WHITE PAPER

# Large insights into microorganisms

Contents

Introduction

Microorganisms are the most abundant and diverse forms of life on Earth, with estimates ranging from millions to trillions of species (**Figure 1**)[1,2]. They are essential components of all ecosystems and play a crucial role in health, disease, and many industrial processes.

The field of microbiology has seen enormous impact from sequencing technologies. Traditionally, microorganisms have been studied through the culturing of individual species or strains using artificial culture media; however, of the ~10 million species so far catalogued, only ~10,000 (0.1%) have been cultured in the laboratory[2]. The advent of modern sequencing technologies, which allow high-throughput genomic analysis of cultured and, importantly, uncultured microbes, has drastically increased our capability to identify and characterise microorganisms.

Genome sequences for ~490,000 microbial strains are now publicly available[3]. However, due to the inherent limitations of traditional short-read sequencing technologies, much of what we know about microbial genomes is based on incomplete data. In fact, approximately 90% of bacterial genomes are considered to be incomplete[3,4].

This review will explore how microbiologists are now utilising the long-read capabilities of real-time nanopore sequencing to overcome the challenges of short-read sequencing technologies to fully characterise microbial genomes — shedding new light on microbial evolution, pathogenicity, and antimicrobial resistance.
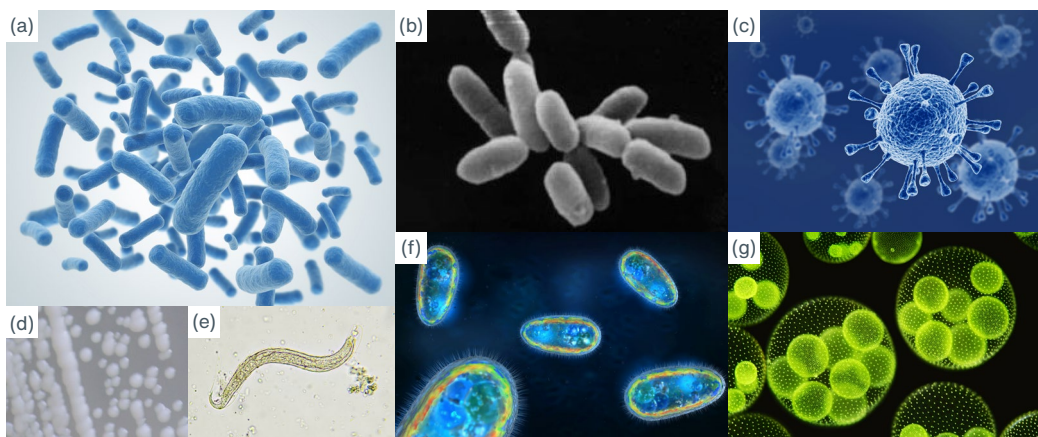
Figure 1
There are seven major types of microorganisms: (a) bacteria (e.g. *Mycobacterium tuberculosis*); (b) archaea (e.g. *Halobacterium salinarum*); (c) viruses (e.g. Herpes simplex virus); (d) fungi (e.g. *Saccharomyces cerevisiae*); (e) helminths (e.g. *Strongyloides stercoralis*) (f) protozoa (e.g. *Paramecium aurelia*); and (g) algae (e.g. *Volvox aureus*).

# The advantages of nanopore sequencing for microbial analysis

Genome assembly

In order to truly understand the diversity of microorganisms, it is imperative to create complete and fully characterised genomes. Repeat regions and structural variants have important functions in microorganisms, including the development and propagation of antimicrobial resistance and virulence factors[4,5,6]. However, it is widely recognised that the short-read sequencing technologies used to generate the majority of existing reference genomes struggle to resolve such regions[5,7] (**Figure 2**). Furthermore, these traditional sequencing technologies have been shown to exhibit GC bias, where sequences with low or high levels of GC content are underrepresented[8]. It is for these reasons that most reference genomes are incomplete, containing gaps where either sequencing or alignment is not possible.

In contrast to traditional sequencing platforms, which typically process short (<300 bp) DNA fragments,

> **With nanopore sequencing, ultra-long read lengths in excess of 4 Mb have been generated[9]**

nanopore sequencing technology processes the entire DNA (or RNA) fragment that is presented to the pore, 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[9]. Clearly, such long reads are more likely to span entire regions of repetitive DNA and structural variation, allowing more complete, gap-free genome assembly[7,10] (**Figure 3**). In addition, 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[8].
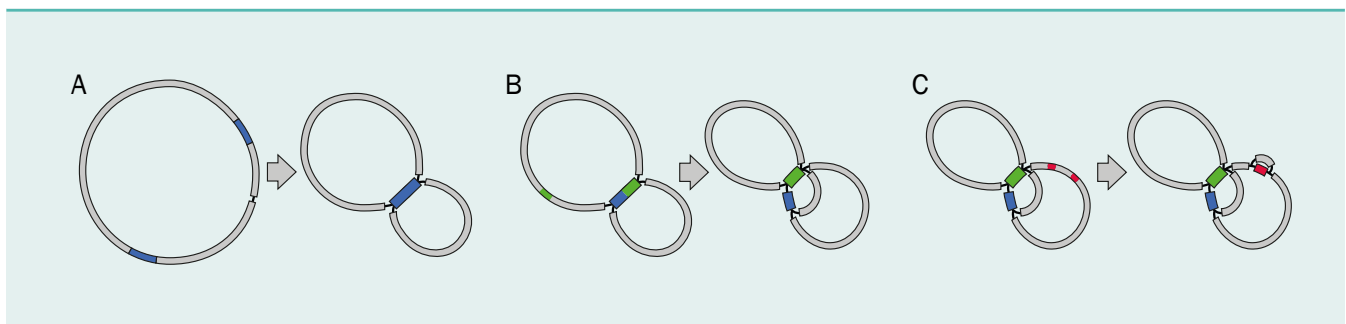
Figure 2
Schematic highlighting the challenge of correct genome assembly using short-read data. The higher the number of repeats (blue, green, and red regions), the more tangled the assembly graphs become, resulting in fragmentation of the assembly. Image courtesy of Ryan Wick, University of Melbourne, Australia.
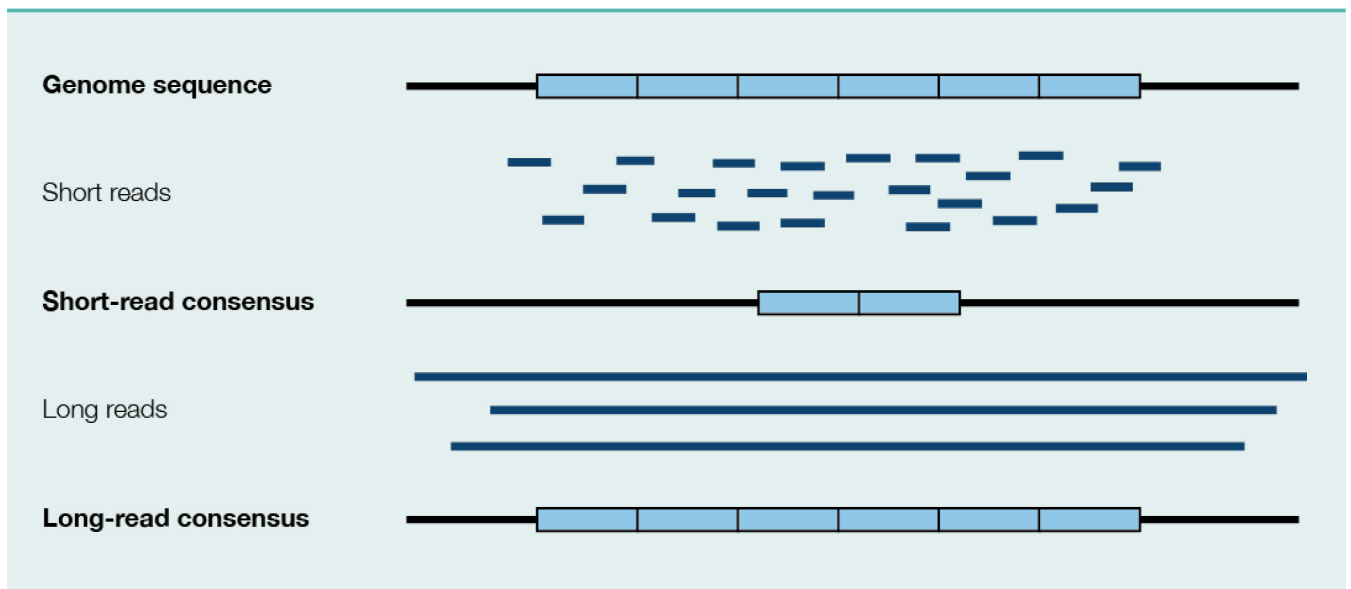
Figure 3
A schematic highlighting the advantages of long reads in *de novo* assembly of repetitive regions. Repeat and high-identity regions are often collapsed during the assembly process when using short-read sequencing technology. In contrast, long read lengths are more likely to incorporate the whole repetitive region (blue boxes) allowing more accurate assembly. Image adapted from Kellog[11].

