Oxford NANOPORE Technologies

INFECTIOUS DISEASE

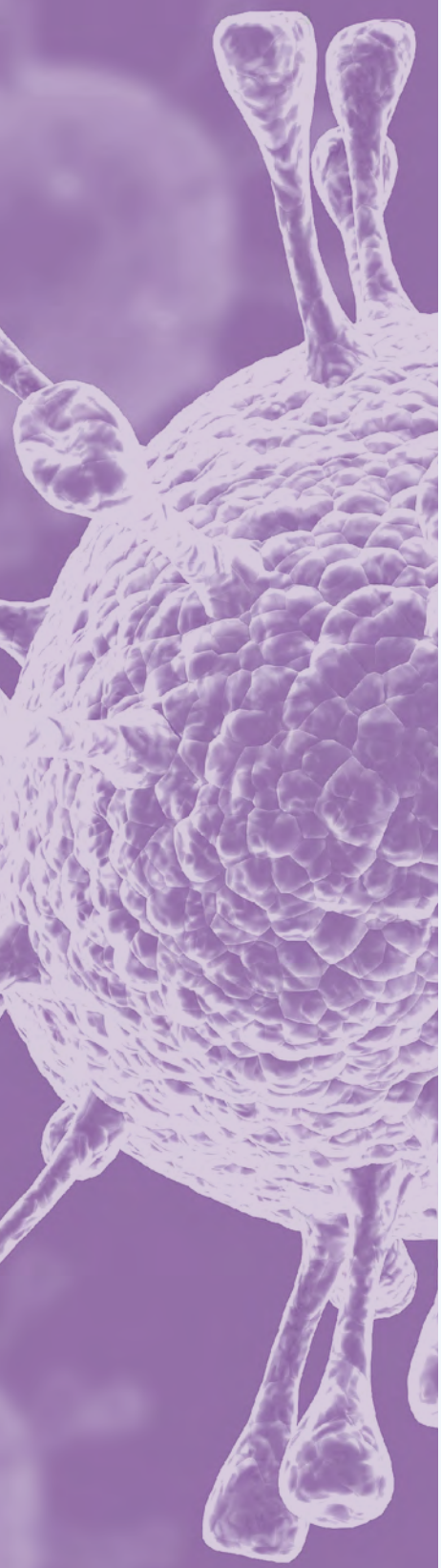
Delivering the future of genomic pathogen surveillance

WHITE PAPER

October 2022

nanoporetech.com/publications

nanoporetech.com



Contents

Introduction

Towards the end of 2019, news first broke of a novel, highly virulent pathogen that was capable of causing serious human respiratory distress and, in some cases, death. That pathogen, SARS-CoV-2 — a name now familiar to all — has since been attributed to over 6.5 million deaths and will cost the global economy an estimated $28 trillion in lost output by the end of 2025[1].

Serious infectious disease outbreaks have been an ever-present threat throughout human history and, driven by factors such as increasing globalisation, population growth, urbanisation, and climate change, that threat is increasing. The last two decades alone have witnessed numerous infectious disease epidemics, including severe acute respiratory syndrome (SARS; 2003), H1N1 swine flu (2009), Ebola (2014–18), and monkeypox (2022) — each with significant potential for global concern (**Figure 1**).

The rapid advancement of genomic technologies means that it is now possible to characterise pathogens more easily and comprehensively than ever before; however, due to high costs and infrastructure requirements, global access to such technologies has traditionally been limited to well-funded, centralised laboratories. Furthermore, in order to inform public health response to an infectious disease outbreak, it is critical that genomic information is generated and delivered quickly. This is a significant challenge for traditional sequencing technologies that require time-consuming sample preparation workflows and lengthy, fixed run times, only after which can the data be delivered and analysed.

This white paper reviews how nanopore sequencing overcomes the challenges associated with traditional sequencing technologies to deliver comprehensive, affordable, and real-time analysis of human and animal pathogens. Case studies reveal how researchers around the world are now deploying nanopore sequencing to better identify, characterise, and track a range of pathogens to help minimise the threat and impact of disease outbreaks.

The importance of pathogen surveillance

Genomic epidemiology — the study of how variations in genomes of pathogens or their hosts influence health and disease — is a crucial weapon in the public health fight against infectious diseases. As evidenced by the COVID-19 pandemic, genomic epidemiology supports precise tracking of pathogen evolution, providing detailed insights into sources of infection, routes of transmission, and the potential association of novel pathogen variants with changes to disease severity, transmission, and diagnostic and therapeutic efficacy.

Significant new outbreaks are an ongoing risk and may stem 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.

> ***'We are entering a very new phase of high impact epidemics ... This is a new normal, I don't expect the frequency of these events to reduce'[2]***

New human pathogens often arise from microbes that normally infect animals but have acquired the ability to infect humans (see Case study 1). 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.

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. For example, it has been suggested that the 2003 SARS outbreak would have tripled in size had there been a delay of just one week in applying control measures[3].

Examination of the relative success of New Zealand in controlling 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[5]. These measures minimised the spread of disease, providing more time to develop and roll out an effective vaccination programme.

> ***'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'[4]***

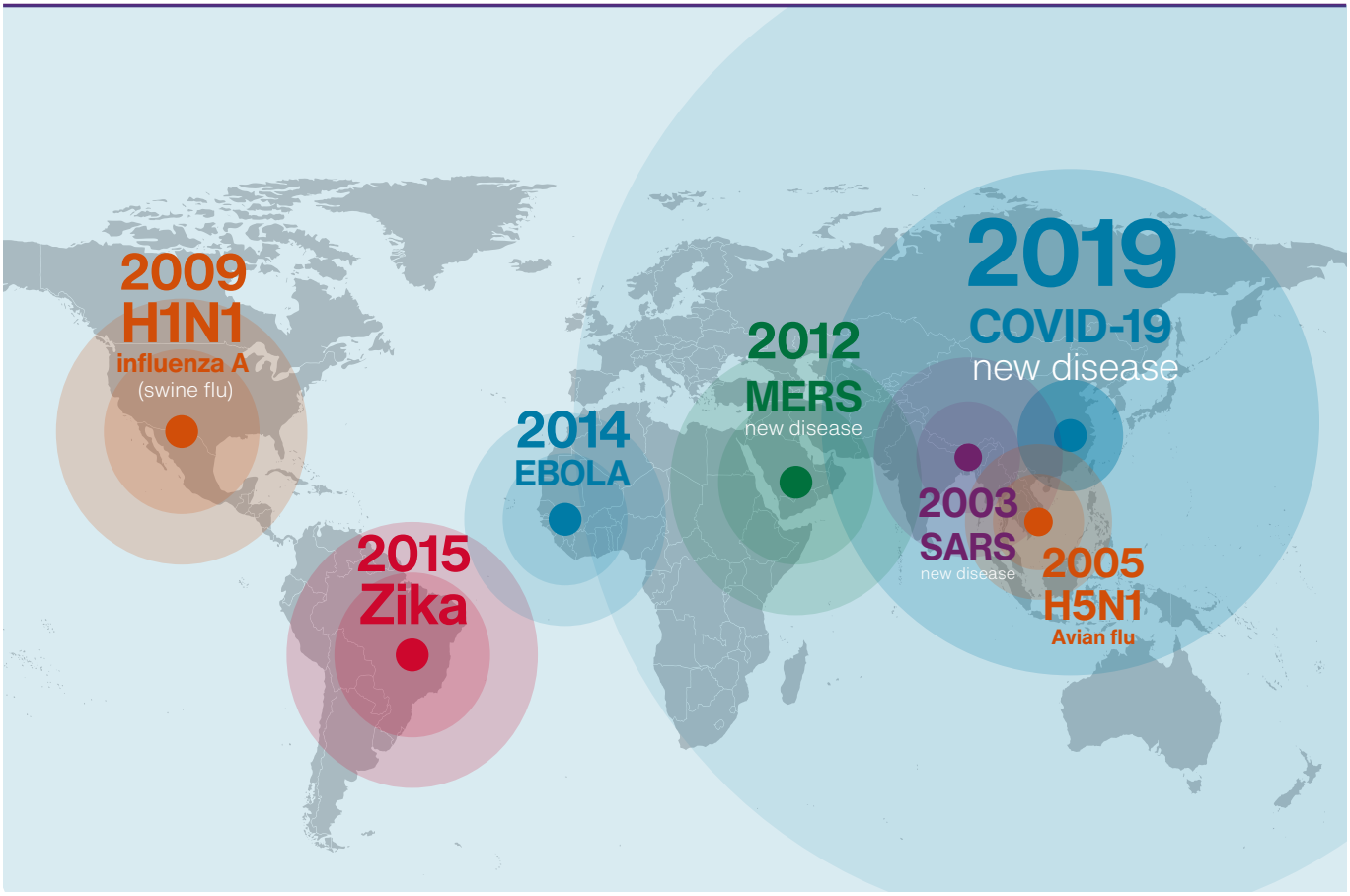


Figure 1
Selected infectious disease outbreaks in the 21st century. Infectious diseases are an increasing threat, for example in 2019 alone the World Health Organization (WHO) recorded over 100 outbreaks of 19 different infectious diseases, each posing a potential epidemic or pandemic threat[6].

