

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# An Introduction to Nanopore Sequencing: Past, Present, and Future Considerations

Morgan MacKenzie<sup>1</sup> and Christos Argyropoulos<sup>1,2,\*</sup>

<sup>1</sup> Department of Internal Medicine, Division of Nephrology, School of Medicine, University of New Mexico, Albuquerque, NM, 87131

<sup>2</sup> University of New Mexico Clinical & Translational Science Center, University of New Mexico, Albuquerque, NM, 87131

\*Corresponding author: cargyropoulos@salud.unm.edu

**Abstract:** There have been several developments in the field of nanopore biosensor development and sequencing applications that address previous limitations that restricted widespread nanopore use. These innovations, paired with the large-scale commercialization of biological nanopore sequencing by Oxford Nanopore Technologies, are making these platforms a mainstay in contemporary research labs. Equipped with the ability to provide long and short-read sequencing information with quick turn-around times and simple sample preparation, nanopore sequencers are quickly improving our understanding of unsolved genetic, transcriptomic, and epigenetic problems. However, there remain some key obstacles that have yet to be improved. In this review, we provide a general introduction to nanopore sequencing principles, discussing biological and solid-state nanopore developments, obstacles to single-base detection, and library preparation considerations. We conclude with examples of important clinical applications to give perspective on the potential future of nanopore sequencing in the field of molecular diagnostics.

**Keywords:** next-generation sequencing; nanopore sequencing; biosensors; single-molecule analysis; molecular diagnostics; genetics; transcriptomics; epigenetics

## 1. Introduction

In the past three decades, methods for single-molecule detection – the 3<sup>rd</sup> generation of sequencing methodologies – have broadened the scope of scientific research, making accessible cutting-edge sequencing technologies that enable a range of applications from clinical discoveries to the characterization of protein kinetics [1]. The utilization of biological nanoscale pores to detect nucleic acid molecules promised the potential of making single molecule sensing more accessible to a wider audience of researchers. A myriad of discoveries in the late 90s surrounding nanopore use – including the theoretical conceptualization of using nanopores for nucleic acid sequencing [2], the solving of the structure of staphylococcal alpha-hemolysin nanopore – the first biological pore used for nucleic acid translocation experiments [3] – and the proof of concept with alpha-hemolysin pore ( $\alpha$ HL) [4] – in many ways marked the beginning of subsequent research in both biological and solid-state nanopores throughout the early-mid 2000s. Their large-scale application by Oxford Nanopore Technologies (ONT) in the 2010s [2] has made Nanopore sequencing widely available, as these sequencers enable long-read sequencing and remain competitively priced compared to other platforms. The simplicity of these long-read sequencing systems makes these devices attractive for a range of applications, including genomic phenotype detection, structural variant detection, molecular biomarker discovery, and epigenetic research.

This review will provide a brief background on the development of nanopore sequencing technologies, starting with an overview of the general construction and sequencing principles involved in nanopore devices. We will then discuss multiple biological pore variants that have historically been used and those more recently developed, before

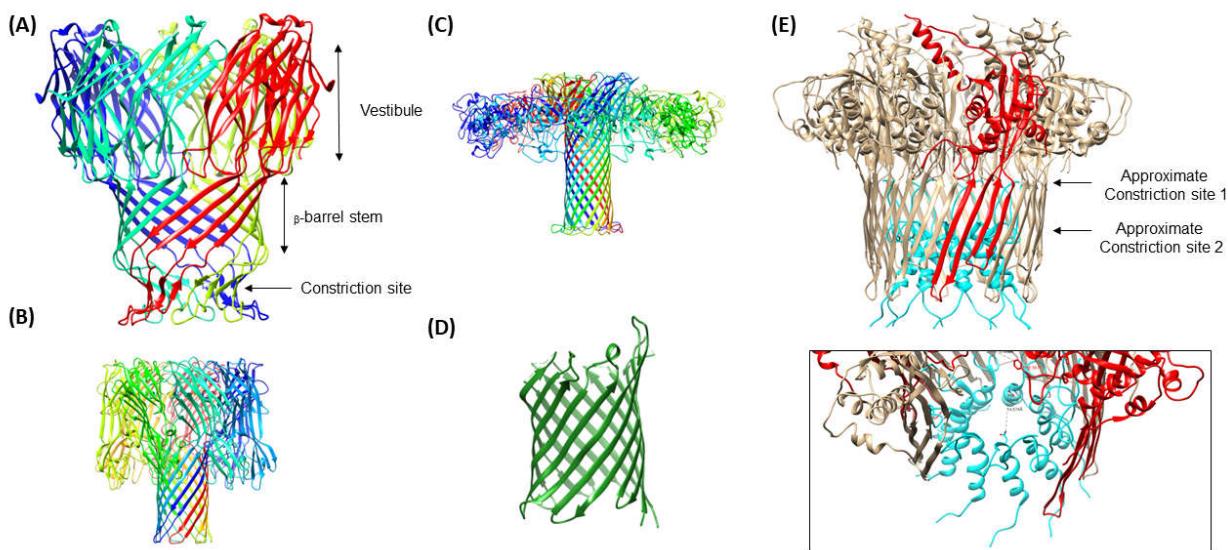


---

delving into an overview of solid-state nanopore fabrication principles and materials. Using this background as a jumping-off point, we will then discuss common applications utilizing nanopores, putting a specific focus on nucleic acid sequencing with high clinical relevancy. To introduce this discussion, we will open with library preparation technologies and recent modifications to these protocols that permit sequencing complicated sample types. We will use this to conclude our review with an investigation of how innovative sequencing approaches have permitted the identification of rare genotypes, biomarkers, and epigenetic phenomena associated with disease. We hope that this review will introduce nanopore sequencing for the curious reader, while providing perspective on the promise of this tool for progressing the field of molecular medicine.

## 2. Nanopore Sequencing Principles

The idea of single molecule detection on nanopore systems was independently conceptualized in the 1980s by several laboratories, including investigators David Deamer, George Church, and Hagan Bayley [5,6]. The original postulation was rooted in the theory that if exposed to an electrical current, oligomers could be driven through a protein nanopore channel, disrupting the current as they passed through in a manner characteristic of their base composition. In principle, nanopore sequencing relies on a biological or synthetic nanoscopic pore spanning the length of a membrane that separate two chambers filled with electrolytic fluid (for example, KCl, or Ag/AgCl systems). The sequencing chamber lies on the *cis* side, while the chamber into which an analyte exists the nanopore is termed the *trans* side [7,8]. Either chamber is connected to a voltage bias that distributes an ionic current throughout the nanopore, from the vestibule to the constriction site (Figure 1A) [6,8]. The mechanism is attached to a patch-clamp amplifier to permit the detection of the resultant signal (though this system has been compacted into portable ASIC chip systems by ONT [9]). In the case of nucleic acid analysis, the negative charge of the molecules causes them to drift away from the negative electrode, towards the anode and through the nanopore. As they do so, each nucleic acid base interacts with the ionic current to cause a disruption in the current. These nucleotide fingerprints can be mapped back to both the length of the strand, generally, and the characteristics of its component bases, specifically [4]. The translocation of DNA or RNA through nanopores can be characterized by event duration (the time the molecule takes to move through the length of the pore), and the magnitude of the current blockade during translocation [1]. These quantities underline the conversion of the electrical signal into a readout appropriate for the sequencing application at hand.