Long sequencing reads offer a further advantage in that they provide greater overlap between reads, making it easier to assemble DNA fragments into their correct order. This can be visualised much like a jigsaw: the larger the pieces, the easier the puzzle.

> **Consensus accuracies in excess of Q50 (99.999%) have been generated using metagenomic nanopore sequencing at 20x sequencing depth, supporting the generation of highly accurate genome assemblies and functional studies[12]**

Using ultra-long nanopore reads, it is possible to sequence entire viral genomes in a single read, bypassing the requirement for assembly completely[13]. As sample preparation protocols develop, allowing even longer, intact DNA and RNA molecules to be generated, it is feasible that even larger genomes will be sequenced in single reads.

Metagenomics

The advent of metagenomic sequencing techniques, whereby all organisms in a given sample are studied at nucleotide resolution, has revolutionised the field of microbiology. Now, previously unculturable microorganisms can be analysed in their natural environment, providing new insights into microbial biology, ecology, and evolution. However, the limitations of traditional short-read sequencing technologies, such as GC bias and the inability to resolve repeat regions, are exacerbated when analysing mixed microbial samples[8,14]. According to Sereika *et al*., *'This is especially problematic in metagenome samples, in which related species or strains often contain long sequences of near-identical DNA'*[14] (see Case studies 1 and 3).

Long, amplification-free nanopore sequencing reads can span repeat regions and highly homologous regions of DNA from related species or strains, supporting the generation of high-quality metagenome-assembled genomes (MAGs) from complex, mixed microbial samples[14,15].

> *'Long-read Oxford Nanopore sequencing has democratized microbial genome sequencing and enables the recovery of highly contiguous microbial genomes from isolates or metagenomes'[14]*

A considerable obstacle associated with metagenomic analysis of pathogen samples is the high proportion of host DNA present in the sample. To overcome this challenge, researchers have successfully deployed a range of methodologies, including optimised sample lysis techniques, host DNA depletion, and enrichment of target microbial genomes; however, all of these approaches require some level of additional sample processing. Recently, researchers have been exploring the use of adaptive sampling — a unique capability of nanopore sequencing that utilises real-time analysis to selectively sequence or reject DNA molecules as they pass through a nanopore — as an alternative approach for metagenomic sample enrichment. In a recent study, Cheng *et al.*[16] utilised adaptive sampling on 11 clinical research samples to reject sequencing reads that matched the human reference genome. The team were able to increase the microbial sequence yield by at least 8-fold in all samples and, within 4.5 hours, accurately identify the pathogens and their associated resistance genes. Furthermore, five samples could be processed simultaneously on a single MinION™ Flow Cell, offering considerable time and cost savings.

Further sample multiplexing and the generation of high-quality MAGs from complex samples can be achieved using PromethION™ devices (**Figure 4**). Offering the flexibility of two independently operable high-output PromethION Flow Cells — each delivering up to five times the sequencing yield of a MinION Flow Cell — PromethION 2 devices bring affordable metagenomic analyses to every microbiology lab.

For more detailed information about metagenomics and microbiome analysis, download our Metagenomics white paper: nanoporetech.com/ resource-centre/metagenomics-white-paper.

Antimicrobial resistance profiling

According to the World Health Organisation: '*Antimicrobial resistance (AMR) is a global health emergency that will seriously jeopardise progress in modern medicine*'[17]. As such, it is vital that researchers can accurately and rapidly characterise pathogen samples to gain insights into the evolution and transmission of drug resistance and to further understand potential therapeutic intervention strategies.

Nanopore sequencing provides a streamlined approach for AMR profiling, both in the lab and in the field. The portable MinION and MinION Mk1C devices can be utilised anywhere, including resource-limited or remote environments, making them ideal for analysis at point of sampling (**Figure 4**). The benchtop GridION™ and PromethION devices provide flexible, on-demand sequencing suitable for higher-throughput requirements (**Figure 4**). 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**).

> **A significant advantage of nanopore sequencing is that it streams data in real time, allowing immediate analysis, drastically reducing time to result**

Unlike traditional sequencing technology that delivers all data at the end of a run, a significant advantage of nanopore sequencing is the facility for real-time data streaming and analysis. In addition to drastically reducing time to result, real-time analysis provides immediate validation of sample quality and verification that the correct sample has been collected. Furthermore, once sufficient data has been generated, flow cells with active pores remaining can be washed and reused — providing further experimental and cost efficiencies.

Oxford Nanopore has developed the EPI2ME™ ARMA analysis workflow for real-time microbial species identification and AMR profiling[18]. This workflow aligns sequencing reads against the Comprehensive Antibiotic Resistance Database (CARD) providing a detailed report highlighting read alignments that indicate resistance to a given antibiotic (**Figure 5**). With no requirement for in-depth bioinformatics expertise, the intuitive ARMA workflow allows any researcher to rapidly characterise AMR in their samples.

Nanopore sequencing has enabled researchers to reduce the time required to characterise AMR in the tuberculosis-causing organisms *Mycobacterium tuberculosis* and *Mycobacterium bovis* from multiple weeks, using traditional culture-based methods, to just 12 hours[19,20].

Using ultra-long reads and real-time analysis, nanopore sequencing has allowed rapid drug-resistance profiling for a range of pathogens, including bacteria[22,23,24], viruses[25,26], and fungi[27,28]. This has been achieved on both cultured isolates and uncultured metagenomic samples.

Figure 4
Oxford Nanopore sequencing platforms (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 and 2 Solo; and the high-throughput PromethION 24 and 48 (shown) platform, capable of delivering up to 7 and 14 Tb of data, respectively[*].

* 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.
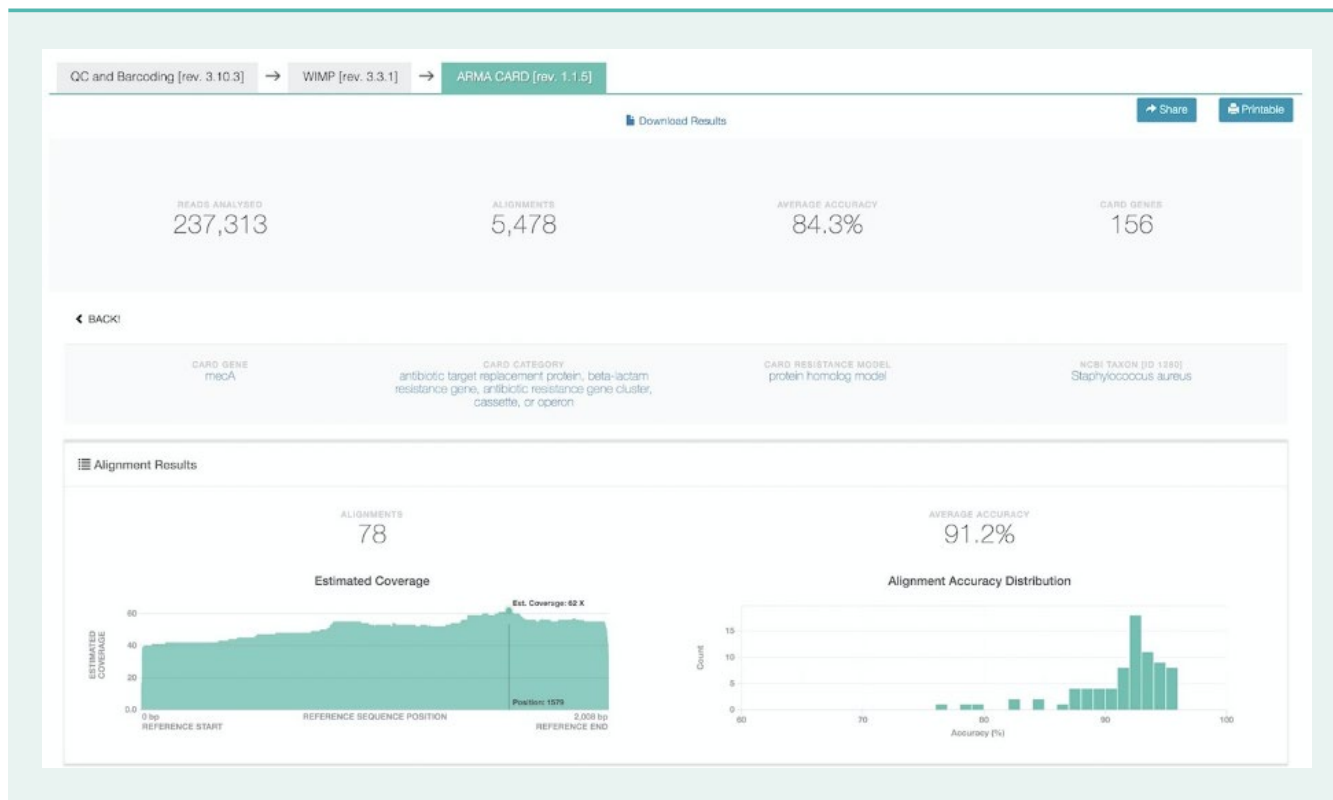
Figure 5
The EPI2ME ARMA workflow enables real-time microbial identification and AMR profiling from nanopore data.

