a cross-sectional study. Lancet Infect Dis. 20(11): 1273-1280 (2020).

[31] Oreshkova, N., Molenaar, R. J., Vreman, S., et al. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. Eurosurveillance. 25(23):2001005 (2020).

[32] Koopmans, M. SARS-CoV-2 and the human-animal interface: outbreaks on mink farms. Lancet Infect Dis. 21(1):18-19 (2020).

[33] BBC. Call for coronavirus screening at mink farms. Available at: https://www.bbc.co.uk/news/science-environment-55048418. [Accessed 02 February 2021]

[34] Rambo-Martin, B. L., Keller, M. W., Wilson, M. M., et al. Influenza A virus field surveillance at a swine-human interface. mSphere. 5:1 (2020).

[35] Peacock, S. J. Reflections on the achievements of COG-UK. COG Consortium. 2021. Available at: https://www.cogconsortium.uk/news_item/reflections-on-the-achievements-of-cog-uk [Accessed: 01 February 2021]

[36] SARS-CoV-2 genomes. GISAID. 2021. Available at: https://www.gisaid.org [Accessed: 26 January 2021]

[37] COG Consortium. COG Consortium COG-UK report on SARS-CoV-2 Spike mutations of interest in the UK. 2021. Available at: https://www.cogconsortium.uk/wp-content/uploads/2021/01/Report-2_COG-UK_SARS-CoV-2-Mutations.pdf [Accessed: 01 February 2021]

[38] Rambaut, A., Loman, N., Pybus, O. G., et al. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. Virological. 2020; Available at: https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563 [Accessed: 01 February 2021]

[39] Peacock, S. J. Reflections on the achievements of COG-UK. COG Consortium. 2021; Available at: https://www.cogconsortium.uk/news_item/reflections-on-the-achievements-of-cog-uk

[40] Tegally, H., Wilkinson, E., Giovanetti, M. , et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. medRxiv. 20248640 (2020).

[41] Faria, N. R., Morales Claro, I., Candido, D., et al. Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. Virological. 2021; Available at: https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586 [Accessed: 01 February 2021]

[42] Pater, A. A., Bosmeny, M. S., Barkau, C. L., et al. Emergence and evolution of a prevalent new SARS-CoV-2 variant in the United States. bioRxiv. 426287 (2021).

Oxford Nanopore Technologies
phone +44 (0)845 034 7900
email sales@nanoporetech.com
twitter @nanopore

www.nanoporetech.com

WP_1097(EN)_V1_09Feb2021