**Figure 1. Anatomy of a Nanopore** Depiction of biological nanopores employed in sequencing. (A) MspA pore (PDB ID 1UUN) with a constriction site diameter of 1.2 nm, and a stem length of ~3.7 nm (ref. 18). (B) Alpha Hemolysin (PDB ID 7AHL) with a constriction site diameter of 2.6 nm and a stem length of 5.2 nm (ref. 2). (C) Aerolysin porin (PDB ID 5J2T) with a constriction site diameter of 1 nm, and a stem length ~10 nm (ref. 13,14). (D) A simplified depiction of OmpG pore (PDB ID 2F1C) with a constriction site of 1.3 nm (ref. 15). (E) CsgG-CsgF mutant (PDB ID 6S17) with two constriction sites—the original CsgG constriction of 1 nm diameter (chain monomer of the original CsgG pore is depicted in red) and a secondary constriction caused by insertion of CsgF (Cyan residues) with a 1.5 nm diameter (ref. 25, 26). Insert shows the approximated diameter of the second constriction site, ~1.5 nm. Protein chains are depicted in different colors to help with distinction. All proteins were recreated in Chimera (ref. 99) utilizing PDB IDs from published protein structures (noted references).

### 3. Biological Nanopores

#### 3.1. Biological Nanopore Variants

Original experimentation of detection of homopolymers and single stranded nucleic acids by Kasianowicz et al. utilized *Staphylococcus aureus* alpha hemolysin ( $\alpha$ HL), a secreted pore-forming toxin that inserts into the bilipid membrane of its host, causing osmotic disruption and cell-lysis (Figure 1B) [3,4]. The original  $\alpha$ HL pore characterized by Song et al. in their 1996 study demonstrated that the pore was a heptamer consisting of a ~2.6 nm diameter transmembrane channel composed of 14 anti-parallel beta strands, an area narrow enough to accommodate single stranded DNA (ssDNA). This early experimentation verified the ability of the  $\alpha$ HL nanopore to allow passage of ssDNA, while also demonstrating that each ssDNA translocation disrupted the current in a measurable fashion [4,10]. Single base recognition was later demonstrated using  $\alpha$ HL [11], and the utilization of multiple recognition sites (sites within the pore where base recognition occurs) demonstrated a potential method to improve distinction between base (the “letters of the nucleic acid alphabet”) identities [12].

However, other biological nanopores with similar activities have been characterized and used for single molecule detection studies. Aerolysin protein, a pore forming toxin secreted by *Aeromonas hydrophila* has a stem that is 1.0-1.7 nm in diameter, an ideal size for nucleic acid sequencing (Figure 1C) [13,14]. Cao et al. successfully used aerolysin to detect variably sized deoxyadenosine chains and characterize the catalytic activity of endonuclease I [14]. This work notes that the small diameter of the aerolysin pore, in addition to electrostatic interactions between the pore wall and individual nucleotides, enable sensitive single base pair discrimination.

Outer membrane protein G of *Escherichia coli* has also been investigated as a potential nanopore for sequencing, given the fact that its constriction site measures at around 1.3 nm (Figure 1D) [15,16]. However, this 33kDa protein has open and closed conformation states that are pH and voltage dependent, and therefore it undergoes spontaneous gating events that makes it difficult to implement for single molecule sensing [15]. Through molecular dynamics simulations, Chen et al. identified a key aspartate residue at position 215 in the protein that, when mutated, decreased overall gating events per second [16]. With a double mutant consisting of the D215 deletion and an engineered disulfide bridge in beta chains 12 and 13, the Chen group was able to reduce gating events that might otherwise complicate single molecule detection. With the addition of a cyclodextrin adapter

within the pore's barrel, this mutant is able to detect ADP molecules as well. This pore has been more recently optimized through an alternative double mutation – the deletion of residues 221-227 and a mutation of the arginine at position 228 – which also appear to decrease gating events [17].

A popular choice in nanopore sequencing use is produced by *Mycobacterium smegmatis*, which has a smaller constriction site (~1.2 nm [18]) than the  $\alpha$ HL pore, making it better for single nucleotide resolution (Figure 1A) [19]. However, negative amino acid residues around the rim of the pore originally complicated analyte detection. Manrao et al. engineered a mutant of this pore, designed with a neutral instead of negatively charged mouth [19]. Utilizing a NeutrAvidin anchor to immobilize ssDNA into the pore, MspA was found to a) sensitively detect residual current levels characteristic of immobilized homopolymers of each nucleotide type, b) distinguish them based on their orientation (5' or 3' entry), and c) detect characteristic current differences between methylated and unmethylated cytosines [19]. The same group found that the region of sensitivity for this pore was approximately 14.5 nucleotides away from its anchor, and using this information, reported the detection of single nucleotide polymorphisms associated with breast or prostate cancer in genomic segments. The detection of the mutant MspA was improved with the addition of a molecular motor to the DNA that can dock onto the pore, a design motivated by the need to slow the translocation speeds of DNA below their natural rates (see below) [20]. More recently, the sensitivity of the MspA system has been used to characterize the kinetics of helicase enzymatic activity [21].

ONT recently introduced mutants of the curli transport lipoprotein, CsgG, to their nanopore devices [22]. This pore has reportedly been used in DNA sensing [23] and direct RNA sequencing applications [24]. While the constriction site of CsgG is narrow enough to permit sensitive base pair discrimination (~1.5 nm), a recently developed mutant has introduced a second constriction by inserting the naturally occurring accessory protein – CsgF [25] – to the interior of CsgG (Figure 1E) [26]. This mutant demonstrates improved single base resolution during DNA sequencing on ONT platforms [26].

### 3.2. Slowing Translocation Speeds in Biological Nanopores

Modulating the translocation speed through motor proteins facilitates the detection of electrical signatures associated with the passage of specific nucleotides through the pores. Too fast a translocation speed, and the current blockade will not be able to be detected without complicating high frequency clocked electronics. On the other hand, slow translocation implies that the single molecule process itself will take an extremely long time to conclude, thus compromising the use of nanopore devices for real time sensing. In many *biological* nanopore systems, the bacteriophage phi29 DNA polymerase (phi29 DNAP) has been used as a molecular ratcheting system to slow the translocation of nucleic acids through the nanopores via controlled 5'-3' synthesis. Translocation through MspA was slowed by docking a phi29-DNAP-DNA complex to the nanopore, where the polymerase synthesizes the DNA complement into the pore in a controlled manner [20]. The pairing of MspA to phi29 DNAP was able to slow DNA translocation and enable single-base discrimination. However, more recently helicase enzymes have been employed for this purpose, as they have been found to produce more sensitive current alterations and slow translocation to approximately 450 bp/sec for DNA [27–29]. To this end, it should be noted that these molecular motors are ATP dependent, and therefore the continuity of the sequencing experiment will depend on a consistent fuel source. Currently, a fixed amount of fuel is loaded at the beginning of the experiment and its depletion over time eventually leads to the termination of sequencing.

## 4. Solid-State Nanopores

While biological nanopores have been extensively developed and are a robust system, their relative shelf life, their limited reuse potential and the difficulty in engineering them to exacting levels make them an less than ideal route of nanopore sequencing.

Additionally, both the size of the constriction site of the nanopore trunk and the thickness of the membrane employed should ideally be close to the size of the analyte in question to increase sensitivity of detection and creating biological nanopores with this stringency is difficult [30]. Recent work with solid state nanopores (SSNPs) in silicon-based and 2D atomic-sized membranes are promising solutions to this problem and are broadly considered to hold a place in the future of nanopore sequencing, as they can be fabricated with high precision, are robust against high voltages and other experimental parameters, and can be integrated within microfluidic devices [22,30,31].