Completing plasmid assemblies

In bacteria, AMR genes can be carried on the bacterial genome or on plasmids. Understanding the location of these genes provides more clarity on the mechanisms of AMR transfer between pathogens and is a vital consideration in epidemiological tracking and containment strategies[29,30,31]. Identifying plasmid-encoded resistance is especially important as it can be rapidly transferred throughout bacterial populations via horizontal gene transfer.

> *'The ability for Oxford Nanopore to sequence long fragments of DNA has significantly aided the assembly of bacterial genomes and plasmids'* [32]

Due to a high level of DNA repeats, which may also be present in the genome, it is particularly challenging to produce complete plasmid assemblies and distinguish them from genome sequence using short-read sequencing technology[31,33]. Long nanopore sequencing reads resolve these challenges, allowing the generation of complete and distinct plasmid and genome assemblies. Using nanopore sequencing, researchers have been able to rapidly depict the specific locations of resistance genes, giving more detailed insight into their transmission and evolution[29,30].

In an experiment by Li *et al.*, a single nanopore read from a carbapenem-resistant *Escherichia coli* strain was shown to span an entire plasmid of >90 kb in length[30].

Bacterial plasmid sizes are extremely diverse and can range from a few hundred bases to several megabases in length. With nanopore technology, read length is only limited by the size of the DNA fragment presented to the nanopore. As a result, it is theoretically possible to sequence even the largest plasmids in single reads. Such a scenario would remove the need for assembly, further streamlining the analysis workflow.

Plasmids are also a fundamental tool in molecular biology, and verification of the cloned plasmid sequence is essential to ensuring that no errors have arisen that may disrupt the expression of inserted genes. Sequencing is the gold-standard approach for clone verification; however, traditional technologies can be limited in regard to turnaround time, primer requirements, sequence bias, sample throughput, and, critically, may preclude the analysis of the whole plasmid construct. Nanopore sequencing has been shown to overcomes these challenges — providing a fast, secure, and cost-effective whole-plasmid sequencing workflow that can be easily implemented in any lab[35].

Find out more at nanoporetech.com/plasmid-verification.

Virulence

Virulence genes are often clustered into pathogenicity islands (PAIs) that can be either incorporated in the genome or located extra-chromosomally (i.e., on plasmids). As discussed for AMR genes, PAIs are also commonly flanked by repetitive insertion sequences that support their movement within and between species (**Figure 6**). The large size of PAIs (typically 10–200 kb[36]), coupled with their repetitive content, makes their accurate characterisation particularly difficult using short-read sequencing techniques[37,38]. Long nanopore sequencing reads can span entire PAIs, thereby providing authoritative characterisation and localisation of these regions.
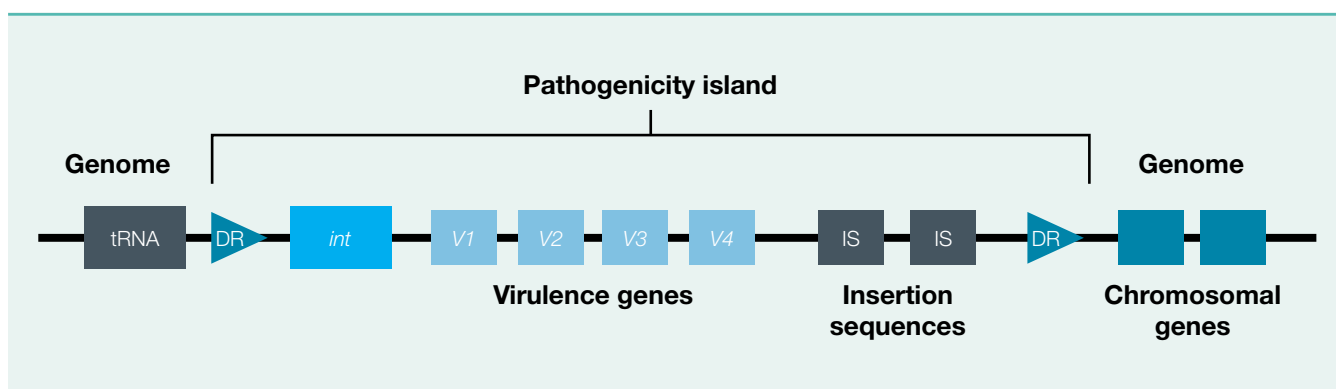
Figure 6
The large size and repetitive content of PAIs makes accurate assembly and location of these regions challenging using short-read sequencing technologies. Mobility genes, such as integrases (*int*), are frequently located at the beginning of the island, close to the tRNA locus or the respective attachment site. PAIs contain virulence genes (V1 to V4) and are frequently interspersed with mobility elements such as insertion elements/sequences (IS), which may be complete or partial. PAI boundaries are frequently flanked by direct repeats (DR), which are used for insertion and deletion processes. Image adapted from Schmidt and Hensel[39].

RNA genomes

RNA viruses lack the proofreading mechanisms typically found in organisms with DNA-based genomes and, as a result, exhibit much higher mutation rates. This high error rate confers an evolutionary advantage, allowing RNA viruses to adapt rapidly to evade the host immune response to infection and subsequent antiviral therapy. It is perhaps unsurprising, then, that many emerging human viral diseases are caused by RNA viruses (e.g. Ebola, influenza, COVID-19, severe acute respiratory syndrome [SARS], and Zika), adding more urgency to accurate genome analysis efforts for these organisms. Through analysing the genome, the genetic and geographic origin of the disease can be elucidated, and novel variants that enhance virulence, allow cross-species spread, or sustain human-to-human transmission can be identified (see Case study 2).

The sequencing of cDNA copies of RNA virus genomes has driven many important discoveries, but it is widely understood that the process of converting RNA to cDNA through reverse transcription and amplification can introduce bias[40,41].

> **'Much of what we know about RNA biology is seen through a cDNA lens'[42]**

Nanopore technology uniquely enables direct sequencing of native RNA molecules, without the requirement for amplification or reverse transcription — removing two potential sources of sequencing bias (**Figure 7**). Utilising nanopore technology, researchers have been able to sequence the complete influenza A RNA genome, with each of the eight genome segments represented by full-length reads[42]. Furthermore, the streamlined direct RNA sequencing workflow allowed the virus to be sequenced in a single day rather than several days as required for the standard short-read technology surveillance pipeline[42].
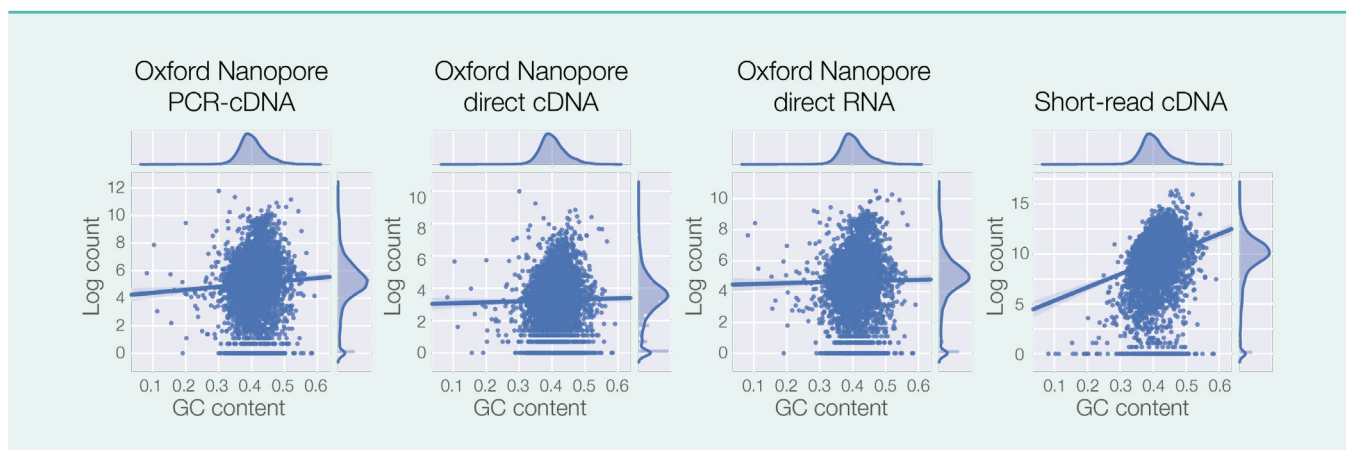
Figure 7
Sequencing workflows that incorporate amplification are vulnerable to sequence-specific biases. Yeast transcriptome libraries were prepared using three nanopore sequencing techniques (PCR-cDNA, direct cDNA, and direct RNA) and a typical short-read cDNA method. In all cases, GC bias in the nanopore data sets was lower than in the short-read data set.

Long, full-length transcripts

Long nanopore sequencing reads enable unambiguous, full-length analysis of RNA transcripts, capturing the full complexity of microbial transcriptomes. Full-length reads have been shown to enable the accurate identification and quantification of transcripts and, in the case of eukaryotes, transcript isoforms. They have also provided new insights into virus-host interactions[43,44], novel viral splicing mechanisms[45,46], and expression of bacterial resistance genes[32].