The advantages of nanopore sequencing for pathogen surveillance

Rapid access to results

Time to result is of critical importance during a novel infectious disease outbreak. Firstly, the causative pathogen must be identified. This is most commonly achieved using a metagenomic sequencing approach, whereby the nucleic acids from all the organisms in a clinical research sample are analysed. Metagenomic sequencing requires no prior knowledge of the infectious organism and is also useful for the identification of co-infections, where multiple different pathogens may be present in a single sample. For example, deploying real-time metagenomic sequencing to research samples from COVID-19 patients in intensive care units has been shown to enable the identification of secondary infections, antimicrobial resistance (AMR), and hospital-associated transmission[7].

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 and lower sequencing and analysis costs. They also enable the analysis of samples with lower levels of pathogen.

> ***'Identification of pathogens was possible after 10 minutes of sequencing and all predefined AMR-encoding genes and plasmids from monoculture experiments were detected within one hour using raw nanopore sequencing data'[8]***

Nanopore sequencing offers several advantages over alternative sequencing technologies for pathogen genome analysis, including the facility for real-time data streaming. Unlike traditional sequencing technologies that deliver data in bulk at the end of a lengthy sequencing run, nanopore technology provides access to data as soon as it is generated. Using nanopore sequencing, researchers have been able to identify pathogens from complex metagenomic samples in minutes, with further insights, such as AMR profiling, becoming available as soon as sufficient data has been generated[9,10,11]. The facility of real-time analysis to deliver significantly faster access to actionable results can be crucial in outbreak scenarios.

Oxford Nanopore Technologies offers a range of streamlined library preparation kits that further shorten time to result when compared to traditional sequencing technologies. For example, Rapid Sequencing Kits, which include barcoding options for cost-effective analysis of multiple samples in a single run, allow researchers to prepare their purified sample for nanopore sequencing in as little as 10 minutes. In addition, the Field Sequencing Kit, which utilises the same rapid sequencing chemistry, does not require refrigeration, enabling researchers to more easily deploy nanopore sequencing in the field or remote, resource-limited environments.

Further workflow efficiencies can be obtained using Oxford Nanopore's growing number of real-time EPI2ME™ data analysis pipelines, which include species identification, AMR profiling, and SARS-CoV-2 variant identification (**Figure 2**). These best practice pipelines are available as simple push-button, cloud-based workflows, as well as more customisable, locally hosted workflows and interactive tutorials.

Find out more about data analysis at nanoporetech.com/analyse.

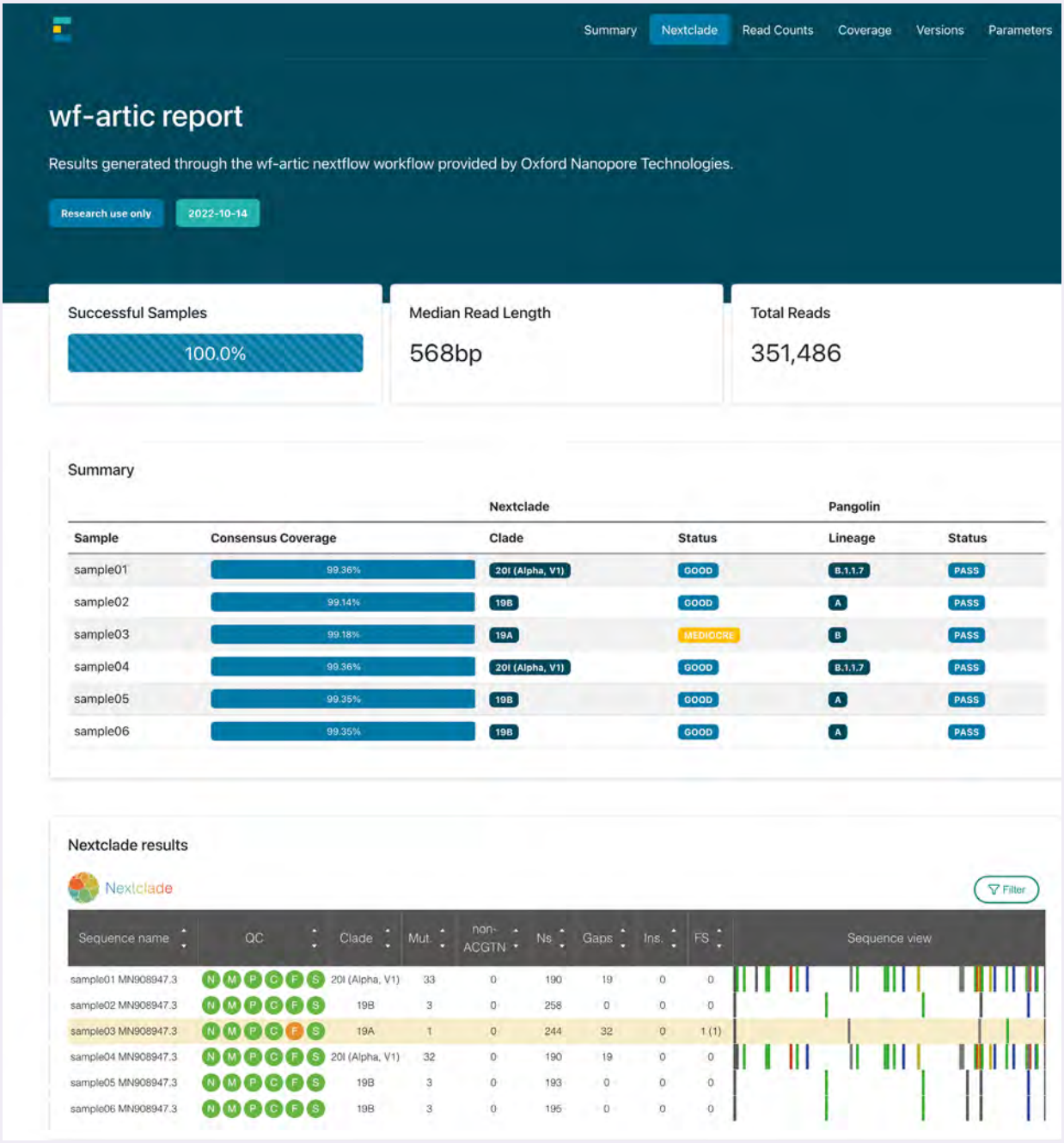


Figure 2

The EPI2ME Labs wf-artic workflow, which is based on the FieldBioinformatics workflow from the ARTIC network[12], provides a streamlined pipeline for analysing SARS-CoV-2 nanopore sequencing data. The workflow supports multiplexed sequencing runs, provides essential QC information, and enables variant identification using both NextClade and Pangolin, which deliver clades and pango lineages, respectively.