#### 4.1. Fabrication Techniques

In constructing SSNPs, drilling nanopores into the deposited material of choice should be precise and reproducible to facilitate accurate sequencing. Several techniques have been established, including but certainly not limited to focused ion beam (FIB) drilling/sculpting (often with  $\text{Ar}^+$  [32] or  $\text{Ga}^+$  [33] ions), transmission electron microscopy (TEM) drilling/sculpting [34] laser pulling of glass pipettes to create glass capillaries [35,36], evaporation induced self-assembly [37], and controlled dielectric breakdown [38,39]. (For thorough reviews on the creation of SSNP fabrication methods, we encourage readers to refer to the following reviews [30,40–42].)

Ion-beam sculpting was originally tested for SSNP construction [24]. Li and their collaborators demonstrated in 2001 that by exposing a  $\text{Si}_3\text{N}_4$  membrane with concave indentations to 3-KeV  $\text{Ar}^+$  ions, atoms can be stripped from the surface of the silicon-based membrane in a feed-back controlled manner, thinning the membrane and eventually forming a pore with the indentations on the opposite side [32]. This technique was utilized to create a ~5 nm diameter pore capable of detecting dsDNA with current reductions of up to 88% of the pore's center. Of note, this work demonstrated that under excessive ion-beam exposure, lateral atomic flow re-deposits the material across the nanopore opening, effectively closing it. Using this feed-back strategy, subsequent work drilled 100 nm diameter nanopores with FIB in silicon nitride membranes, then shrunk to diameters near 3 nm [43]. This small size (for reference, just slightly larger than the diameter of  $\alpha$ HL) permitted the analysis of various levels of dsDNA folding and intermolecular pairing [43]. FIB drilling has further been paired with ion scanning to create and modulate the size of an array of nanopores below 20 nm, and down to 5 nm [33].

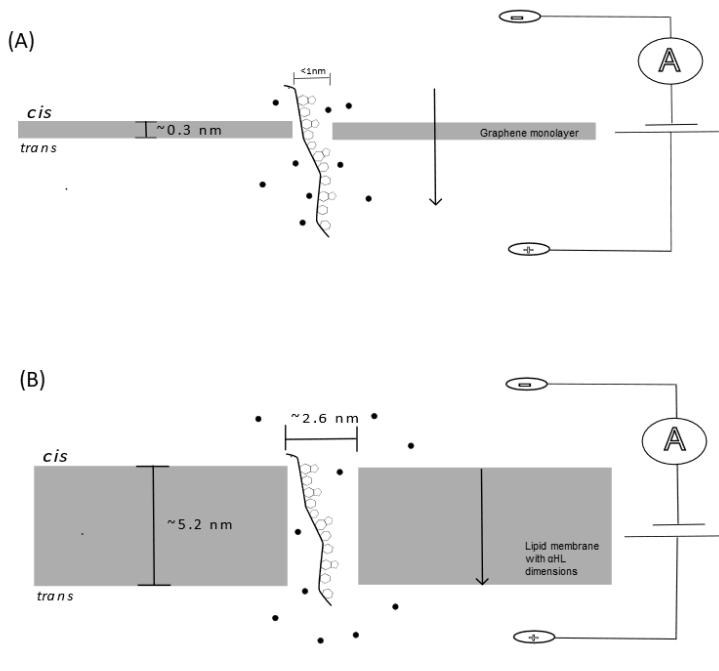
Transmission electron microscopy (TEM) has also been employed as an electron beam sculpting system to permit real-time analysis of the pore size and direct sculpting of the pores [34,40,44]. Storm's group utilized electron beam lithography to create pores within a silicon oxide membrane, which were then shaped to variable dimensions with TEM [45]. Their group found that the mechanism of pore size alteration depended on the starting material's thickness, where pores >80 nm could be widened and those under 40 nm would shrink. This controllable shrinking strategy was later determined to be the result of surface-tension induced mass flow resulting from the fluidization of the  $\text{SiO}_2$  material induced by TEM [46]. (For a thorough review of controllable shrinking strategies in SSNP systems, we direct the reader to other reviews [42,47].)

While TEM and FIB are powerful techniques for SSNP construction, they are expensive and require specialized equipment that may not be available for every lab. Additionally, while nanopores can be crafted before implementation in a fluidic chamber, introduction of prefabricated nanopores into electrolyte solutions may alter the characteristics of pores of certain materials [48]. *In situ* fabrication methods like controlled dielectric breakdown alleviate this concern, while also being more accessible and inexpensive methods for a wider range of laboratories [41]. Using the work of Kowk et al. as an example, controlled dielectric breakdown permits the formation of nanopores in solution by distributing a potential difference across the surface of a dielectric membrane, creating a strong electric field that creates nanopores in the surface of the starting material (in their case, silicon nitride) as a consequence of the induced charge build up [39]. While these pore sizes can be fabricated to ~1 nm diameters, it is difficult to control their location on

the surface of the membrane, which may result in irregularly sized pores [30,38]. Notably, controlled breakdown has been applied to silicon nitride membranes embedded in microfluidic devices, permitting enhanced detection of dsDNA and proteins [48].

#### 4.2. Materials for Construction

Although the list of explored SSNP materials is extensive, a few of the commonly tested materials of interest are silicon-based nanopores, glass nanopore capillaries, graphene monolayer assemblies, and molybdenum disulfide ( $\text{MoS}_2$ ) layers [30,49]. The choice of the material largely influences the maximum voltage that can be used during translocation experiments, (as certain materials are much more robust than others, which will start to erode under high voltage stressors [50]) as well as what chemical modifications can be added to the surface for enhanced analyte sensing (see below) [51]. Beyond this, constructing membranes of significant thinness is important for maximizing single-base resolution, as it limits the number of bases that contribute to the current disruption (Figure 2) [52].



**Figure 2.**

Comparison of membrane thickness of (A) theoretical, ideal graphene monolayer dimensions (membrane thickness  $\sim 0.3$  nm, with a pore diameter of  $<1$  nm, (ref. 52)) to (B)  $\alpha$ HL pore dimensions (membrane thickness  $\sim 5.2$  nm, pore diameter  $\sim 2.6$  nm (ref. 3)). Black arrows depict the direction of nucleic acid translocation from the *cis* to *trans* side of the membrane. In the case of the graphene monolayer membrane, only a single base contributes to the current disruption. Approximately 10 bases at one time can fit in the pore of  $\alpha$ HL, all of them contributing to the signal.

Silicon based systems have been proposed as a material for this application, because their theoretical thickness can be narrowed down to extremely thin sizes, they can withstand high voltage biases, and can be operated under greater bandwidths with decent resolution [50,53]. Rodríguez-Manzo et al. developed an electron irradiation-based technique for silicon membrane nanopore preparation with the goal of maximizing the conductance detection of the nanopore while decreasing baseline signal noise [54]. They scanned Silicon nitride films with scanning TEM to decrease the thickness of the silicon membrane, then used an electron probe to bore the nanopores into the surface of the synthetic membrane. The thinnest membrane achieved in this work was  $1.4 \pm 0.1$  nm, and each pore had diameters within  $1.3$ - $2.4$  nm (comparable to the diameter of the  $\alpha$ HL pores). Of note for silicon membrane construction, molecular dynamics simulations have demonstrated that the physical bottle neck for Si nanopore membranes is around  $0.7$  nm, after which point they are no longer stable [54]. Silicon based nanopores have also been constructed on glass chips to enhance stability and reduce the capacitance of the silicon membrane, allowing for translocation events to be detected with short event durations [55].