> *'Our work highlights how long-read sequencing technologies can reveal further complexity within viral transcriptomes'[46]*

Oxford Nanopore offers three streamlined RNA sequencing kits: PCR-cDNA, direct cDNA, and direct RNA, each combining full-length reads with low bias (**Figure 7**).

For more information, download the RNA sequencing white paper at nanoporetech.com/resource-centre/ rna-sequencing.

Modified base detection

Modified bases (e.g. 5-methylcytosine, adenine N6-methylation) have been found in virtually all organisms studied. The specific roles of many of these modified bases are still to be fully elucidated; however, they are known to influence gene expression and, in prokaryotes, protect against bacteriophages[47]. Researchers have also proposed a role for modified bases in antimicrobial resistance[41].

> *'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'[48]*

The requirement for nucleic acid amplification in traditional short-read sequencing technology erases these base modifications, meaning they cannot be detected without additional time-consuming and often inefficient sample processing methods[41].

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 wide a number of modified bases, including pseudouridine[41,49], N6-methyladenosine (m6A)[40], 4-methylcytosine (4mC)[50], 5-methylcytosine (5mC)[40,42], and 7-methylguanosine (m7G)[41].

High-quality, low-cost, nanopore-only bacterial genome sequences

To obtain reference-quality bacterial genome assemblies, data is often used from the sequencing of either pure cultures or metagenomic samples. Short-read sequencing has been the technology of choice for this application in previous years but has limited ability to resolve repetitive sequences that are longer than the library insert size. Consequently, technology capable of producing long sequencing reads, including the Oxford Nanopore platform, has '*recently emerged as the choice*' for assembling genomes derived from such samples[14].

Professor Albertson and colleagues, based at Aalborg University in Denmark, investigated whether nanopore sequencing data alone could be used to obtain reference-quality bacterial genome assemblies[14]. Their work noted that, in the past, there has been a preference to use either short-read or reference polishing of nanopore data to obtain near-complete microbial genome assemblies, yet this is an undesirable option as it adds cost and complexity[14].

The team evaluated the performance of R9 and the more recent R10 nanopore chemistry in bacterial genome assembly, obtaining sequence data derived from 'pure cultures' (in this case, a mock community) and an activated sludge sample.
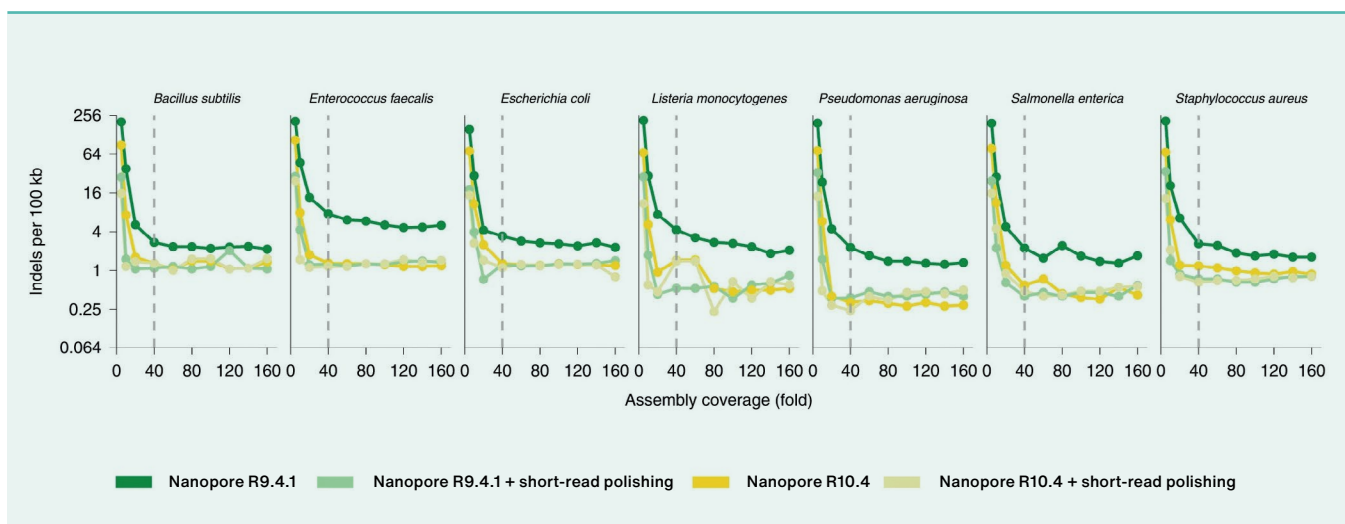
Figure 8

Indels observed per 100 kb in the *de novo* bacterial isolate assemblies, at different depths of coverage, with and without short-read polishing. The authors noted that short-read polishing of nanopore data obtained using R10.4 chemistry provided no significant improvement in assembly quality. Image adapted from Sereika *et al*.[14] and available under Creative Commons license (creativecommons.org/licenses/by/4.0).

They introduced the term 'near-finished' genome to indicate the generation of a high-quality genome assembled with only long nanopore reads, for which the application of short-read polishing would not significantly improve the consensus sequence. They found that R10.4 data alone could generate near-finished bacterial genomes, without polishing (**Figure 8**). The depth of coverage required to achieve this was approximately 40-fold. To assess performance on metagenomic genome assembly, the team sequenced a sample of activated sludge; a similar conclusion was made — that R10.4 chemistry enabled the generation of near-finished microbial genomes, without short-read polishing[14].

> *'Oxford Nanopore R10.4 enables the generation of near-finished microbial genomes from pure cultures or metagenomes at coverages of 40-fold without short-read polishing'[14]*

A notoriously challenging bacterial genome to sequence and assemble is that of *Mycobacterium tuberculosis*. *M. tuberculosis* is the pathogen responsible for tuberculosis (TB), which remains one of the deadliest infectious diseases, with 1.5 million human deaths attributed to TB in 2020[51]. Drug-resistant *M. tuberculosis* is a particularly significant threat for effective TB control[51,52]. Genome sequencing of the pathogen has gained traction in recent years for both clinical research and epidemiological investigations. Such efforts have provided valuable insights into circulating strains, including mutations underlying drug resistance

and virulence, and the dynamics of person-to-person transmission — conferring high-resolution analyses when compared with culture-based phenotyping or targeted sequencing assays[52].

Previously, short-read sequencing technology was typically used to investigate the genetic basis of resistance and the genomics underpinning TB transmission. However, the genome of *M. tuberculosis* is challenging to resolve with short reads due to its high GC content and repetitive nature — including the highly variable and GC-rich *pe/ppe* genes associated with drug resistance, which are often excluded from analysis due to difficulties in accurately mapping these regions to the genome when using short reads. Furthermore, the high capital cost and centralisation associated with these sequencing platforms has limited access to whole-genome analysis in many areas with a high TB burden and lower income[52,53].

In contrast, the Oxford Nanopore platform can produce sequencing reads of any length, and a scalable range of devices is available, including portable options suitable for *in situ* sequencing; the technology has therefore been recognised as a *'promising platform for cost-effective application'* to TB genome analysis[52]. However, few studies have investigated the performance of nanopore sequencing for *M. tuberculosis* genome analysis for drug susceptibility prediction or outbreak investigation.

In light of this, Gómez-González *et al.* and Hall *et al.* compared the performance of Oxford Nanopore and short-read sequencing platforms for these applications[52,53]. Gómez-González *et al.* sequenced 10 *M. tuberculosis* clinical research isolates with both nanopore and short-read technology, obtaining 93.6-fold short-read and 72.2-fold nanopore depth of coverage, after mapping. The team highlighted the improved coverage of long nanopore reads in repetitive regions where short reads failed to accurately align. As expected, a higher number of large variants were detected with long nanopore reads, (median 81 vs. 24, across the isolates); regarding single nucleotide polymorphisms (SNPs), for all sample pairs, >99% of SNPs identified were called in both samples, with few platform discrepancies. All lineage predictions were identical between the two platforms (**Figure 9a**); however, looking specifically at the nanopore data, as the *pe/ppe* gene regions were successfully resolved with long nanopore reads, SNPs could also be incorporated from these regions for lineage analysis, which led to an improved resolution that '*would be of special interest in outbreak settings, where transmission analysis of closely related isolates can be potentially better established*' (**Figure 9c**). They also suggested that the ability to cover repetitive regions with long reads could contribute a better understanding of drug-resistance mechanisms in *M. tuberculosis*[52].

Hall *et al.* aimed to establish whether nanopore sequence data could be used to reproduce equivalent transmission clusters and drug susceptibility profiles to those generated with short-read data[53]. To investigate this, the team obtained matched nanopore and short-read data from 151 isolates. The study found that isolate clustering was the same between the two platforms, and in terms of genotyping resistance-associated SNPs and INDELs, they obtained near-identical results, with a concordance of >99.99% between the two technologies[53].