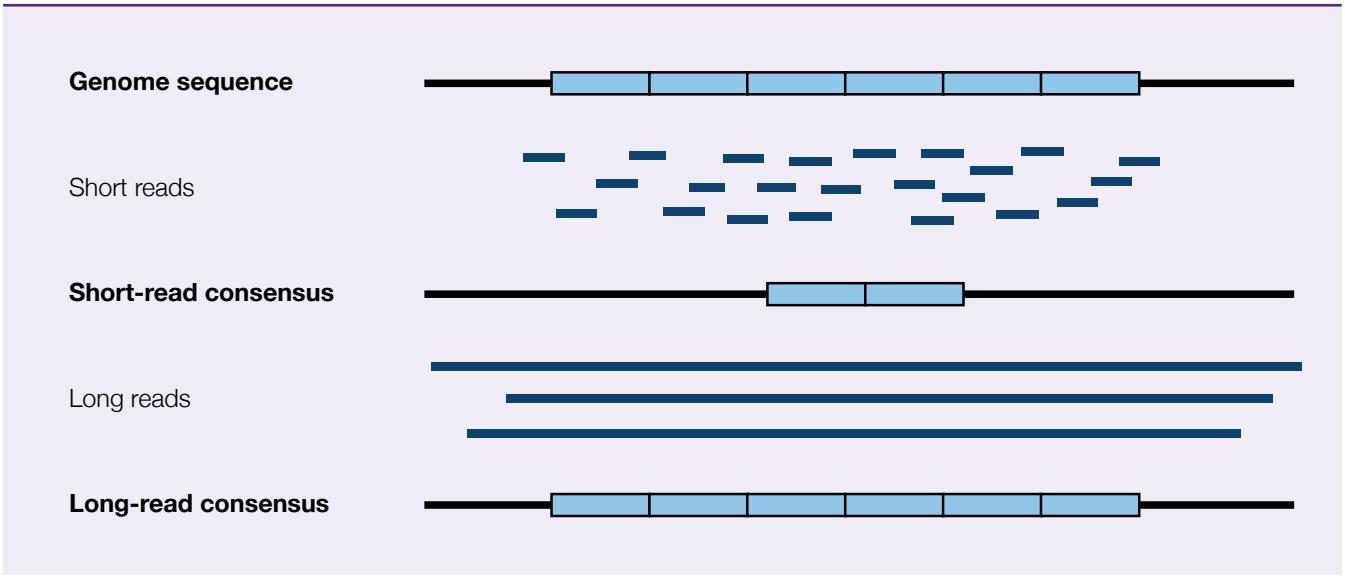


Figure 3

A schematic highlighting the advantages of long reads in *de novo* assembly of repetitive regions. Repeat and high-identity regions are often collapsed during the assembly process when using short-read sequencing technology. In contrast, long read lengths are more likely to incorporate the whole repetitive region (blue boxes), allowing more accurate assembly. Image adapted from Kellog[17].

The importance of long sequencing reads

The facility of a pathogen to cause disease or evade treatment is often encoded in highly repetitive genomic regions, such as those flanking AMR and virulence factors. Genomic repeats also frequently serve as sites for structural variation (SV), which contribute to the facility for rapid microbial evolution[13].

In order to fully characterise and functionally annotate a given pathogen, it is imperative to create a complete genome assembly; however, traditional sequencing technologies, which typically process short (<300 bp) DNA fragments, struggle to resolve repetitive regions and SVs.

In contrast, nanopore sequencing technology processes the entire DNA (or RNA) fragment that is presented to the nanopore, regardless of its size. Complete fragments of hundreds of kilobases are routinely processed and ultra-long read lengths in excess of 4 Mb have been generated[14]. Clearly, such long reads are more likely to span entire regions of repetitive DNA and SVs, allowing more complete, gap-free genome assembly[15](**Figure 3**). In addition, unlike traditional techniques, nanopore sequencing does not require DNA amplification, eliminating a key source of sequencing bias.

In a study by Browne *et al*., out of five sequencing technologies tested, only nanopore sequencing exhibited no GC bias[16].

A further advantage of long reads is the facility to distinguish between closely related species or strains. This supports both microbial identification and the generation of high-quality metagenome-assembled genomes (MAGs) from complex, mixed microbial samples[18].

Accessible sequencing for anyone, anywhere

Establishing and maintaining local capacity for rapid deployment of pathogen genomic sequencing can provide a critical time advantage in containing outbreaks; however, traditional next generation sequencing (NGS) technologies are typically expensive (>$100k–$1M)[19], bulky, fragile, cumbersome to ship, and require substantial power infrastructure. This, coupled with a general lack of facilities and trained personnel, make such technologies difficult to deploy in remote locations from where many novel infectious diseases emerge. Furthermore, the requirement for precise calibration by trained engineers means that most sequencing platforms are incompatible with rapid deployment.

> ***'Accessibility to genome sequencing has typically been limited to large, well-funded institutions due to large capital requirements of DNA sequencing platforms'[20]***

Nanopore sequencing technology resolves these challenges, making sequencing accessible to anyone, anywhere. Starting at just $1,000, MinION™ Starter Packs (which include a MinION device, flow cell, and sequencing reagents), offer affordable sequencing for every lab. In addition, MinION devices (**Figure 4**; **Table 2**) are uniquely portable, allowing sequencing to be undertaken at sample source, removing the considerable time and cost implications of shipping samples to a centralised sequencing facility.

> ***'The fact that almost half of all SARS-CoV-2 sequencing in Africa was performed using the Oxford Nanopore technology, which is relatively low-cost compared to other sequencing technologies and better adapted to modest laboratory infrastructures, illustrates one component of how this rapid scale-up of local sequencing was achieved'[21]***

MinION and MinION Mk1C — which includes a high-resolution touchscreen and powerful integrated compute for onboard, real-time data analysis — are ideal for local implementation and resource-limited settings that ordinarily might not have access to such powerful and versatile sequencing technology.

Scalable to all sample throughput requirements

Routine pathogen sequencing during an outbreak is vital to identify novel variants that may increase pathogenicity or reduce the effectiveness of approved diagnostic assays or therapeutics. In addition, the genomic data allows patterns of transmission to be elucidated to support outbreak management. The power of genomic epidemiology to answer these questions relies on the rapid analysis of high numbers of samples. For example, at an early stage of the COVID-19 pandemic, the European Commission urged member states to sequence 10% of positive test results[22].

> ***'The key strengths of Oxford Nanopore's sequencing platforms are that they are a low capital investment, have competitive per sample sequencing cost when multiplexing, the possibility of bias-free PCR library preparation, cold-chain-free sequencing reagents, fast turnaround time, use of long reads to resolve complex genomic loci, and the ability to investigate methylation status'[23]***

As with traditional sequencing platforms, nanopore sequencing allows multiple samples to be analysed on a single run. For example, the Midnight protocol enables up to 96 whole-genome SARS-CoV-2 samples to be analysed on a single MinION Flow Cell. However, unlike traditional platforms, nanopore flow cells, with active pores remaining after a sequencing run, can be washed and reused. This allows sample throughput to be scaled to suit the number of samples available — further reducing time to result and costs.

For higher sample throughput requirements, the benchtop GridION™ device can run up to five MinION Flow Cells simultaneously (**Figure 4**; **Table 2**). Each flow cell can be operated independently, allowing sequencing runs to be started as and when required, and offering the facility to rapidly adapt throughput for smaller-scale pathogen surveillance and higher-scale outbreak scenarios.

The latest additions to the portfolio, the PromethION™ 2 and PromethION 2 Solo (**Figure 4**; **Table 2**), offer the flexibility of two high-output PromethION Flow Cells, each capable of delivering up to five times the sequencing yield of a MinION Flow Cell. These yields are highly suited to the delivery of high-quality MAGs from complex samples, such as bodily fluids, faecal samples, and environmental samples, where the

Figure 4

Oxford Nanopore sequencing devices (from top left to bottom right): Flongle™, a flow cell adapter for MinION and GridION; the portable MinION and MinION Mk1C; GridION, with capacity for up to five Flongle or MinION Flow Cells; PromethION 2 Solo and PromethION 2; and the high-throughput PromethION 24 and 48 (shown) devices, capable of delivering up to 7 and 14 Tb of data, respectively*.

pathogen DNA/RNA may represent just a small fraction of the nucleic acids present.

For even higher throughput applications, the PromethION 24 and 48 devices enable up to 24 or 48 high-output PromethION Flow Cells to be run, respectively (**Figure 4**; **Table 2**).

> ***'We demonstrated here that multiplexing on a Flongle may further present significant cost reduction for simple assays, such as amplicon sequencing, without deterioration of the sequencing quality'[24]***

For cost-effective analysis of smaller assays, Oxford Nanopore offers Flongle, a flow cell adapter for MinION and GridION devices, enabling the use of low-cost, single-use Flongle Flow Cells (**Figure 4**; **Table 2**). Researchers have demonstrated the use of Flongle for cost-effective pathogen analysis, including enterovirus genotyping[24] and the discrimination of monkeypox virus from other closely related pathogens[25].

The same sequencing chemistry and kits are used across all nanopore sequencing devices, allowing researchers to easily scale their assays as required. Furthermore, nanopore sequencing devices are available in capital-free Starter Packs, where the only cost is the consumable flow cells and kits, which can be scheduled to match usage.

Proven performance

One of the earliest examples of the suitability of nanopore sequencing for pathogen and outbreak sequencing was evidenced in the 2015 Ebola epidemic in West Africa, where researchers from the University of Birmingham and Public Health England utilised the MinION as part of a mobile sequencing unit on site in Guinea[26]. Within two days of arriving on site, the team were able to perform real-time genomic surveillance, delivering sequencing data within just 15–60 minutes of sample preparation.

> ***'We show that real-time genomic surveillance is possible in resource-limited settings and can be established rapidly to monitor outbreaks'[26]***

* Theoretical max output (TMO). Assumes system is run for 72 hours (or 16 hours for Flongle) at 420 bases/second. Actual output varies according to library type, run conditions, etc. TMO noted may not be available for all applications or all chemistries.