Graphene, a monoatomic layer of carbon grid structures, has many attractive physical qualities including electrical conductivity, malleability, and its impenetrability to ions

and experimental parameters (pH, temperature) [56]. Unlike silicon, graphene is stable as a monolayer system, and can exist comfortably at thicknesses of 0.3 nm [52]. Techniques have successfully grown graphene membranes on silicon chips and used electron beam drilling to create nanopores ~3.3 nm, enabling measurable differences between ssDNA and dsDNA [57]. However, similar 2D materials have been proposed for use as nanopore systems as well. MXene membranes – atom-thick layers of transition metal carbides [58] – have also been proposed as a potential substitute for graphene. Specifically,  $Ti_3C_2(OH)_2$  nanopore membranes have been studied in theoretical evaluations of sequencing efficiency and found to allow for base distinction between all four bases in molecular dynamics simulations [58].

#### 4.3. Controlling Noise and Translocation Speeds in Solid-State Nanopores

The implementation of SSNPs in nucleic acid sequencing is currently limited by the difficulty associated with slowing translocation speeds, which is needed to enhance the signal-to-noise ratio (SNR) of analyte translocation. Previous investigation has proposed that background, low frequency ( $1/f$ ) noise interferes with SNR in both biological and solid state nanopore systems, likely due *in part* to a) nanoscopic gaseous bubbles within the pore (nanobubbles) [59] or irregularities in the pore structure in the case of SSNPs, b) conformational changes in the case of biological pores, and/or c) electrode noise in either case [50]. While other methods of enhancing the signal-to-noise ratio of solid-state pores exist, an importance is placed on slowing down translocation time, enhancing the signal and minimizing these background frequencies. DNA naturally threads through nanopores at a rate of ~ 1 million bases/second, which is too fast for single base resolution with current electronics and computational techniques for signal processing. There are techniques which can be employed to slow nucleic acid translocation, including molecular motors and mutations within biological pores (as discussed above) and chemical group additions to the surfaces of synthetic nanopores and their membranes [51,60].

Such chemical modifications have been extensively studied, and include chemical and physical vapor deposition, atomic layer deposition (ALD), chemical group modifications of the surface layer, coating of solid-state surfaces with lipid bilayers, and the creation of hybrid nanopores by inserting biological pores with solid state tunneling [30,61–63]. By altering the charges of the surface and/or adding chemical components or probes for specific analytes, the interactions between the sequencing nucleic acid strands and the SSNPs are thought to be enhanced, improving the likelihood of successful and trackable translocation. For example Wang and their colleagues recently developed Gold- $Fe_3O_4$  nanoparticles, adapting them with peptide nucleic acids that can bind to targeted short RNA molecules [64]. These complexes enabled the detection of their translocation through glass quartz nanopores. The range of these applications is broad, and the methodology for chemical functionalization can be chosen by the investigator to produce desired interactions with the analyte of choice.

### 5. Library Preparation Considerations

Having been one of the preliminary motivations for the 1990's exploration of nanopore platforms, nucleic acid sequencing is perhaps the most common application of nanopores. Both direct and PCR-amplification methods have been used to execute DNA and RNA sequencing on nanopore platforms [see ref. 65 for comparison of direct and reverse transcription-PCR methods]. To facilitate sequencing, samples must be converted into the proper format for the sequencing platform in question through a process called library preparation. Library preparation effectively functions as both pre-analytic signal filtering (excluding molecules that are of no interest to the user) and as signal amplification of low input samples through nucleic acid amplification techniques such as polymerase chain reaction and rolling circle amplification. Library construction methods for DNA samples can easily be modified and adapted to the user's sample to permit the highest possible coverage of an analyte (e.g. genome or transcriptome) or to focus attention to a

particular molecule (or groups of molecules), i.e. targeted sequencing. The library preparation principle is similar across nanopore and non-nanopore based sequencing platforms, though we will exemplify the process according to Oxford Nanopore Technologies' numerous library preparation protocols [which are described here: 66]. More library preparation details can be found in the following review [29].

For many long-read DNA samples, first steps involve fragmentation of long sequences of DNA into smaller units, followed by end-repair enzymatic reactions of the damaged DNA [29,66]. This ensures a degree of uniformity of the molecules that pass through the pores, which in many cases allow a more thorough utilization of the fuel mix used in sequencing. This mechanism often involves the adenylation of DNA ends, making them compatible to hybridize with single-thymine overhangs attached to *sequencing adapters*. Said adapters are subsequently attached via enzymatic covalent ligation or rapid-attachment chemistries [63]. These sequencing adapters make the DNA library compatible with the nanopore platform, but also serve as a) reverse transcription and strand switching primers in the case of RNA → cDNA conversion protocols, and/or b) PCR primers in the case of amplification techniques. These primers can simultaneously serve as molecular barcodes for multiplexing experiments, which have the potential to drop the costs of sequencing significantly by enabling sequencing of multiple samples at a time. DNA can also be sequenced natively (without amplification) to avoid the introduction of PCR bias into the sample. However, in the case of limited DNA input (<100 ng [66]), amplification is useful to avoid sample lost throughout the library preparation process. This is particularly of consequence in the case of sequencing rare species that may be present in low concentration.

Following the attachment of sequencing adapters and the optional reverse transcription and/or PCR amplification, side reaction products (e.g. primer dimers) can be enzymatically degraded with a DNase enzyme (an important step to maximize the library depth of target sequences), and libraries are cleaned using either bead-based or column-based methods, both of which can be adapted to the sample content for sufficient retention. Finally, a second sequencing adapter is added to both 3' and 5' ends of the polymers just before initiating a sequencing experiment – this adapter includes a helicase motor that attaches to the nanopore and helps 'unzip' double stranded polymers, translocating them as they enter the pore. This step is also crucial to fine tuning the resolution of current disruptions, as it slows the translocation of the DNA molecules down to approximately 450 bp/s [29], a speed that allows the sequencing of DNA molecules and analysis of the electrical signals with simple electronics.

In the case of RNA library preparation, the general layout is similar: cDNA libraries can be created from RNA samples using reverse transcription, or RNA can be sequenced directly [65,66]. In the latter case, an optional single stranded cDNA synthesis can be performed to limit the formation of complex RNA secondary and tertiary structures (though this strand is not sequenced through the nanopores). Prior to sequencing, sequencing/motor protein adapters are attached through enzymatic ligation to the 3' end of the RNA product (direct sequencing) or rapid attachment to both strands (PCR-cDNA sequencing). While direct RNA (and DNA) sequencing can be utilized to preserve epigenetic modifications during sequencing, (as in [24]), large amounts of RNA (>500 ng) must be input to account for inevitable losses [65,66]. RNA is also sequenced at slower rates (~70 bp/second), so the overall yield from these experiments is lower than DNA runs [29].

Optimizations to library preparation methods have been executed to permit sequencing diverse libraries. Often, nucleic acid enrichment strategies are necessary to increase the library depth of targets. For example, running samples through ribosomal RNA or transfer RNA depletion prior to library preparation can limit the input of undesirable species to the library prep [67,68]. Cas9-mediated sequence specific adapter addition has also been proposed to select targets for downstream analysis [69]. A modification of this Cas9 protocol has been utilized alongside custom bioinformatics pipelines to detect fusion-pairs and breakpoint locations in cancer cell lines [70].