‘Our analysis shows that it is now possible to obtain high-precision SNP calls in M. tuberculosis with current nanopore data’[53]
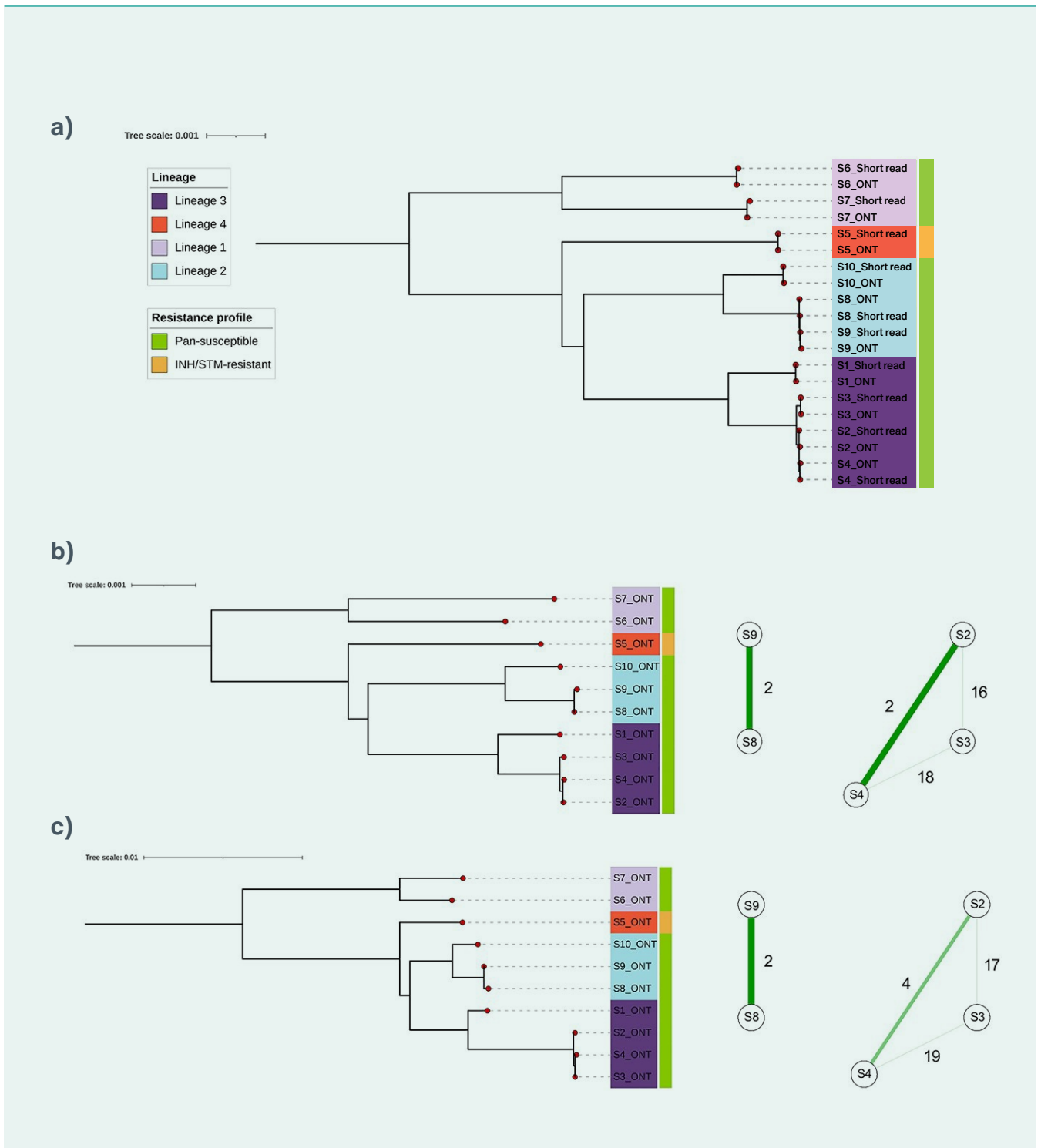
Figure 9

Phylogenetic trees representing the branching order for the *M. tuberculosis* clinical research isolates studied, showing equal branch lengths for the 10 pairs of sequenced isolates **(a)**, and only the 10 nanopore-sequenced isolates **(b)**. **(c)** Phylogenetic tree of nanopore-sequenced samples using the 3,955 SNPs as well as the 568 sites located in the *pe/ppe* genes, achieving a deeper separation of transmission clusters. Drug resistance profiles (obtained by phenotyping drug susceptibility testing) are shown by coloured strip labels alongside each tree; INH = Isoniazid, STR = Streptomycin. Image adapted from Gómez-González *et al.*[52] and available under Creative Commons license (creativecommons.org/licenses/by/4.0).

CASE STUDY 2

One size fits all: development of a simple workflow to characterise any pathogen

Successful surveillance of pathogen outbreaks requires a rapid, scalable, and cost-effective response. To overcome the challenges of monitoring pathogen outbreaks, such as generating real-time epidemiological information, researchers are utilising the benefits of portable, real-time nanopore sequencing technology to perform successful surveillance at sample source.

Genomic epidemiology, the study of how variants in the genomes of pathogens, or their hosts, influence health and disease, aims to track pathogen transmission, spread, and evolution. Sequence information also enables the identification of drug resistance factors. Nanopore technology has played a pivotal role in global pathogen surveillance, from the Ebola outbreak in western Africa in 2013-16 to the present-day COVID-19 pandemic, with greater than one million SARS-CoV-2 genomes sequenced using nanopore technology across over 85 countries[54]. Many of the known RNA viruses, such as SARS, influenza, and HIV, are fast evolving, which means they continually accumulate changes in their genome. Sequence data can guide important control measures, but only if the results are generated quickly enough to inform interventions. Conventional sequencing technologies are typically not accessible in low-resource settings, where requirements, such as continuous power, lab infrastructure, and trained personnel are often not available within outbreak areas, leading to practical difficulties transporting samples to remote sequencing facilities. With this in mind, Josh Quick and his team based at the University of Birmingham, took advantage of the accessibility, affordability, and portability of nanopore sequencing, to successfully perform scalable, field-based genomic surveillance of the Ebola virus in western Africa. Using portable

equipment, including the MinION, which they termed a 'lab-in-a-suitcase'[55], the team highlighted how easily nanopore technology can be deployed to '*remote and resource-limited locations*' to monitor pathogen outbreaks with rapid turnaround times, '*with the sequencing process taking as little as 15–60 min*'[55].

The ARTIC network has since been set up to expand international viral surveillance, enabling worldwide collaboration in sequencing and analysing pathogen outbreaks[56]. Leveraging the portability and streamlined sample preparation workflows of the MinION, they aimed to target a wide range of DNA and RNA viruses with a simple workflow that could be used worldwide. The approach needed to be applicable, scalable, and rapidly deployable. Sequencing viral genomes directly from clinical research samples can be challenging for viruses with a low viral load, such as the Zika virus. The workflow developed by the ARTIC network used a PCR enrichment approach, suitable for samples containing as few as 50 genome copies per reaction[57]. PCR can provide both target enrichment and amplification in a single step. For complete coverage, a tiled amplicon scheme was devised, and to reduce time, complexity, and cost, a multiplexed sequencing approach was used. The end-to-end workflows allow any researcher to amplify and sequence high- and low-abundance

Virus	Genome size (bp)	Amplicon length (bp)	Number of amplicons
Ebola[55]	~20,000	2,000	11
Zika[57]	~11,000	400	35
Monkeypox[59]	~200,000	2,500	88
SARS-CoV-2 (ARTIC)[60]	~30,000	400	98
SARS-CoV-2 (Midnight)[61]	~30,000	1,200	29
Chikungunya[62]	~12,000	300	44
Yellow fever[63]	~10,000	500	27
Dengue[64]	~11,000	1,000	13
Africa Swine Fever[65]	~190,000	7,000	32
Avian influenza[66]	~13,500	900–2,300	8
Rabies[67]	~12,000	400	39
Tick-borne encephalitis[68]	~10,000	450	37
Hepatitis C[69]	~9,600	~9,000	1
HIV-1[70]	~9,600	5,000–9,000	1–2
Adenovirus-F41[71]	~34,000	1,200	92

Table 1
Examples of tiled amplicon protocols applied to viral genomes with nanopore sequencing, detailing genome size, amplicon length, and number of amplicons to complete whole-genome sequencing.

viruses directly from clinical research samples. Users can design their own primer schemes for their virus of interest using the web-based primer design tool, primal scheme[58].