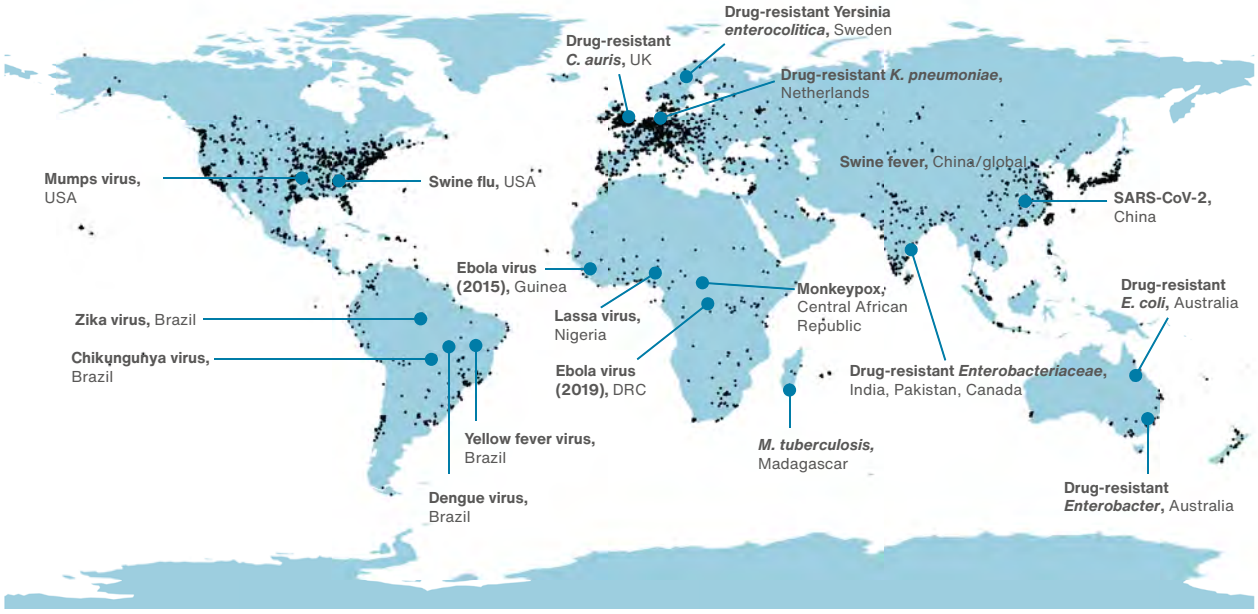


Figure 5

Selected human pathogen outbreaks analysed using nanopore sequencing. Affordable, real-time nanopore sequencing is being used extensively around the world for rapid insights into human and animal pathogens. Dots represent nanopore users for a range of applications, including outbreaks. Data correct as of October 2022.

Since then, nanopore sequencing has delivered actionable insights into human and animal pathogen surveillance and outbreaks around the world, including viral pathogens (e.g. Zika virus[27], rabies virus[28,29], Lassa virus[30], swine influenza[31], monkeypox virus[32], and SARS-CoV-2[33]), bacterial pathogens (e.g. *Mycobacterium tuberculosis*[34], *Salmonella spp.*[35], *Enterobacteria spp.*[36]) and fungal pathogens (e.g. *Candida spp.*[37,38]) (**Figure 5**). Demonstrating the global utilisation of nanopore sequencing, as of October 2022, over 1.4 million SARS-CoV-2 genomes had been generated using nanopore technology across 197 countries and territories[39].

Genomic surveillance of circulating viruses can enable the development of candidate vaccine viruses, which form the basis for a new vaccine before an outbreak occurs. Using the MinION device as part of a mobile sequencing station at a large swine exhibition, researchers in the USA identified three different swine influenza A viruses, one of which was genetically distinct to the most similar candidate vaccine virus — including differences in parts of the genome known to determine vaccine efficiency[31]. Using this information, the US Centers for Disease Control and Prevention (CDC) were able to develop a new candidate vaccine virus. Importantly, all data were obtained within 18 hours of setting up the mobile facility and, according to the researchers, their '*proactive surveillance efforts and vaccine derivation would have provided an approximate 8-week time advantage for vaccine manufacturing*'[31].

> ***'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'[31]***

Methylation calling as standard

Base modifications (e.g. 5-methylcytosine and adenine N6-methyladenosine) in microbial genomes have been linked to a range of functions, including gene expression, AMR, and, in prokaryotes, protection against bacteriophages. Furthermore, research has also suggested that modified bases within host and pathogen genomes

underlie intricate biological interactions that could provide new markers for disease severity or novel therapeutic targets[40]. As a result, the detection of base modifications in the genomes of pathogens and their hosts is becoming an area of increasing interest; however, DNA/RNA amplification, which is a requirement for traditional sequencing technologies, removes these modifications, meaning they can only be studied by using additional time-consuming and costly sample processing and sequencing steps.

Nanopore sequencing does not require amplification or strand synthesis, meaning that both the base and its modification can be detected in the same sequencing run, in real time. To date, researchers have utilised nanopore sequencing to detect a wide a range of modified bases, including pseudouridine[41], N6-methyladenosine (m6A)[42], 4-methylcytosine (4mC)[43], 5-methylcytosine (5mC)[43], and 7-methylguanosine (m7G)[44].

'With longer reads as well as richer information of additional epigenetic modifications, developments in sequencing technology could bring another round of revolution in metagenome as well as virome investigations'[45]

Base modification detection not only provides greater depth of genomic characterisation but, as demonstrated by Tourancheau *et al.*, can also be used to support metagenomic contig binning, association of mobile genetic elements, and the identification of misassembled metagenomic contigs[43].

CASE STUDY 1

Rapid surveillance of monkeypox virus from complex metagenomic samples

International travel and urbanisation help novel pathogens enter new geographical areas. Researchers around the world are applying metagenomic analyses to identify and characterise emerging pathogens to better understand their epidemiology, transmission patterns, evolution, and adaptation to human transmission.

Since May 2022, monkeypox virus (MPXV) — a zoonotic virus endemic to the African continent — has been reported in over 30 countries outside of Africa. At the University of São Paulo in Brazil, Ingra Morales Claro and her team used nanopore sequencing to study the first confirmed case of monkeypox in Brazil[46]. The team isolated viral DNA from a clinical research sample taken from skin lesions, and prepared sequencing libraries using the Rapid PCR Barcoding Kit. Libraries were then sequenced on a MinION, producing a total of >950,000 reads and enabling a turnaround of 18 hours from DNA extraction to consensus sequence generation.

> ***'The average [read] depth was 277.7x, covering 100% of the viral genome with at least 1 read, and an N50 of 4,493'[46]***

The almost complete 197 kb genome assembly revealed that it clustered within clade 3 (the newly proposed B.1 lineage), which is associated with West Africa, and is closely related to MPXV sequences from Portugal, Germany, USA, and Spain. MPXVs in clade 3 are associated with lower disease severity than those in clade 1, which cluster in Central Africa.

Also combining metagenomics with nanopore sequencing are Joana Isidro and her colleagues at the National Institute of Health in Portugal. The team generated the first draft genome sequence of an MPXV associated with the recent major outbreak in Portugal[47]. Using DNA isolated from skin lesions, libraries were prepared using the PCR-free Rapid Barcoding Sequencing Kit and sequenced on a MinION device. Sequencing reads were analysed and mapped to a reference genome in real time using RAMPART — delivering immediate visualisation of genome coverage. Approximately 800,000 reads were generated within 18 hours, of which about 0.5% were identified as MPXV.

A comparison with the reference sequence from the 2018–19 multi-country MPXV outbreak confirmed the presence of approximately 50 SNPs in the newly sequenced MPXV from Portugal. This substitution rate is far greater than expected for Orthopoxviruses, suggesting that '*viral genome sequencing might provide sufficient resolution to track the transmission dynamics and outbreak spread, which seemed to be challenging for a presumably slow-evolving double-stranded DNA virus*'[48].

Phylogenetic analyses suggested the new sequence was closely related to viruses previously imported from Nigeria to the UK, Israel, and Singapore in 2018–19, and belonged to clade 3 (**Figure 6**)[47,48].