Further optimizations can be employed before or during library preparation to permit the sequencing of otherwise ignored species. For example, “Phospho-seq” utilizes a T4 polynucleotide kinase enzyme to add 5' phosphates and 3' hydroxyl groups to the ends of rarer subspecies of RNA that lack 5' phosphates or contain 3' phosphate groups [71]. This makes them compatible for the adapter attachments that are necessary during library preparation and may also improve the polyadenylation-base enrichment methods [72]. Recent work by our group has demonstrated the ability to perform highly effective co-sequencing of short and long coding and non-coding RNA species through universal poly-adenylation tailing, enabling contextualized quantification of all RNA species [73]. To further enhance the detection of target species, spike-in synthetic nucleotide mixes permit quantification, which we have demonstrated with mixed RNA samples using ERCC spike-in mixes [73], and others have demonstrated with RNA isoform identification and quantification using synthetic sequin RNAs [74].

## 6. Clinical Applications: Nucleic Acid Sequencing on Nanopores

While nanopores have been applied to numerous different areas of focus, including water purification [75], protein identification and characterization [31,76] and recently data storage [77] – we will focus on nucleic acid sequencing for the applications section of our review, though other reviews explore the wide field of nanopore technologies beyond this scope [78].

### 6.1. Genomics & Structural Variants

In 2018, Jain et al.’s group successfully sequenced a reference human genome *de novo* from the GM12878 cell line using Oxford nanopore’s CsgG mutant R.9.4.1 [79]. Their method included native DNA sequencing for the sake of accurately detecting repetitive elements and epigenetic modifications. This enabled them to achieve 30x genome coverage, 99.88% sequence accuracy and high levels of agreement with competitive short and long read platforms. The groups also successfully profiled complete MHC locus (an application directly applicable for optimal matching of human organs between donor and recipients in clinical transplantation), estimated telomere lengths, and methylation profiles [79]. This accomplishment demonstrates the adaptability of the nanopore system to optimize the platform for long-reads, while helping fill in the holes of sequences currently un-attainable by short-read sequencing platforms.

Fusion genes have also been a targeted application of nanopore sequencing, as fusion events are responsible for several forms of cancer. Identifying fused genes quickly in clinical samples is important to influence rapid treatment responses. The gold standard of fusion gene identification is fluorescence *in situ* hybridization, but the technique has a turn-around time of up to 48 hours and may be insensitive to some mutations [80,81]. Nanopore sequencing has been applied to these sample types because of its potential to deliver fusion gene readouts rapidly within 12 hours. Using a DNA adapter-ligation sequencing approach combined with modified bioinformatics, Jeck et al. were able to successfully identify the BCR-ABL1 gene rearrangement (diagnostic hallmark for chronic myeloid leukemia [81]) and the PML-RAMA fusion within seconds of sequencing even with low library depth of the target fusions. While a common complaint of nanopore sequencing involves its high error rates, this group found that even low-quality base calls were mappable to the regions of interest, something that our group has also found [73]. Using similar methodology, the same group was later able to sequence these same libraries on ONT’s Flongle device – their smallest, single-use flow cell that generates ~2.8 Gb of data with a cost below \$100 – and were able to capture all of the previously-identified fusion genes and the fusion *CIC-DUX4*, which is embedded in a locus with a high number of repeats [82]. This is a promising finding, as with improvements to the device structure and pore design, the inexpensive Flongle flow cell may prove to be an accessible diagnostic tool for both genomics and transcriptomic applications.

Among the sequencing problems that have been historically difficult to solve are short tandem repeat sequences, as they contain repetitive sequences difficult for short-read sequencers to localize and properly identify. The use of nanopores for STR sequencing remains a challenging prospect, as they tend to generate reads with high proportion of errors that among other considerations, a) increase with length of the repeat [83] and b) are unpredictable, varying with location of the repeat [84]. However, because the electrical squiggle signals from Nanopore sensors elucidate characteristics of base identity regardless of the composition of the nucleic acid, this issue is likely more so a result of the *base-callers* used during sequencing to convert electrical signals to bases of DNA than the platform's capabilities itself. This implies that algorithmic developments in the absence of any radical platform innovations may expand the application of Nanopore sequencing in this space.

In that regards, recent work compared Guppy to Bonito – two frequently used base-callers that utilize recurrent neural networks and convolutional neural networks, respectively [85] – to genotype autosomal and non-autosomal STR loci, and SNPs within them. While the investigators were able to genotype most STR loci correctly, the basecalling was easily obfuscated by the presence of homopolymers near the STR, highly repetitively expressed elements, and high sequence similarity – in other words, the success of the basecaller was dependent on the sequence. In support of this, previous work has demonstrated the successful sequencing of the mitochondrial genome of *Schistosome haematobium* with the Guppy basecaller, notably sequencing a tandem repeat region 18.5 kb long [86]. Thus, the success of this commonly used basecaller may be dependent on the applied scenario. Methods have been developed to improve error rates of basecalling. For example, a recent development translates each electronic signal from an STR unit into a matrix that is then converted to chromatic channels [87]. Deep convolutional networks are then used to infer the identity of each signal and allocate them to their designated STR regions.

While the above applications have been performed on a commercialized protein pore system, there have been investigations into the application of nucleic acid sequencing through solid state nanopores. Recently, Athreya and their colleagues used molecular dynamics simulations and electronic transport models to model the detection of single strand breaks in DNA molecules in graphene and MoS<sub>2</sub> pores [88]. It was found that in graphene, DNA nicks cause increased dwell times of a few nanoseconds, which they attribute to hydrophobic interactions between the un-phosphorylated DNA backbone point and the hydrophobic nature of the graphene. They found that depending on the nucleotide characteristics of the strand break, DNA strands will denature within the pores at different voltage biases. Therefore, they propose that the nick intersection location and nature of the surrounding bases can be identified by finding the characteristic voltage that causes strand dislocation. Interestingly, this effect was not found in the MoS<sub>2</sub> membrane, and was in fact the opposite - the DNA nicks cause the DNA to translocate to one side of the pore, leaving more room for the translocation of ions through the pore, thus decreasing the detected electronic signal while increasing the current passage. Applying these principles to a 2D-material sequencing experiment is a task for future SSNP studies.

## 6.2. Epigenetic Modifications

Epigenetic modifications of genomic material is likewise an attractive field for discovering novel mechanisms of disease development and control of gene expression. As an example, methylation of DNA is a well-studied mechanism of transcriptional silencing, however historically the transient nature of methylation and low sequence complexity of highly methylated regions have complicated sequencing methylated regions [89]. The gold standard for methylation studies is the bisulfite conversion technique, which converts methylated cytosines (5meC) into uracil residues – however, this approach may risk confounding these converted methylation patterns with experimental error [90]. Methylation studies have successfully been performed on nanopore instruments. Davenport et al. utilized standardized nanopore sequencing kits with the older R9.5 chemistry to

identify hypermethylation of cytosines, enabling the discovery of potential tumor suppressor genes that may be epigenetically silenced in hepatocellular carcinomas [91]. Nanopore identification of genome wide 5meC had high levels of agreement with standard bisulfite conversion and managed to discover 482 methylated genes that were invisible to short read sequencing platforms. However, the accuracy of these methods needs further verification and improvements to bioinformatic pipelines. Additional methods have been developed to improve the detection of methylated residues from the ionic current signal data of ONT nanopore sequencers using hidden Markov models combined with hierarchical Dirichlet processes [92].