Due to easy adaptability for the amplification and sequencing of different viruses, tiled PCR approaches using nanopore sequencing have been widely adopted (**Table 1**).
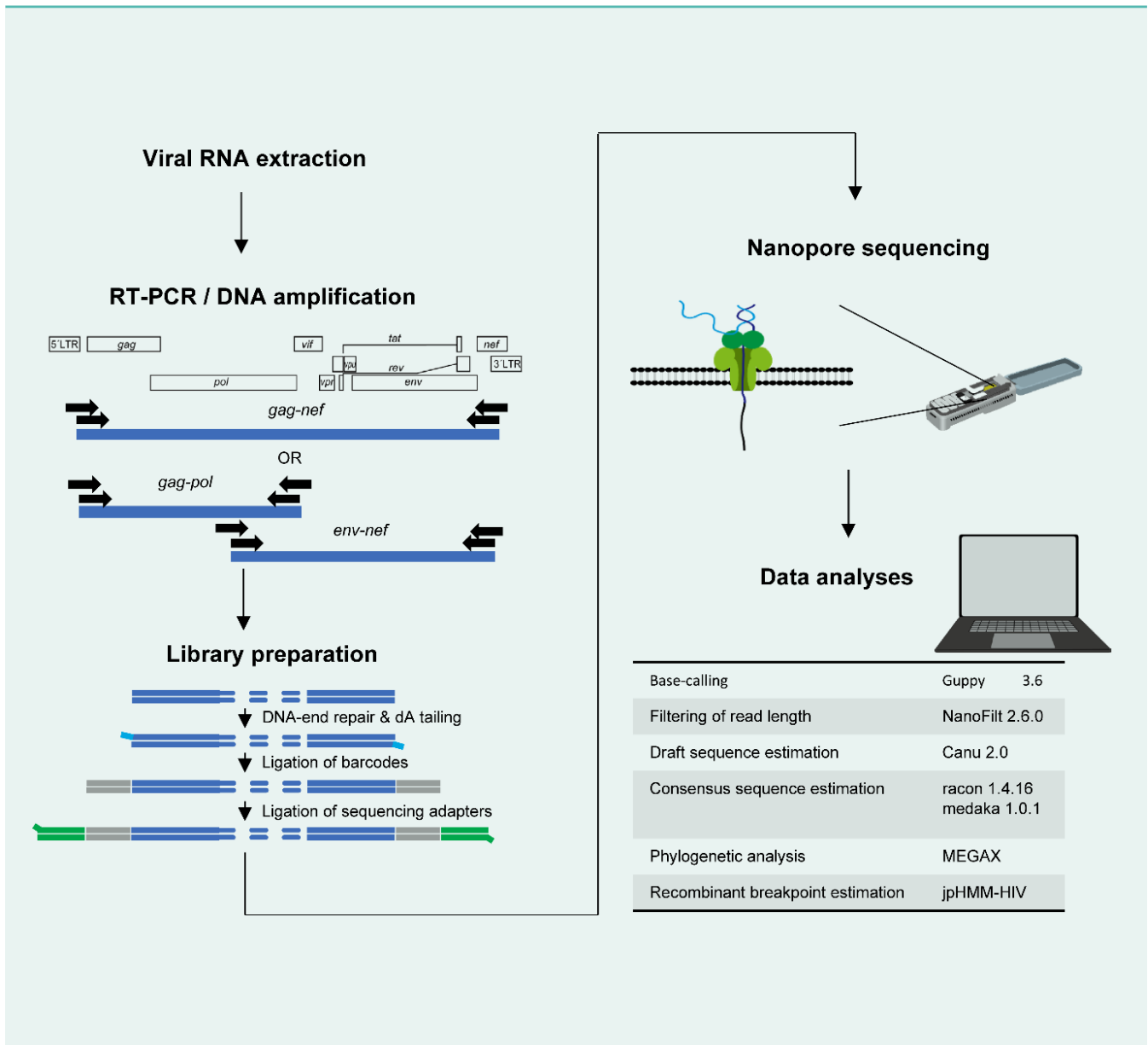
Figure 10
Nanopore sequencing protocol applied to the near full-length HIV-1 genome. Image from Mori *et al.*[70] and available under Creative Commons license (creativecommons.org/licenses/by/4.0).

With nanopore sequencing, read length is equal to fragment length, enabling tiled PCR and sequencing of viral genomes with long amplicons. Mori and coworkers[70] generated near full-length HIV genomes and obtained distinctive genetic information for the highly genetically diverse recombinant forms (RFs) of HIV-1 using nanopore sequencing of amplicons (**Figure 10**). The team highlighted that drug resistance-associated mutations are located outside the target sequence regions of the traditional HIV-1 genotyping assay and suggested that long sequencing reads are now required for full analysis.

> **'Our new nanopore sequencing platform is applicable to identify the full-length HIV-1 genome structure of intersubtype RFs as well as dual infection heterologous HIV-1'**[70]

Outbreaks of mosquito-borne viruses, such as Zika, chikungunya, yellow fever, and dengue, have all been characterised using tiled amplicon workflows with nanopore sequencing, and together with broader epidemiology data could potentially be used to inform vaccination strategies. When

vaccine levels are low, critical decisions on geographic areas to be targeted require a detailed understanding of the spatial spread of the virus and predictions of where it is most likely to spread.

Nanopore workflows have also been adapted to study viruses known to affect animals, such as rabies, avian influenza, and African swine fever virus (ASFV). ASFV is highly contagious and has a mortality of up to 100% in domestic pigs, which severely impacts pork production and local economies. There have been major outbreaks throughout the world, making pathogen surveillance essential. However, its large genome size (170–190 kb), ability to acquire large deletions and insertions, and the presence of highly mutagenic hypervariable regions make sequencing the ASFV genome very challenging. Using the MinION, Amanda Warr and colleagues[65] developed a tiled-amplicon approach to sequence ASFV that could enable near-live monitoring of outbreak situations with rapid turnaround times, and no challenging transportation of samples. Amanda described how '*multiplexing,... washing and reuse of the most expensive component of sequencing, the flow cells'* allowed for lower cost sequencing than other methods and the long reads improved assembly potential, particularly of highly repetitive genomes[65]. Amplification of the genome in 32 fragments of 7 kb, with 1 kb overlaps, generated near-complete genomes.

Studying virus populations within infected hosts can provide essential information in understanding virus-host interactions and new approaches to outbreak response. During the recent SARS-CoV-2 pandemic, Oxford Nanopore and the ARTIC network were quick to refine the targeted amplicon approach with the Midnight protocol[8]. With the ability to multiplex up to 96 samples, both time and costs have been substantially reduced. The Midnight protocol has become a popular method within the public health community because hands-on time is minimal, workflows are simple, and researchers can go from RNA to real-time sequencing within one working day.

Oxford Nanopore is proud to have worked with the scientific community in pathogen surveillance and, in collaboration with the ARTIC network and other research groups, continuously develops its protocols, kits, and bioinformatic pipelines. The cost-effective workflows and scalability to large numbers of samples during outbreak surges suggests the multiplex, tiled PCR approach with Oxford Nanopore technology will remain an essential method in pathogen surveillance.

# Metagenomic analysis of microbial communities in permafrost thaw

Taking advantage of the long sequencing reads generated by nanopore technology, Devin Drown and his team at the University of Alaska Fairbanks, USA, have been researching how the thawing of permafrost may affect soil microbial communities[72]. Permafrost is a permanently frozen layer of soil, gravel, and sand, on or under the Earth's surface.

To be considered permafrost, the area must have been continuously frozen for at least two consecutive years but, in most cases, it has been frozen for hundreds or thousands of years. Permafrost covers large regions of the Earth, in areas with cold climates at high latitudes or altitudes. Sitting on top of the permafrost is a thin layer of soil, known as the active layer. Devin and his team wanted to understand how permafrost thawing, caused by climate change, affects the microbial communities in the active layer and how, in turn, this affects the crops which grow within the soil. Interactions between microbes and plants are critical for the acquisition and cycling of nutrients, and even small changes in soil microbial communities can disrupt plant-microbial interactions[73].

Based in Alaska, the team applied nanopore metagenomic sequencing to the study of microbial communities in soils associated with differing permafrost conditions that can be found in boreal forests, also known as snow forests. These forests are home to many plants that are integral to the diets of native Alaskan communities. Using metagenomics means microorganisms do not need to be cultured. Devin and his team chose nanopore technology for a rapid, and thorough characterisation of the complex microbial communities in the active layers. Long, PCR-free nanopore sequencing reads enable access to regions that are difficult to sequence with traditional short-read sequencing technologies, facilitating the assembly of accurate microbial genomes from complex communities. Native DNA can be sequenced in real time, enabling streamlined library preparation and rapid turnaround times.