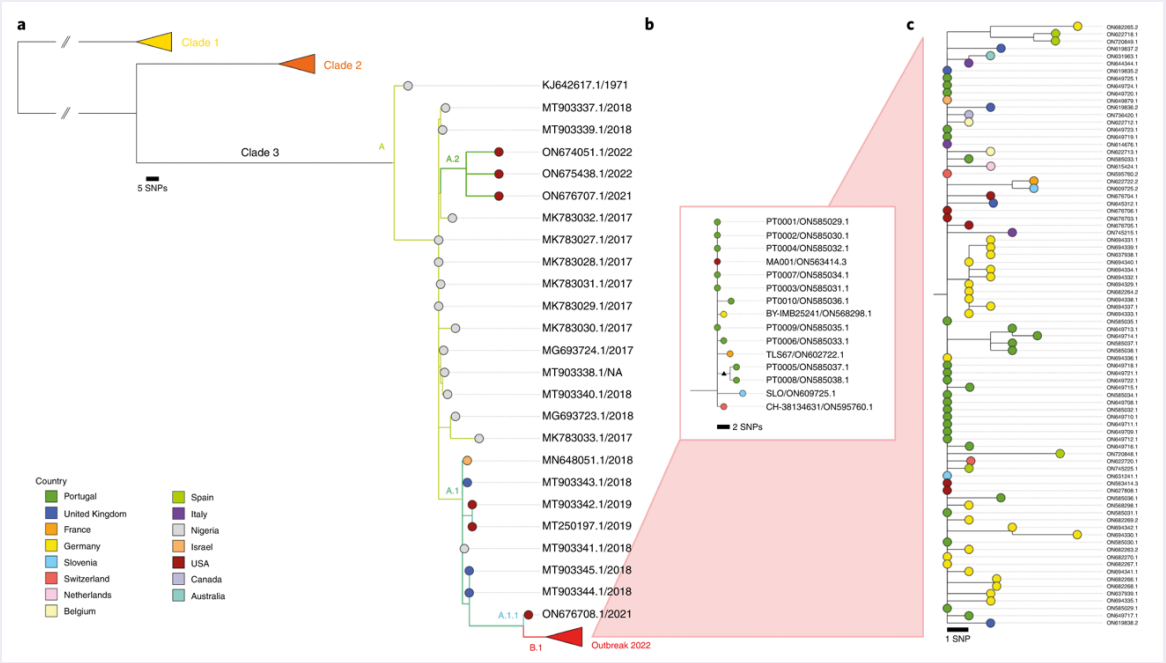


Figure 6

Phylogenetic analysis of MPXV viral sequences associated with the 2022 worldwide outbreak. a) Global phylogeny showing that the 2022 outbreak cluster (lineage B.1) belongs to clade 3. b) Genetic diversity within the outbreak cluster, including the first outbreak-related MPXV genome sequenced in Portugal and all available outbreak genome sequences released to the National Center for Biotechnology Information (NCBI) by 27 May 2022. c) Updated phylogenetic tree with sequences available in the NCBI as of 15 June 2022. Image taken from Isidro *et al.*[48] and available under Creative Commons license (creativecommons.org/licenses/by/4.0).

'Emergence of new viral infections with significant public health impact are frequent events, which re-enforces the need for comprehensive methodologies to detect rare, novel or emerging pathogens'[49]

Using nanopore technology, Adela Alcolea-Medina and her team at the Centre for Clinical Infection & Diagnostics Research, UK, developed a simple metagenomic workflow to detect low-abundance RNA and DNA viruses in different sample types[49]. Their approach applied genome sequencing using long nanopore reads and required only seven hours from sample receipt to answer — offering a significant advantage over metagenomic detection using traditional short-read sequencing technologies, which typically take three to five days. The method also demonstrated the potential to differentiate between viruses that present with similar symptoms. The streamlined workflow comprised library preparation, using the Rapid PCR Barcoding Kit, before sequencing on a GridION device, and live basecalling using MinKNOW™ for real-time analysis. Testing the workflow on four MPXV-positive nasopharyngeal research samples, the team detected MPXV in all four samples within 30 minutes of sequencing. To test their method using samples of unknown etiology, the team analysed four research samples from blistering skin rashes: MPXV was detected in three samples within minutes of sequencing. In two of the samples, 94% and 96% of the viral genome were recovered at ≥10x depth after 16 hours of sequencing. For the third research sample, only 49% of the genome was recovered, which may reflect the stage of disease at sample collection — day 18 of MPXV infection and receiving antiviral treatment. Varicella zoster virus (VZV), which causes chickenpox, was detected within 30 minutes of sequencing the fourth sample: 100% of the genome was sequenced to ≥10x depth of coverage.

These studies reveal how nanopore sequencing is supporting researchers to rapidly identify and fully characterise pathogens from complex metagenomic samples — providing actionable results to inform public health responses.

CASE STUDY 2

Wastewater sequencing — an early warning system for infectious disease outbreaks

Monitoring the genomic characteristics of pathogens circulating in a population can reveal important insights into the epidemiological dynamics of an outbreak. Unfortunately, sequencing every confirmed positive sample in a densely populated area is both challenging and expensive. Since viruses are shed in faeces, routine sequence analysis of wastewater samples could act as an early warning system for the silent transmission of viruses, helping public health bodies make quick, informed decisions to reduce their spread. Wastewater-based epidemiological analysis using nanopore sequencing could provide a rapid, accurate, and cost-effective solution to augment pathogen surveillance.

> ***'We ... built in nanopore sequencing so we could ... remove that requirement for shipping a sample to a specialised facility. Nanopore sequencing can be performed within countries where the samples are collected'[50]***

Outbreaks of poliovirus — wild-type or vaccine-derived — risk reversing the colossal progress of global eradication programmes. Successful control of any pathogen requires effective surveillance, involving rapid and accurate detection. At the Vaccine Epidemiology Research Group, Imperial College London, UK, Alex Shaw and colleagues were concerned with the time required to transport poliovirus samples to specialist sequencing facilities[50]. RNA viruses, such as poliovirus and SARS-CoV-2, are fast evolving, so they continually accumulate changes in their genome. These mutations can provide essential information on virus transmission, spread, and evolution but only if samples are analysed quickly enough. To overcome these challenges, they developed Direct Detection by Nanopore Sequencing (DDNS), which utilises the accessibility, scalability, and portability of nanopore sequencing to deliver faster access to results. DDNS could provide actionable results '*from sample to sequence in as little as 2 days', at a cost of around '$15 per sample, [including] reagents and kits*'[50]. Their approach used a nested PCR[51] to detect poliovirus in stool and wastewater samples (**Figure 7**). The two-step reaction first amplified the entire capsid region of the enteroviruses present in the sample. In the second step, the team used poliovirus-specific primers with barcodes[52] allowing the amplified products from multiple samples to be pooled and sequenced on a single flow cell. Such sample multiplexing enables very high-throughput analyses and reduces cost. Leveraging the streamlined Ligation Sequencing Kit library preparation workflow, the team then used both the MinION and GridION devices to sequence samples, identifying poliovirus in real-time using RAMPART[53] software, which maps reads according to serotype. The poliovirus-specific primers target the VP1 region of poliovirus, which is '*required for informing a vaccination response'* and can help *'map the movement of wild viruses*'[50].

Long, accurate sequencing reads enable easier analysis of complex environmental surveillance

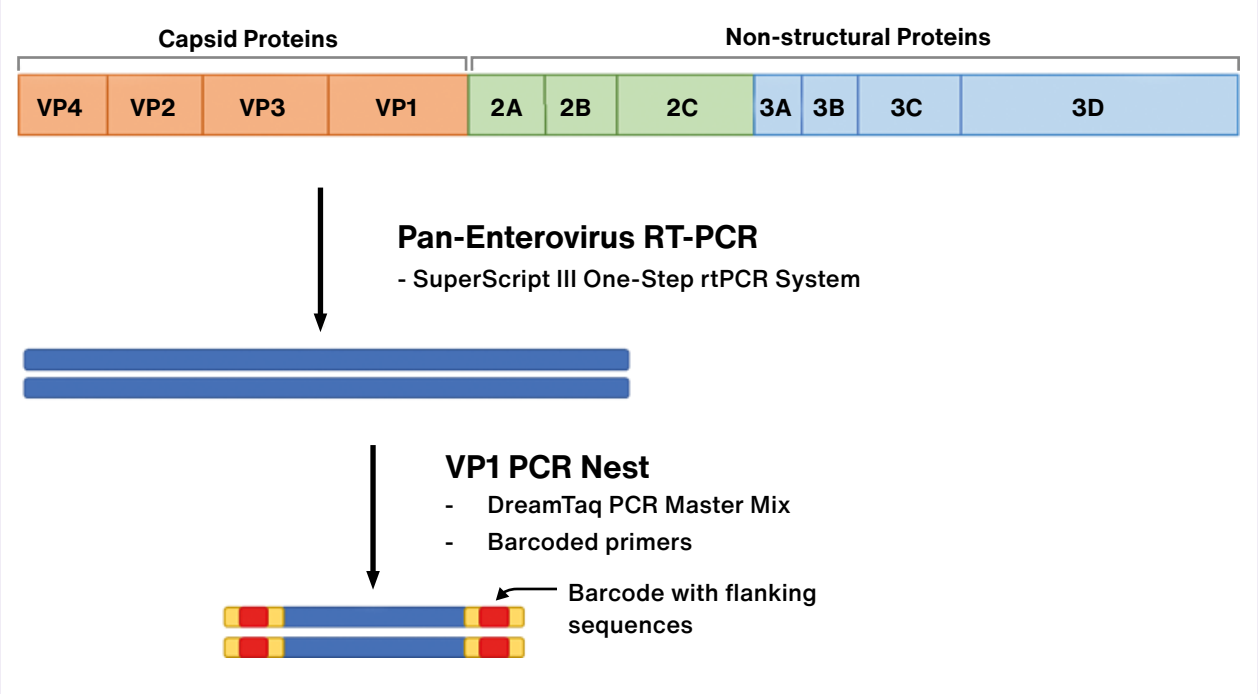


Figure 7

Poliovirus direct detection by nanopore sequencing (DDNS). Image kindly provided by Dr. Alex Shaw, Imperial College London, UK.