#### 6.3. Infectious Disease Characterization and Detection

Rapid detection of viruses and other pathogens is important to mitigate outbreaks and improve treatment in clinical environments. The current methodology is slower and more expensive than optimal, while targeting specific species may inhibit the detection of low-concentration targets and miss important species [93]. A platform that offers improvements to both techniques could dramatically improve disease control. Utilizing long-read sequencing permits the detection of full-length pathogen genomes, while also enabling the characterization and identification of variants. Of note, ONT MinIon sequencers were recently used to track COVID-19 Alpha and Delta variants in Ukraine utilizing reverse transcription-driven cDNA sequencing [94]. Similarly, a group successfully sequenced a full monkeypox viral genome – including ITR sequences previously missed by other short read platforms – within 8 hours, obtaining sequencing depths of ~12-57x genome coverage [95]. Although the platform struggled with detecting homopolymers greater than a few base pairs long, which is still an issue of the *basecaller* more than the basic sequencing principles, the authors note that they had success utilizing HomoPolish to clean and correct some mismatches in their readouts. While these innovations are very promising, the library preparation and the data analysis pipelines of these systems will need to be improved and normalized before this becomes a standard diagnostic tool

#### 6.4. RNA Seq and Transcriptomics

Improvements to library preparation approaches, as discussed above, have permitted the sequencing of numerous RNA species, ranging from microRNAs, tRNAs, and circular RNA sequencing, to full-length transcriptomic isoform identification. Depledge et al. employed direct RNA sequencing to study the HSV-I transcriptome, employing a novel method to correct erroneous base calls [96]. Their method used proovread to align nanopore reads to a previously sequenced Illumina sequence, and from those corrections generated pseudotranscripts to identify read identities. While this approach resulted in improved mapping rates, the authors note the limited applicability of the technique outside of transcript isoform studies [96]. Recent investigation of microRNA sequencing coupled a MspA porin to phi-29 DNAP ratcheting protein and created chimeric microRNA-DNA hybrids for sequencing [97]. Their method permitted discrimination of isoforms and methylation markers, though did not capture the full microRNA body. Interestingly this work suggests that the microRNA sequencing on ONT sequencers may be complicated by their short length, though subsequent investigation by our group has demonstrated the platform is capable of detecting even primer-dimer sized transcripts, and thus microRNA sequences [73]. Circular RNA sequencing has also been performed on ONT sequencers, with protocols utilizing rolling-circle reverse transcription to capture full-length transcripts, enabling the detection of isoforms and fusion reads [98].

### 7. Conclusions

Within the past decade, there have been several improvements in the field of nanopore sequencing, making it a competitive technology for fundamental science research and clinical diagnostics. While innovations to protein pores have been commercially used to this date and offer high single-base resolution, there is motivation for further

development of solid state nanopores, as they offer highly customizable platforms that are more robust than protein pores. Additionally, the potential of perfecting 2D materials like graphene monolayers for nanopore fabrication would enable higher single-base pair resolution. Further research is needed to optimize these platforms, focusing on providing reproducible nanopore arrays by simpler, more accessible means. An important obstacle in this field is slowing translocation speeds of the analyte, which can be explored through chemical functionalization of the membrane surface.

A significant body of work has gone into improving and modifying library preparation techniques for diverse sample types. With the proper biochemical innovations, sample types that would otherwise go “unseen” by the sequencing platform can be captured and quantified. This is an important step in enabling the contextualization of disease processes, a crucial point for correctly analyzing epigenetic mechanisms and gene expression control. Additionally, the accuracy of nanopore sequencing can still be improved through changes in stranded sequencing. With ONT sequencers, the common “1D” sequencing approach does not currently permit sequential sequencing of both strands of a (c)DNA helix [30]. Previous attempts to permit this utilized a hairpin adapter that would anchor the other strand in place, waiting for the complete translocation of one strand before threading through the sister strand [99]. This approach enables an internal checking system for the basecaller, as it helps verify the primary sequence with the information garnered from the complement. While it was phased out from the company’s platforms for a few years, in 2022 this duplex sequencing seems to have been reintroduced, with some preliminary results suggesting improved sequencing accuracy [99].

Finally, while we introduced a handful of recent applications of this technology, the list is by no means complete. The potential of sequencing continuous long-read assemblies alongside shorter transcripts truly offers the potential of closing gaps in our knowledge of human genetics, transcriptomics, epigenetics, and infectious disease. In effect, this sequencing technology has the potential to permit biochemical reconstitution of human disease processes. Standardization of library preparation techniques and bioinformatics pipelines, combined with improvements in basecalling methods in particular, will be an essential part of making this technology widely used as a molecular diagnostic tool.

**Acknowledgements and Financial Support:** This grant was financially supported by DCI Inc. #C-3765 “A community based study of the Epidemiology of CKD in rural New Mexico” to CA. Research performed at the University of New Mexico Clinical and Translational Science Center (CTSC) was supported by an award from the National Center for Advancing Translational Sciences, National Institutes of Health under grant number UL1TR001449.

Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

**Author Contributions:** C.A. conceptualized the idea for the review. M.M. performed the literature review, wrote the manuscript, and made the figures. C.A. helped edit the review and provided feedback for content.

## References

1. Shi, W.; Friedman, A.K.; Baker, L.A. Nanopore Sensing. *Anal. Chem.* **2017**, *89*, 157–188, doi:10.1021/acs.analchem.6b04260.
2. Company History Available online: <https://nanoporetech.com/about-us/history> (accessed on 14 January 2023).
3. L Song; Song, L.; Hobaugh, M.R.; Shustak, C.; Cheley, S.; Cheley, S.; Cheley, S.; Bayley, H.; Gouaux, J.E.; Gouaux, J.E. Structure of Staphylococcal  $\alpha$ -Hemolysin, a Heptameric Transmembrane Pore. *Science* **1996**, *274*, 1859–1866, doi:10.1126/science.274.5294.1859.