Using plots of artificially induced permafrost thaw, created over 60 years ago at the Fairbanks Permafrost Experiment Station in Alaska, the team conducted plant growth experiments with five plant species found in boreal forests: *Vaccinium vitis-idaea* (low-bush cranberry), *Vaccinium uliginosum* (bog blueberry), *Picea mariana* (black spruce), *Ledum groenlandicum* (Labrador tea), and *Chamerion angustifolium* (fireweed). The
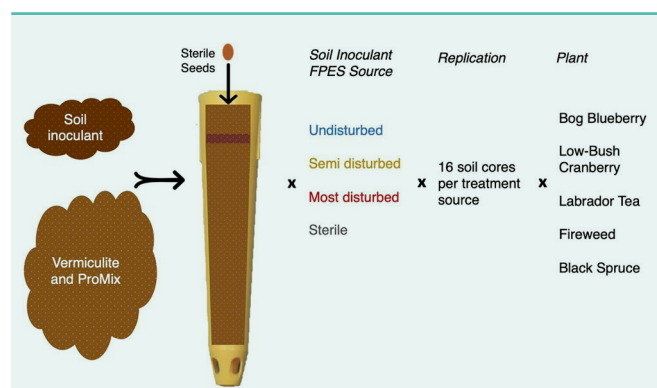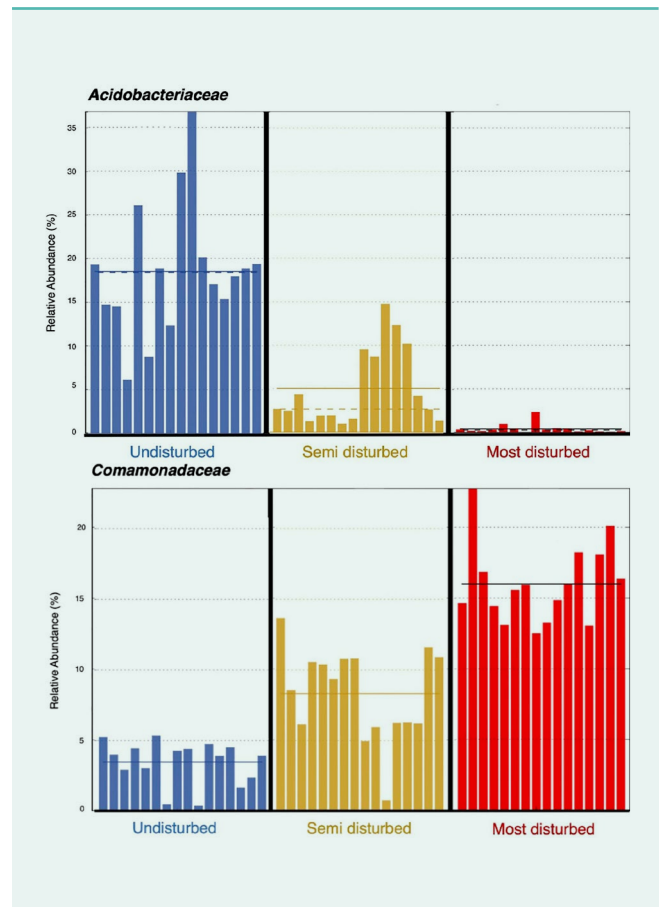

Figure 11
Experimental design of the plant growth experiment. Image from Seitz *et al.*[72] and available under Creative Commons license (creativecommons.org/licenses/by/4.0).

plants were grown in soils with differing degrees of permafrost thaw, termed 'undisturbed', 'semi-disturbed', and 'most disturbed' (**Figure 11**).

Most of the plants inoculated with microbial communities from the active layer above thawing permafrost showed decreased productivity compared to plants inoculated with microbes from the undisturbed active layer. To analyse the microbial communities within the differing active layers, the team used the Ligation Sequencing Kit to sequence 48 metagenomes across four MinION Flow Cells. The combined datasets showed a mean read length of 2,594 bp, and a read length N50 of 5,531 bp. The reads were then processed using Kraken[74] and Bracken[75] to detect taxa and estimate abundance, respectively. The team identified 24 bacterial phyla within the microbial communities. Visualisation of the data showed that the microbial community populations differed between the different soil types, with the largest difference observed between the undisturbed and most disturbed active soil layers.

Beneficial soil microbes are known to enhance nutrient availability to plants, enabling increased plant growth and productivity[73]. The team analysed biomarkers associated with healthy plant growth and found microbes that have been positively correlated with plant productivity, such as *Acidobacteriaceae* and *Bacillales,* were highly represented in undisturbed soil samples compared with most disturbed soil samples. In contrast, members of the *Comamonadaceae* family, which are known to exhibit pathogenic effects on a variety of plants, were more abundant in the most disturbed soils compared with undisturbed soils (**Figure 12**). The team hypothesised that the known

Figure 12
Relative abundances of bacterial families in the different soil types. Solid lines represent mean relative abundance. Image from Seitz *et al.*[72] and available under Creative Commons license (creativecommons.org/licenses/by/4.0).

plant pathogens found in the most disturbed soil led to reduced plant growth, due to a disruption in nutrient cycling and direct alterations to the plant rhizosphere, the area around a plant root.

This environmental metagenomics study formalises the links between climate change, thawing permafrost, changing microbial communities, plant health, and broader community health[73].

Summary

Microbes are ubiquitous and form an essential part of all ecosystems. To enhance our understanding of health and the environment, it is vital to fully characterise these organisms. While the advent of high-throughput, short-read sequencing technologies has improved our ability to detect and analyse microbes, the inherent limitations of these methods present significant challenges to complete and rapid genome and transcriptome characterisation.

The long, real-time, and direct sequencing reads provided by nanopore technology overcome these challenges, delivering unprecedented insight into microorganisms — allowing complete genome assembly, real-time species identification, AMR and virulence profiling, modified base detection, and unambiguous transcriptome studies.

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 sample preparation and data analysis workflows provides rapid access to results for time-critical applications, such as outbreak tracking, and AMR, and virulence profiling.

A range of sequencing platforms is available, suitable for all microbial research applications and sample throughput requirements (**Table 2**; **Figure 4**).

	Flongle	MinION & MinION Mk1C	GridION Mk1	PromethION 2/2 Solo	PromethION 24/48
Read length	Fragment length = read length. Longest read now >4 Mb[9]				
Run time	1 min – 16 hrs	1 min – 72 hrs	1 min – 72 hrs	1 min – 72 hrs	1 min – 72 hrs
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 7Tb /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 microbiology applications — from complete assembly of microbial genomes to real-time AMR characterisation, in-field outbreak tracking, and transcriptomics. Data correct at time of print. Visit www.nanoporetech.com for the latest information.

For more information about utilising nanopore technology for microbial analysis, visit www.nanoporetech.com/applications/microbiology.

References

1. Schloss, P.D., Girard, R.A., Martin, T., Edwards, J., and Thrash, J.C. Status of the archaeal and bacterial census: an update. mBio. 7(3) (2016).

2. Locey, K.J. and Lennon, J.T. Scaling laws predict global microbial diversity. Proc. Natl. Acad. Sci. USA 113(21):5970-5 (2016).

3. NCBI National Center for Biotechnology Information. Genomes information by organism. Available at: https://www.ncbi.nlm.nih.gov/genome/browse/#!/overview [Accessed: 16 August 2022]

4. Land, M. et al. Insights from 20 years of bacterial genome sequencing. Funct. Integr. Genomics 15(2): 141–161 (2015).

5. Abrahams, J.S. et al. Duplications drive diversity in *Bordetella pertussis* on an underestimated scale. bioRxiv 937284 (2020).

6. Duc Cao, M. et al. Scaffolding and completing genome assemblies in real-time with nanopore sequencing. Nat. Commun. 8 (2017).

7. Díaz-Viraqué, F. et al. Nanopore sequencing significantly improves genome assembly of the protozoan parasite *Trypanosoma cruzi*. Genome Biol. Evol. 11(7):1952–1957 (2019).

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

9. Oxford Nanopore Technologies. Ultra-Long DNA Sequencing Kit. Available at: https://store.nanoporetech.com/ultra-long-dna-sequencing-kit.html [Accessed: 16 August 2022]

10. 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).

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

12. Oxford Nanopore Technologies. Accuracy. [online]. Available at: https://nanoporetech.com/accuracy [Accessed: 26 August 2022]

13. Beaulaurier, J. et al. Assembly-free single-molecule sequencing recovers complete virus genomes from natural microbial communities. Genome Res. 30(3):437–446 (2020).

14. 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).

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

16. Cheng, H. et al. An ultra-sensitive bacterial pathogen and antimicrobial resistance diagnosis workflow using Oxford Nanopore adaptive sampling sequencing method. medRxiv 22277093 (2022).

17. World Health Organisation. The world is running out of antibiotics, WHO report confirms. [online] Available at: https://apps.who.int/mediacentre/news/releases/2017/running-out-antibiotics/en/index.html [Accessed: 16 August 2022]

18. Oxford Nanopore Technologies. EPI2ME ARMA workflow: real-time antimicrobial resistance profiling. [online] Available at: https://nanoporetech.com/resource-centre/epi2me-arma-workflow-real-time-antimicrobial-resistance-profiling [Accessed: 16 August 2022]

19. Zhao, K et al. Rapid identification of drug-resistant tuberculosis genes using direct PCR amplification and Oxford Nanopore technology sequencing. Can. J. Infect. Dis. Med. Microbiol. 7588033 (2022).

20. Votintseva, A.A. et al. Same-day diagnostic and surveillance data for tuberculosis via whole genome sequencing of direct respiratory samples. J. Clin. Microbiol. 55(5):1285-1298 (2017).

21. Tafess, K. et al. Targeted-sequencing workflows for comprehensive drug resistance profiling of *Mycobacterium tuberculosis* cultures using two commercial sequencing platforms: comparison of analytical and diagnostic performance, turnaround time, and cost. Clinical chemistry, 66(6), 809–820 (2020).

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

23. Leggett, R.M. et al. Rapid MinION profiling of preterm microbiota and antimicrobial-resistant pathogens. Nat. Microbiol. 5, 430-442 (2020).