samples. DDNS identifies poliovirus from metagenomic samples, providing a sample-by-sample report of which polioviruses are present and any mutations that have been acquired, compared to a reference sequence. For added flexibility, the first round RT-PCR product can be used to identify species A and C enteroviruses. The second primer set amplifies species B and D enteroviruses from the same RNA sample — providing cost-effective enterovirus sequencing when combined with amplicon pooling and native barcoding. Global training efforts are enabling other labs to perform DDNS for genomic surveillance, such as at the Institut National de Recherche Biomédicale (INRB) in the Democratic Republic of the Congo. The team at INRB compared the results obtained for 2,365 samples analysed using the traditional cell culture method with the metagenomic DDNS technique, reporting '*excellent accuracy when comparing the two methods*'[50]. The sensitivity between the two approaches was as expected, based on previous work, and the specificity was '*excellent*', indicating the absence of contamination despite performing nested PCRs. The turnaround time was much faster using DDNS (**Figure 8**) and the median sequence identity was 100% when comparing DDNS with Sanger sequencing.

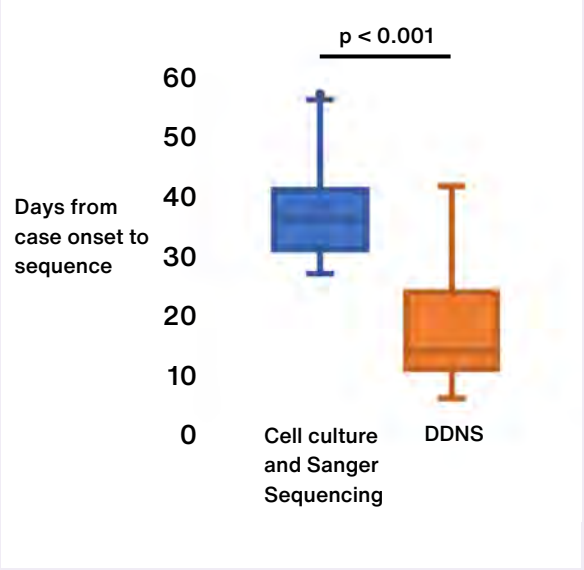


Figure 8

Comparison of the time taken for the traditional cell culture and Sanger sequencing method to detect poliovirus compared with DDNS at the INRB. Median time between stool collection and sequence generation was 36 days using cell culture and Sanger sequencing, and 14 days by DDNS. Image kindly provided by Dr. Alex Shaw, Imperial College London, UK.

Key samples that were used to confirm three poliovirus outbreaks were identified by DDNS 23 days faster than the cell culture method. This rapid turnaround time suggests DDNS could replace the traditional cell culture and Sanger sequencing method, which '*can take weeks*'. DDNS could provide a rapid, scalable, and cost-effective response for effective poliovirus surveillance, guiding public health initiatives in countries where the virus is still endemic.

> ***'[DDNS can] potentially ... lead to the ability to respond to outbreaks a lot more quickly. Quicker responses lead to smaller outbreaks'[50]***

Also leveraging the low cost and rapid turnaround times of nanopore sequencing in wastewater pathogen surveillance are Xuan Lin and Ryan Ziels with their colleagues from The University of British Columbia, Vancouver, Canada. Using a PCR enrichment approach to target SARS-CoV-2, the team assessed different multiplex primer schemes and wastewater sample types — influent wastewater and primary sludge[54].

Accessible, rapid, and cost-effective workflows to detect the emergence of variants of concern (VoC) within municipal areas would be a powerful tool during any infectious disease outbreak. Using a MinION to deliver real-time results, the team identified the use of influent wastewater and a 400 bp primer scheme as optimal conditions to accurately detect VoC within SARS-CoV-2 circulating in the Metro Vancouver region of British Columbia (**Figure 9**). The rapid turnaround time of approximately three days (from sampling to data generation), low capital cost, and high portability of nanopore sequencing, combined with the highly multiplexed tiling PCR sequencing approach, showed great promise in wastewater surveillance to complement genomic epidemiology efforts.

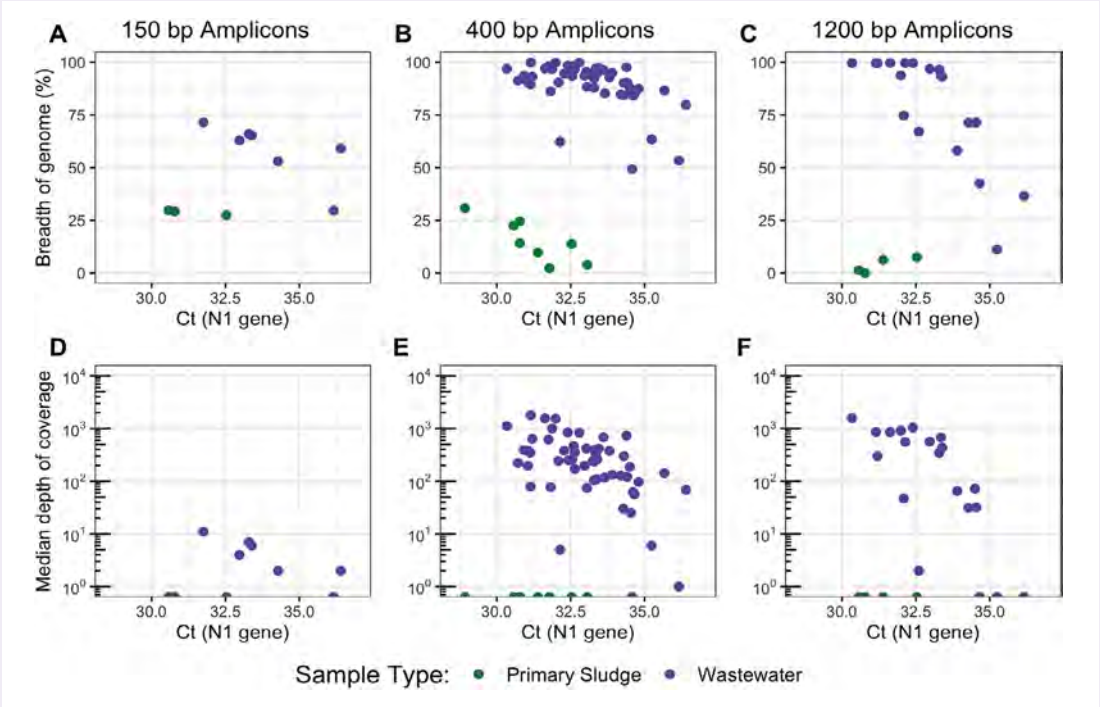


Figure 9

SARS-CoV-2 whole-genome sequencing coverage for three multiplex tiling PCR primer schemes (150 bp, 400 bp, and 1,200 bp), including breadth of genome coverage (A–C) and median depth of coverage across the genome (D–F). Use of the 400 bp tiling scheme on influent wastewater provided optimal breadth and depth of genome coverage. Image taken from Lin *et al.*[54] and available under Creative Commons license (creativecommons.org/licenses/by/4.0).

CASE STUDY 3

Nanopore-only assemblies for simple, cost-effective surveillance of antimicrobial resistance

At the London School of Hygiene and Tropical Medicine, UK, Ebenezer Foster-Nyarko and colleagues harnessed the long-read capability of nanopore sequencing to generate nanopore-only genome assemblies of *Klebsiella pneumoniae* — a common intestinal bacterium capable of causing life-threatening infections[55]. Over recent decades, *K. pneumoniae*, along with many other pathogens, has rapidly acquired antimicrobial resistance (AMR), and the World Health Organization now recommends robust surveillance of AMR as part of their critical Global Action Plan[56].