4. Kasianowicz, J.J.; Brandin, E.; Branton, D.; Deamer, D.W. Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. *Proc. Natl. Acad. Sci.* **1996**, *93*, 13770–13773, doi:10.1073/pnas.93.24.13770.
5. Branton, D.; Deamer, D.W.; Marziali, A.; Bayley, H.; Benner, S.A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X.; et al. The Potential and Challenges of Nanopore Sequencing. *Nat. Biotechnol.* **2008**, *26*, 1146–1153, doi:10.1038/nbt.1495.
6. Deamer, D.; Akeson, M.; Branton, D. Three Decades of Nanopore Sequencing. *Nat. Biotechnol.* **2016**, *34*, 518–524, doi:10.1038/nbt.3423.
7. Fu, J.; Wu, L.; Qiao, Y.; Tu, J.; Lu, Z. Microfluidic Systems Applied in Solid-State Nanopore Sensors. *Micromachines* **2020**, *11*, 332, doi:10.3390/mi1103032.
8. Stoloff, D.H.; Wanunu, M. Recent Trends in Nanopores for Biotechnology. *Curr. Opin. Biotechnol.* **2013**, *24*, 699–704, doi:10.1016/j.copbio.2012.11.008.
9. How Nanopore Sequencing Works Available online: <https://nanoporetech.com/support/how-it-works> (accessed on 14 January 2023).
10. Deamer, D.W.; David W. Deamer; Deamer, D.W.; Branton, D. Characterization of Nucleic Acids by Nanopore Analysis. *Acc. Chem. Res.* **2002**, *35*, 817–825, doi:10.1021/ar000138m.
11. Ashkenasy, N.; Sánchez-Quesada, J.; Ghadiri, M.R.; Bayley, H. Recognizing a Single Base in an Individual DNA Strand: A Step Toward Nanopore DNA Sequencing\*. **2007**.
12. Stoddart, D.; Maglia, G.; Mikhailova, E.; Heron, A.J.; Bayley, H. Multiple Base-Recognition Sites in a Biological Nanopore: Two Heads Are Better than One. *Angew. Chem. Int. Ed.* **2010**, *49*, 556–559, doi:10.1002/anie.200905483.
13. Iacovache, I.; De Carlo, S.; Cirauqui, N.; Dal Peraro, M.; van der Goot, F.G.; Zuber, B. Cryo-EM Structure of Aerolysin Variants Reveals a Novel Protein Fold and the Pore-Formation Process. *Nat. Commun.* **2016**, *7*, 12062, doi:10.1038/ncomms12062.
14. Cao, C.; Ying, Y.-L.; Hu, Z.-L.; Liao, D.-F.; Tian, H.; Long, Y.-T. Discrimination of Oligonucleotides of Different Lengths with a Wild-Type Aerolysin Nanopore. *Nat. Nanotechnol.* **2016**, *11*, 713–718, doi:10.1038/nnano.2016.66.
15. Liang, B.; Tamm, L.K. Structure of Outer Membrane Protein G by Solution NMR Spectroscopy. *Proc. Natl. Acad. Sci.* **2007**, *104*, 16140–16145, doi:10.1073/pnas.0705466104.
16. Chen, M.; Khalid, S.; Mark S.P. Sansom; Sansom, M.S.P.; Bayley, H. Outer Membrane Protein G: Engineering a Quiet Pore for Biosensing. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 6272–6277, doi:10.1073/pnas.0711561105.
17. Sanganna Gari, R.R.; Seelheim, P.; Liang, B.; Tamm, L.K. Quiet Outer Membrane Protein G (OmpG) Nanopore for Biosensing. *ACS Sens.* **2019**, *4*, 1230–1235, doi:10.1021/acssensors.8b01645.
18. Faller, M.; Niederweis, M.; Schulz, G.E. The Structure of a Mycobacterial Outer-Membrane Channel. *Science* **2004**, *303*, 1189–1192, doi:10.1126/science.1094114.
19. Manrao, E.A.; Derrington, I.M.; Pavlenok, M.; Niederweis, M.; Gundlach, J.H. Nucleotide Discrimination with DNA Immobilized in the MspA Nanopore. *PLoS ONE* **2011**, *6*, e25723, doi:10.1371/journal.pone.0025723.
20. Manrao, E.A.; Derrington, I.M.; Laszlo, A.H.; Langford, K.W.; Hopper, M.K.; Gillgren, N.; Pavlenok, M.; Niederweis, M.; Gundlach, J.H. Reading DNA at Single-Nucleotide Resolution with a Mutant MspA Nanopore and Phi29 DNA Polymerase. *Nat. Biotechnol.* **2012**, *30*, 349–353, doi:10.1038/nbt.2171.
21. Craig, J.M.; Laszlo, A.H.; Brinkerhoff, H.; Derrington, I.M.; Noakes, M.T.; Nova, I.C.; Tickman, B.I.; Doering, K.; de Leeuw, N.F.; Gundlach, J.H. Revealing Dynamics of Helicase Translocation on Single-Stranded DNA Using High-Resolution Nanopore Tweezers. *Proc. Natl. Acad. Sci.* **2017**, *114*, 11932–11937, doi:10.1073/pnas.1711282114.
22. Brown, C.G.; Clarke, J. Nanopore Development at Oxford Nanopore. *Nat. Biotechnol.* **2016**, *34*, 810–811, doi:10.1038/nbt.3622.
23. Wellcomeopenres-2-16062.Pdf.
24. Stephenson, W.; Razaghi, R.; Busan, S.; Weeks, K.M.; Timp, W.; Smibert, P. Direct Detection of RNA Modifications and Structure Using Single-Molecule Nanopore Sequencing. *Cell Genomics* **2022**, *2*, 100097, doi:10.1016/j.xgen.2022.100097.
25. Zhang, M.; Shi, H.; Zhang, X.; Zhang, X.; Huang, Y. Cryo-EM Structure of the Nonameric CsgG-CsgF Complex and Its Implications for Controlling Curli Biogenesis in Enterobacteriaceae. *PLOS Biol.* **2020**, *18*, e3000748, doi:10.1371/journal.pbio.3000748.
26. Van der Verren, S.E.; Van Gerven, N.; Jonckheere, W.; Hambley, R.; Singh, P.; Kilgour, J.; Jordan, M.; Wallace, E.J.; Jayasinghe, L.; Remaut, H. A Dual-Constriction Biological Nanopore Resolves Homonucleotide Sequences with High Fidelity. *Nat. Biotechnol.* **2020**, *38*, 1415–1420, doi:10.1038/s41587-020-0570-8.
27. Derrington, I.M.; Craig, J.M.; Eric Stava; Stava, E.; Laszlo, A.H.; Ross, B.C.; Brinkerhoff, H.; Nova, I.C.; Doering, K.; Tickman, B.I.; et al. Subangstrom Single-Molecule Measurements of Motor Proteins Using a Nanopore. *Nat. Biotechnol.* **2015**, *33*, 1073–1075, doi:10.1038/nbt.3357.
28. Caldwell, C.C.; Spies, M. Helicase SPRNTing through the Nanopore. *Proc. Natl. Acad. Sci.* **2017**, *114*, 11809–11811, doi:10.1073/pnas.1716866114.
29. Wang, Y.; Zhao, Y.; Bollas, A.; Wang, Y.; Au, K.F. Nanopore Sequencing Technology, Bioinformatics and Applications. *Nat. Biotechnol.* **2021**, *39*, 1348–1365, doi:10.1038/s41587-021-01108-x.
30. Xue, L.; Yamazaki, H.; Ren, R.; Wanunu, M.; Ivanov, A.P.; Edel, J.B. Solid-State Nanopore Sensors. *Nat. Rev. Mater.* **2020**, *5*, 931–951, doi:10.1038/s41578-020-0229-6.
31. Mohapatra, S.; Lin, C.-T.; Feng, X.A.; Basu, A.; Ha, T. Single-Molecule Analysis and Engineering of DNA Motors. *Chem. Rev.* **2020**, *120*, 36–78, doi:10.1021/acs.chemrev.9b00361.
32. Li, J.; Stein, D.; McMullan, C.J.; Branton, D.; Aziz, M.J.; Golovchenko, J.A. Ion-Beam Sculpting at Nanometre Length Scales. *Nature* **2001**, *412*, 166–169, doi:10.1038/35084037.

33. Fürjes, P. Controlled Focused Ion Beam Milling of Composite Solid State Nanopore Arrays for Molecule Sensing. *Micromachines* **2019**, *10*, 774, doi:10.3390/mi10110774.

34. Muhammad Sajeer P; Simran; Nukala, P.; Manoj M. Varma TEM Based Applications in Solid State Nanopores: From Fabrication to Liquid in-Situ Bio-Imaging. *Micron* **2022**, *162*, 103347, doi:10.1016/j.micron.2022.103347.

35. Bafna, J.A.; Soni, G.V. Fabrication of Low Noise Borosilicate Glass Nanopores for Single Molecule Sensing. *PLOS ONE* **2016**, *11*, e0157399, doi:10.1371/journal.pone.0157399.