24. Břinda, K. et al. Rapid inference of antibiotic resistance and susceptibility by genomic neighbour typing Nat. Microbiol. 5(3) 455-464 (2020).

25. Lewandowski, K. et al. Metagenomic nanopore sequencing of influenza virus direct from clinical respiratory samples. J. Clin. Microbiol. 58(1) (2019).

26. Saranathan, R. et al. Capturing structural variants of herpes simplex virus genome in full length by Oxford Nanopore sequencing. Microbiology Spectr. e0228522 (2022).

27. Allen, J.L. Full chromosome assembly of symbiotic fungal genomes from complex metagenomics samples using nanopore sequencing. Presentation. Available at: https://nanoporetech. com/resource-centre/video/ lc21/full-chromosome-assembly-symbiotic-fungal-genomes-complex-metagenomics-samples. [Accessed: 16 August 2022]

28. Wang, H. et al. High-quality genome resource of *Clonostachys rosea* strain CanS41 by Oxford Nanopore long-read sequencing. Plant disease, 105(8), 2231–2234 (2021).

29. Wick, R.R., Judd, L.M., Gorrie, C.L. and Holt, K.E. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb. Genom. 3(10) (2017).

30. Li, R. et al. Efficient generation of complete sequences of MDR-encoding plasmids by rapid assembly of MinION barcoding sequencing data. GigaScience gix132 (2018).

31. Lemon, J.K, Khil, P.P, Frank, K.M. and Dekker, J.P. Rapid nanopore sequencing of plasmids and resistance gene detection in clinical Isolates. J. Clin. Microbiol. 55 (2017).

32. Pitt, M.E. et al. Evaluating the genome and resistome of extensively drug-resistant *Klebsiella pneumoniae* using native DNA and RNA nanopore sequencing. GigaScience 9(2), giaa002 (2020).

33. Arredondo-Alonso, S., Willems, R.J., van Schaik, W. and Schürch, A.C. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. Microb. Genom. 3(10) (2017).

34. U.S. National Library of Medicine. Genome information by organism. [online] Available at: https://www.ncbi.nlm.nih.gov/genome/browse/#!/plasmids/ [Accessed: 16 August 2022]

35. Liu, J. Cloning validation with the Oxford Nanopore pipeline in AAV research. Presentation. Available at: https://uk.vwr.com/cms/nanophore_plasmid_sequencing [Accessed: 26 August 2022]

36. Hacker, J. and Kaper, J.B. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. 54:641-79 (2000).

37. González-Escalona, N. et al. Nanopore sequencing for fast determination of plasmids, phages, virulence markers, and antimicrobial resistance genes in Shiga toxin-producing *Escherichia coli*. PLoS One. 14(7):e0220494 (2019).

38. Muthuirulandi Sethuvel, D.P. et al. Complete genome analysis of clinical *Shigella* strains reveals plasmid pSS1653 with resistance determinants: a triumph of hybrid approach. Gut Pathog. 11:55 (2019).

39. Schmidt, H. and Hensel, M. Pathogenicity islands in bacterial pathogenesis. Clin. Microbiol. Rev. 17(1): 14–56 (2004).

40. Garalde, D. R. et al. Highly parallel direct RNA sequencing on an array of nanopores. Nat. Methods 15(3):201– 206 (2018).

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

42. Keller, M. Direct RNA sequencing of influenza viral RNA using the MinION. Presentation. Available at: https://vimeo.com/250295802 [Accessed: 16 August 2022]

43. Tombácz, D. et al. Long-read assays shed new light on the transcriptome complexity of a viral pathogen and on virus-host interaction. Sci. Rep. 10, 13822 (2020).

44. Yang, S.L. et al. Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions. Nat. Commun. 12, 5113 (2021).

45. Price, A.M. et al. Direct RNA sequencing reveals m6A modifications on adenovirus RNA are necessary for efficient splicing. Nat. Commun. 11, 6016 (2020).

46. Price. A.M. et al. Novel splicing and open reading frames revealed by long-read direct RNA sequencing of adenovirus transcripts. bioRxiv 876037 (2019).

47. Weigele, P. and Raleigh, E.A. Biosynthesis and function of modified bases in bacteria and their viruses. Chem. Rev. 116(20):12655-12687 (2016).

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

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

50. Tourancheau, A. et al. Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. Nat. Methods 18, 491–498 (2021).

51. WHO. Tuberculosis. [online] Available at: https://www.who.int/news-room/fact-sheets/detail/tuberculosis. [Accessed: 23rd August 2022]

52. Gómez-González, P.J. et al. Portable sequencing of *Mycobacterium tuberculosis* for clinical and epidemiological applications. Briefings in Bioinformatics. bbac256, https://doi.org/10.1093/bib/bbac256 (2022).

53. Hall, M.B. et al. Nanopore sequencing for *Mycobacterium tuberculosis* drug susceptibility testing and outbreak investigation. medRxiv. DOI: https://doi.org/10.1101/2022.03.04.22271870 (2022).

54. GISAID. Home. [online] Available at: http://gisaid.org [Accessed: 18 August 2022]

55. Quick, J. et al. Real-time, portable genome sequencing for Ebola surveillance. Nature. 11;530(7589):228-232 (2016).

56. ARTIC network. Real-time molecular epidemiology for outbreak response. [online] Available at: http://artic.network [Accessed: 18 August 2022]

57. Quick, J. et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat. Protoc. 12(6):1261-1276 (2017).

58. Primalscheme. Primer panels for multiples PCR. [online] Available at: http://primalscheme.com [Accessed: 18 August 2022]

59. Welkers, M., Jonges, M., and van den Ouden, A. Monkeypox virus whole genome sequencing using combination of NextGenPCR and Oxford Nanopore. [online] Available at: https://www.protocols.io/view/monkeypox-virus-whole-genome-sequencing-using-comb-ccc7sszn.html [Accessed 18 August 2022]

60. Tyson, J.R. et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. bioRxiv 283077 (2020).

61. Freed, N.E. et al. Rapid and inexpensive whole-genome sequencing of SARS-CoV-2 using 1200 bp tiled amplicons and Oxford Nanopore Rapid Barcoding. Biol. Methods Protoc. 18;5(1):bpaa014 (2020).

62. Naveca, F.G. et al. Genomic, epidemiological and digital surveillance of Chikungunya virus in the Brazilian Amazon. PLoS Negl. Trop. Dis. 7;13(3):e0007065 (2019).

63. Kraemer, M.U.G. et al. Spread of yellow fever virus outbreak in Angola and the Democratic Republic of the Congo 2015–16: a modelling study. Lancet Infect. Dis. 17(3):330-338 (2017).

64. Neto, Z. et al. Molecular and genomic investigation of an urban outbreak of dengue virus serotype 2 in Angola, 2017-2019. PLoS Negl. Trop. Dis. 18;16(5):e0010255 (2022).

65. Warr, A. et al. No part gets left behind: tiled nanopore sequencing of whole ASFV genomes stitched together using Lilo. bioRxiv 470769 (2021).

66. de Vries, E.M. et al. Rapid, in-field deployable, avian influenza virus haemagglutinin characterisation tool using MinION technology. Sci. Rep. 12;11886 (2022).

67. Kumar, A. et al. Genome sequences of five Indian canine rabies virus isolates obtained using Oxford Nanopore Technologies sequencing. Microbiol. Resour. Announc. 19;11(5):e0124621 (2022).

68. Zakotnik, S. et al. Complete henome sequencing of tick-borne encephalitis virus directly from clinical samples: comparison of shotgun metagenomic and targeted amplicon-based sequencing. Viruses. 10;14(6):1267 (2022).

69. Riaz, N. et al. Evolution of within-host variants of the hepatitis C virus. Infect. Genet. & Evol. 99;105242 (2022).

70. Mori, M. et al. Nanopore Sequencing for characterization of HIV-1 recombinant forms. Microbiol. Spectr. 27:e0150722 (2022).

71. Maes, M. et al. Enteric adenovirus F41 genetic diversity comparable to pre-COVID-19 era: validation of a multiplex amplicon-MinION sequencing method. OSF Preprints. https://doi.org/10.31219/osf.io/6jku5 (2022).

72. Seitz, T.J. et al. Soil Disturbance Affects Plant Productivity via Soil Microbial Community Shifts. Frontiers in Microbiology 12:619711 (2022).

73. Devin Drown. Evaluating the effects of the changing permafrost ecosystem through the lens of genomics. https://nanoporetech.com/resource-centre/video/lc21/evaluating-the-effects-of-the-changing-permafrost-ecosystem-through-the-lens-of-genomics [Accessed 24 August 2022]

74. Wood, D.E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. Genome Biol 20:257 (2019).

75. Lu, J. et al. Bracken: estimating species abundance in metagenomics data. PeerJ Comp. Sci. 3:e104 (2017).

Oxford Nanopore Technologies

phone +44 (0)845 034 7900

email sales@nanoporetech.com

twitter @nanopore

www.nanoporetech.com

Oxford Nanopore Technologies

phone +44 (0)845 034 7900

email sales@nanoporetech.com

WP_1063(EN)_V3_09Sept2022