> ***'Nanopore ... sequencing has rich potential for genomic epidemiology and public health investigations of bacterial pathogens, particularly in low-resource settings and at the point of care, due to its portability and affordability'[55]***

In their publication, Foster-Nyarko *et al*.[55] noted that AMR determinants cannot be accurately delineated using short-read sequencing technologies because the '*fragmented genome assemblies ... cannot accurately resolve plasmids and other mobile genetic elements that drive the dissemination of AMR determinants*'. They further commented that nanopore sequencing reads generate contiguous genome assemblies, as the '*longer reads ... are capable of resolving structural variations, long repeat regions, and genomic copy-number alterations*', at '*low capital cost*'[55]. Although hybrid genome assemblies, containing both short-read and long-read data, have been used to overcome the limitations of short reads, continuous improvements in nanopore chemistry and analysis software now enable the generation of highly contiguous, highly accurate genome assemblies using nanopore sequencing alone.

Ebenezer generated genomic libraries using the Native Barcoding Kit, enabling multiplexing of up to 24 samples in a single run, before sequencing on the MinION. Genome sizes ranged from approximately 5.1 Mb to 6.2 Mb, with a GC content of 49.6% to 57.7%. Exploring both time and accuracy in AMR determination, three different real-time basecalling models were applied: Fast; High accuracy (HAC); and Super accuracy (SUP). A total of 270 acquired AMR genes were present across 54 reference genome assemblies, and the majority were correctly identified using SUP basecalling with Medaka polishing (95.8%). Mutations were identified in the *gyrA* and *parC* genes (substitutions associated with fluoroquinolone resistance), in bla_{SHV} (substitutions associated with extended spectrum β-lactamase activity), and in *ompK35* and *ompK36* (truncations associated with reduced susceptibility to β-lactams) (**Figure 10**). Colistin resistance is associated with truncations in *pmrB* and *mgrB* — these genes were intact in all reference assemblies.

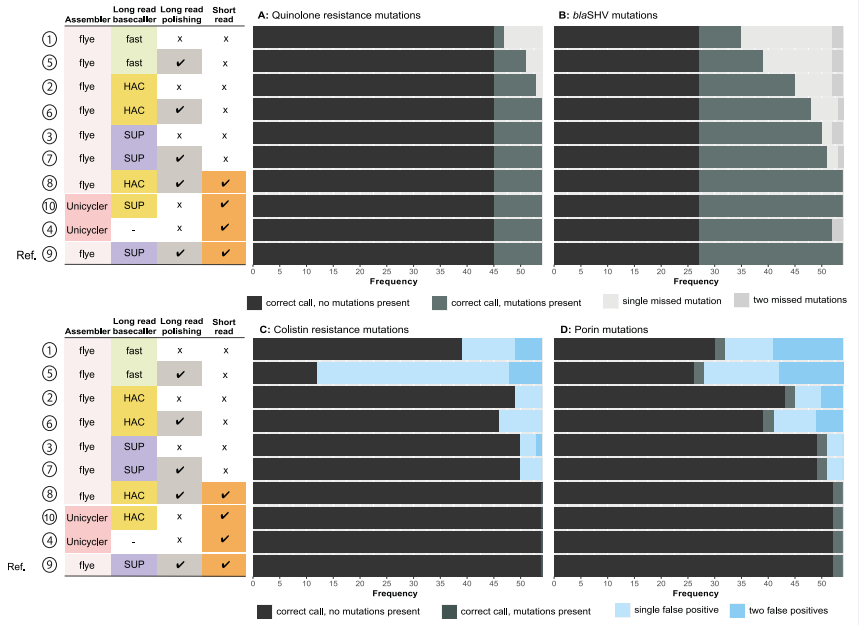


Figure 10

Detection of AMR-associated mutations. Each panel summarises the accuracy of genotyping across 54 samples for each assembly methodology, compared with the reference assembly. A) Detection of fluoroquinolone resistance-associated mutations. B) Detection of mutations in bla$_{SHV}$. C) Detection of colistin resistance-associated mutations. D) Detection of mutations in outer membrane porins which are associated with carbapenem susceptibility. Image kindly provided by Dr. Ebenezer Foster-Nyarko, London School of Hygiene and Tropical Medicine, UK.

	Aminoglycosides	Carbapenems	3rd gen cephalosporins	Fluoroquinolones	Phenicols	Sulfamethoxazole	Trimethoprim	Tetracycline
Reference genomes with AMR determinants N (% of total)	26 (48.1%)	6 (11.1%)	25 (46.3%)	^27 (50.0%)	17 (31.5%)	26 (48.1%)	24 (44.4%)	20 (37.0%)
Nanopore-only assemblies with AMR determinants N (% of reference genome positives)								
SUP+Medaka	**24 (92.3%)**	**6 (100%)**	**22 (88%)**	26 (96.3%)	**16 (94.1%)**	**26 (100%)**	**24 (100%)**	**20 (100%)**
SUP	11 (42.3%)	**6 (100%)**	21 (84%)	**27 (100%)**	14 (82.4%)	25 (96.2%)	**24 (100%)**	11 (55%)
HAC+Medaka	20 (76.9%)	**6 (100%)**	15 (60%)	**27 (100%)**	14 (82.4%)	24 (92.3%)	21 (87.5%)	15 (75%)
HAC	2 (7.7%)	**6 (100%)**	19 (76%)	**27 (100%)**	14 (82.4%)	8 (30.8%)	14 (58.3%)	2 (10%)
Fast+Medaka	3 (11.5%)	**6 (100%)**	8 (32%)	17 (63%)	13 (76.5%)	21 (80.9%)	19 (79.2%)	4 (20%)
Fast	1 (3.8%)	5 (71%)	6 (24%)	13 (48.1%)	12 (70.6%)	4 (15.4%)	8 (33.3%)	1 (5%)

Table 1

Accuracy of identifying AMR based on the detection of AMR determinants in the genome assembly. Data provided by Dr. Ebenezer Foster-Nyarko, London School of Hygiene and Tropical Medicine, UK.

***'There is ... now an unparalleled opportunity to harness Oxford Nanopore sequencing for genomic epidemiology and surveillance of bacterial pathogens, which is particularly attractive for AMR-associated pathogens such as* K. pneumoniae*'*[55]**

The presence of AMR-associated mutations is often used to predict drug resistance. In this research study, assemblies basecalled with SUP, assembled with Flye, and polished with Medaka identified '*88% – 100% of genomes with AMR determinants ... across the various drug classes*' (**Table 1**)[55].

Summary

The COVID-19 pandemic brought the devastation caused by infectious diseases into sharp focus, and improving global access to genomic analysis tools is now a key objective in the World Health Organization's ten-year pathogen surveillance strategy[57]. Pathogens know no borders, so to ensure effective genomic surveillance and timely outbreak response, it is imperative to build regional sequencing capability that interconnects with national and global infrastructure.

> ***'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'[4]***

Nanopore technology meets these critical needs, delivering fast, scalable, and cost-effective pathogen sequencing that can be implemented in any environment — from high-throughput public health and research labs through to rapidly deployable mobile labs at outbreak source.

Oxford Nanopore is proud to support the work of the global pathogen surveillance community to develop rapid and affordable genomic analysis solutions to mitigate the incidence and impact of future outbreaks.

About Oxford Nanopore Technologies

Oxford Nanopore's goal is to enable the analysis of anything, by anyone, anywhere. The company offers the only sequencing technology to combine scalability — from portable to ultra-high throughput formats — with real-time data delivery and the ability to elucidate accurate, rich biological data through the analysis of short to ultra-long fragments of native DNA or RNA.

The facility for real-time analysis combined with streamlined sequencing workflows provides rapid access to results for time-critical applications, such as pathogen surveillance and genomic epidemiology.

A range of sequencing devices is available, suitable for all pathogen surveillance and outbreak analysis requirements — from portable, in-field sequencing in resource-limited environments to high-throughput sample analysis in regional and national sequencing laboratories (**Table 2; Figure 4**).

	Flongle	MinION & MinION Mk1C	GridION	PromethION 2/2 Solo	PromethION 24/48
Read length	Fragment length = read length. Longest read >4 Mb[14]				
Run time	1 min – 16 hours	1 min – 72 hours	1 min – 72 hours	1 min – 72 hours	1 min – 72 hours
Number of flow cells per device	1	1	5	2	24/48
DNA sequencing yield per flow cell*	Up to 2.8 Gb	Up to 50 Gb	Up to 50 Gb	Up to 290 Gb	Up to 290 Gb
DNA sequencing yield per device*	Up to 2.8 Gb	Up to 50 Gb	Up to 250 Gb	Up to 580 Gb	Up to 7 Tb /14 Tb
Multiplexing	1 – 96 samples	1 – >2,000 samples	1 – >2,000 samples	1 – >2,000 samples	1 – >2,000 samples

* Theoretical max output (TMO). Assumes system is run for 72 hours (or 16 hours for Flongle) at 420 bases/second. Actual output varies according to library type, run conditions, etc. TMO noted may not be available for all applications or all chemistries

Table 2
A range of nanopore sequencing devices is available to suit all pathogen surveillance requirements. Data correct at time of print. Visit www.nanoporetech.com for the latest information.

For more information about utilising nanopore technology for pathogen sequencing, visit www.nanoporetech.com/infectious-disease.

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: 06 October 2022]

2. BBC. Large Ebola outbreaks new normal, says WHO. 2019. Available at: https://www.bbc.co.uk/news/health-48547983 [Accessed: 26 September 2022]

3. 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).

4. 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: 03 October 2022]

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

6. 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: 26 September 2022]

7. 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).

8. Taxt, A.M. et al. Rapid identification of pathogens, antibiotic resistance genes and plasmids in blood cultures by nanopore sequencing. Sci. Rep. 10, 7622 (2020).

9. Gu, W. et al. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. Nat. Med. 27(1):115–124 (2021).

10. Břinda, K. et al. Rapid inference of antibiotic resistance and susceptibility by genomic neighbour typing Nat. Microbiol. DOI: 10.1038/s41564-019-0656-6 [Online ahead of print] (2020).

11. Charalampous, T. et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. Nat. Biotechnol. 37 (7), 783–792 (2019).

12. FieldBioinformatics. Software. Available at: https://github.com/artic-network/fieldbioinformatics [Accessed: 14 October 2022]

13. Gorkovskiy, A. and Verstrepen, K.J. The role of structural variation in adaptation and evolution of yeast and other fungi. Genes (Basel) 12(5):699 (2021).

14. Oxford Nanopore Technologies. Ultra-Long DNA Sequencing Kit. Available at: https://store.nanoporetech.com/ultra-long-dna-sequencing-kit.html [Accessed: 03 October 2022]

15. Moss, E.L., Maghini, D.G., and Bhatt, A.S. Complete, closed bacterial genomes from microbiomes using nanopore sequencing. Nat. Biotechnol. 38(6):701–707 (2020).

16. Browne, P.D. et al. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. Gigascience.9(2):giaa008 (2020).

17. Kellog, E.A. Genome sequencing: Long reads for a short plant. Nat. Plants. 1, 15169 (2015).

18. Sereika, M. et al. Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. Nat. Methods. 19, 823–826 (2022).

19. Akaçin, İ., Ersoy, Ş., Doluca, O., and Güngörmüşler, M. Comparing the significance of the utilization of next generation and third generation sequencing technologies in microbial metagenomics. Microbiol. Res., 264, 127154 (2022).

20. McGinnis, J.L. and Giguere, D.J. High-quality genome assembly of a *Pestalotiopsis* fungus using DIY-friendly methods [version 1; peer review: 3 approved with reservations]. F1000Research 11:442 (2022).

21. Tegally, H. et al. The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance. Science, eabq5358 (2022). https://doi.org/10.1126/science.abq5358

22. 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: 03 October 2022]

23. Dippenaar. A et al. Nanopore sequencing for *Mycobacterium tuberculosis*: a critical review of the literature, new developments, and future opportunities. J. Clin. Microbiol. 60(1):e0064621 (2022).

24. Grädel, C. et al. Rapid and cost-efficient enterovirus genotyping from clinical samples using Flongle Flow Cells. Genes, 10(9), 659 (2019).

25. Israeli, O. et al. Rapid amplicon nanopore sequencing (RANS) for the differential diagnosis of monkeypox virus and other vesicle-forming pathogens. Viruses 14(8):1817 (2022).

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

27. 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).

28. 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).

29. 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).

30. 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).

31. 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):e00822–19 (2020).

32. Claro, I. M. et al. Shotgun metagenomic sequencing of the first case of monkeypox virus in Brazil, 2022. Revista do Instituto de Medicina Tropical de Sao Paulo, 64, e48 (2022).

33. Constantinides, B. et al. Rapid turnaround multiplex sequencing of SARS-CoV-2: comparing tiling amplicon protocol performance. medRxiv 21268461 (2022).

34. Gómez-González, P.J., Campino, S., Phelan, J.E., Clark, T.G. Portable sequencing of *Mycobacterium tuberculosis* for clinical and epidemiological applications. Brief. Bioinform.23(5):bbac256 (2022).

35. Wu, X. et al. Evaluation of *Salmonella* serotype prediction with multiplex nanopore sequencing. Front. Microbiol. 12:637771 (2021).

36. Roberts, L. W. et al. Integrating multiple genomic technologies to investigate an outbreak of carbapenemase-producing *Enterobacter hormaechei*. Nat. Commun. 11(1), 466 (2020).

37. Rhodes, J. et al. Genomic epidemiology of the UK outbreak of the emerging human fungal pathogen *Candida auris*. Emerg Microbes Infect. 7(1):43 (2018).

38. Chan, W. S. et al. Potential utility of targeted nanopore sequencing for improving etiologic diagnosis of bacterial and fungal respiratory infection. Diagnostic pathology, 15(1), 41 (2020).

39. GISAID. Website. https://gisaid.org/ [Accessed: 10 October 2022]

40. Gaspersic, J. and Dolzan, V. Viral and host genetic and epigenetic biomarkers related to SARS-CoV-2 cell entry, infection rate, and disease severity. Biology, 11(2), 178 (2022).

41. Begik, O., et al. Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing. Nat. Biotechnol. 39 (10), 1278–1291 (2021).

42. Price, A.M., Hayer, K.E., McIntyre, A.B.R. et al. Direct RNA sequencing reveals m6A modifications on adenovirus RNA are necessary for efficient splicing. Nat. Commun. 11, 6016 (2020).

43. Tourancheau, A, Mead, E.A., Zhang, X.S, and Fang, G. Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. Nat. Methods. 18(5):491–498 (2021).

44. Smith, A.M. et al. Reading canonical and modified nucleotides in 16S ribosomal RNA using nanopore direct RNA sequencing. PLoS One 14:e0216709 (2019).

45. Cao, J. et al. Profiling of human gut virome with Oxford Nanopore technology. Med. in Microecol. 4, 100012 (2020).

46. Claro, I.M. et al. Shotgun metagenomic sequencing of the first case of monkeypox virus in Brazil, 2022. Rev Inst Med Trop Sao Paulo. 24;64:e48 (2022).

47. Isidro, J. Monkeypox virus genome sequencing. Presentation. Available at: https://nanoporetech.com/resource-centre/monkeypox-virus-genome-sequencing-first-draft-genome-associated-2022-multi-country [Accessed 10 Oct 2022]

48. Isidro, J., et al. Phylogenomic characterization and signs of microevolution in the 2022 multi-country outbreak of monkeypox virus. Nat. Med. 28, 1569–1572 (2022).

49. Alcolea-Medina, A. et al. Novel, rapid metagenomic method to detect emerging viral pathogens applied to human monkeypox infections. doi.org/10.2139/ssrn.4132526 (2022).

50. Shaw, A. Implementation of rapid, direct sequencing for poliovirus surveillance. Presentation. Available at: https://nanoporetech.com/resource-centre/implementation-rapid-direct-sequencing-poliovirus-surveillance [Accessed: 5 Oct 2022]

51. Arita, M. et al. Development of an efficient entire-capsid-coding-region amplification method for direct detection of poliovirus from stool extracts. J. Clin. Microbiol. 53(1);73-78 (2015).

52. Shaw, A. G. et al. Rapid and sensitive direct detection and identification of poliovirus from stool and environmental surveillance samples by use of nanopore sequencing. J. Clin. Microbiol. 58(9); e00920-20 (2020).

53. ARTIC Network. Available at https://artic.network/rampart. [Accessed: 14 October 2022]

54. Lin, X. et al. Assessing multiplex tiling PCR sequencing approaches for detecting genomic variants of SARS-CoV-2 in municipal wastewater. mSystems 6(5):e0106821 (2021).

55. Foster-Nyarko, E. et al. Nanopore-only assemblies for genomic surveillance of the global priority drug-resistant pathogen, *Klebsiella pneumoniae*. bioRxiv 498322 (2022).

56. Global action plan on antimicrobial resistance. World Health Organization. 2016. Available at: https://www.who.int/publications/i/item/9789241509763 [Accessed: 12 October 2022]

57. World Health Organization. Global genomic surveillance strategy for pathogens with pandemic and epidemic potential. 2022. Available at: https://www.who.int/publications/i/item/9789240046979 [Accessed: 03 October 2022]

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

WP_1155(EN)_V1_19Oct2022