36. Li, W.; Bell, N.A.W.; Hernández-Ainsa, S.; Thacker, V.V.; Thackray, A.M.; Bujdoso, R.; Keyser, U.F. Single Protein Molecule Detection by Glass Nanopores. *ACS Nano* **2013**, *7*, 4129–4134, doi:10.1021/nn4004567.

37. Chen, Z.; Jiang, Y.; Dunphy, D.R.; Adams, D.P.; Hodges, C.; Liu, N.; Zhang, N.; Xomeritakis, G.; Jin, X.; Aluru, N.R.; et al. DNA Translocation through an Array of Kinked Nanopores. *Nat. Mater.* **2010**, *9*, 667–675, doi:10.1038/nmat2805.

38. Waugh, M.; Briggs, K.; Gunn, D.; Gibeault, M.; King, S.; Ingram, Q.; Jimenez, A.M.; Berryman, S.; Lomovtsev, D.; Andrzejewski, L.; et al. Solid-State Nanopore Fabrication by Automated Controlled Breakdown. *Nat. Protoc.* **2020**, *15*, 122–143, doi:10.1038/s41596-019-0255-2.

39. Kwok, H.; Briggs, K.; Tabard-Cossa, V. Nanopore Fabrication by Controlled Dielectric Breakdown. *PLoS ONE* **2014**, *9*, e92880, doi:10.1371/journal.pone.0092880.

40. Deng, T.; Li, M.; Wang, Y.; Liu, Z. Development of Solid-State Nanopore Fabrication Technologies. *Sci. Bull.* **2015**, *60*, 304–319, doi:10.1007/s11434-014-0705-8.

41. Fried, J.P.; Swett, J.L.; Nadappuram, B.P.; Mol, J.A.; Edel, J.B.; Ivanov, A.P.; Yates, J.R. *In Situ* Solid-State Nanopore Fabrication. *Chem. Soc. Rev.* **2021**, *50*, 4974–4992, doi:10.1039/D0CS00924E.

42. Lei, X.; Zhang, J.; Hong, H.; Yuan, Z.; Liu, Z. Controllable Shrinking Fabrication of Solid-State Nanopores. *Micromachines* **2022**, *13*, 923, doi:10.3390/mi13060923.

43. Li, J.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J.A. DNA Molecules and Configurations in a Solid-State Nanopore Microscope. *Nat. Mater.* **2003**, *2*, 611–615, doi:10.1038/nmat965.

44. Song, B.; Schneider, G.F.; Xu, Q.; Pandraud, G.; Dekker, C.; Zandbergen, H. Atomic-Scale Electron-Beam Sculpting of Near-Defect-Free Graphene Nanostructures. *Nano Lett.* **2011**, *11*, 2247–2250, doi:10.1021/nl200369r.

45. Storm, A.J.; Chen, J.; Ling, X.S.; H.W. Zandbergen; Zandbergen, H.W.; Zandbergen, H.W.; Dekker, C. Fabrication of Solid-State Nanopores with Single-Nanometre Precision. *Nat. Mater.* **2003**, *2*, 537–540, doi:10.1038/nmat941.

46. Storm, A.J.; Chen, J.H.; Ling, X.S.; Zandbergen, H.W.; Dekker, C. Electron-Beam-Induced Deformations of SiO<sub>2</sub> Nanostructures. *J. Appl. Phys.* **2005**, *98*, 014307, doi:10.1063/1.1947391.

47. van den Hout, M.; Hall, A.R.; Wu, M.Y.; Zandbergen, H.W.; Dekker, C.; Dekker, N.H. Controlling Nanopore Size, Shape and Stability. *Nanotechnology* **2010**, *21*, 115304, doi:10.1088/0957-4484/21/11/115304.

48. Tahvildari, R.; Beamish, E.; Tabard-Cossa, V.; Godin, M. Integrating Nanopore Sensors within Microfluidic Channel Arrays Using Controlled Breakdown. *Lab. Chip* **2015**, *15*, 1407–1411, doi:10.1039/C4LC01366B.

49. Wang, Y.; Yang, Q.; Wang, Z. The Evolution of Nanopore Sequencing. *Front. Genet.* **2015**, *5*, doi:10.3389/fgene.2014.00449.

50. Fragasso, A.; Schmid, S.; Dekker, C. Comparing Current Noise in Biological and Solid-State Nanopores. *ACS Nano* **2020**, *14*, 1338–1349, doi:10.1021/acsnano.9b09353.

51. Eggenberger, O.M.; Ying, C.; Mayer, M. Surface Coatings for Solid-State Nanopores. *Nanoscale* **2019**, *11*, 19636–19657, doi:10.1039/C9NR05367K.

52. Wasfi, A.; Awwad, F.; Ayesh, A.I. Graphene-Based Nanopore Approaches for DNA Sequencing: A Literature Review. *Biosens. Bioelectron.* **2018**, *119*, 191–203, doi:10.1016/j.bios.2018.07.072.

53. Venta, K.; Shemer, G.; Puster, M.; Rodriguez-Manzo, J.A.; Balan, A.; Rosenstein, J.K.; Shepard, K.; Drndić, M. Differentiation of Short, Single-Stranded DNA Homopolymers in Solid-State Nanopores. *ACS Nano* **2013**, *7*, 4629–4636, doi:10.1021/nn4014388.

54. Rodriguez-Manzo, J.A.; Puster, M.; Nicolaï, A.; Meunier, V.; Drndić, M. DNA Translocation in Nanometer Thick Silicon Nanopores. *ACS Nano* **2015**, *9*, 6555–6564, doi:10.1021/acsnano.5b02531.

55. Chien, C.-C.; Shekar, S.; Niedzwiecki, D.J.; Shepard, K.L.; Drndić, M. Single-Stranded DNA Translocation Recordings through Solid-State Nanopores on Glass Chips at 10 MHz Measurement Bandwidth. *ACS Nano* **2019**, *13*, 10545–10554, doi:10.1021/acsnano.9b04626.

56. Heerema, S.J.; Dekker, C. Graphene Nanodevices for DNA Sequencing. *Nat. Nanotechnol.* **2016**, *11*, 127–136, doi:10.1038/nnano.2015.307.

57. Garaj, S.; Song, L.; Liu, S.; Song Liu; Golovchenko, J.A.; Branton, D. Molecule-Hugging Graphene Nanopores. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 12192–12196, doi:10.1073/pnas.1220012110.

58. Prasongkit, J.; Jungthawan, S.; Amorim, R.G.; Scheicher, R.H. Single-Molecule DNA Sequencing Using Two-Dimensional Ti<sub>2</sub>C(OH)<sub>2</sub> MXene Nanopores: A First-Principles Investigation. *Nano Res.* **2022**, *15*, 9843–9849, doi:10.1007/s12274-022-4632-8.

59. Smeets, R.M.M.; Keyser, U.F.; Wu, M.Y.; Dekker, N.H.; Dekker, C. Nanobubbles in Solid-State Nanopores. *Phys. Rev. Lett.* **2006**, *97*, 088101, doi:10.1103/PhysRevLett.97.088101.

60. Jia, Z.; Choi, J.; Park, S. Surface Charge Density-Dependent DNA Capture through Polymer Planar Nanopores. *ACS Appl. Mater. Interfaces* **2018**, *10*, 40927–40937, doi:10.1021/acsami.8b14423.

61. Elibol, K.; Susi, T.; O'Brien, M.; Bayer, B.C.; Pennycook, T.J.; McEvoy, N.; Duesberg, G.S.; Meyer, J.C.; Kotakoski, J. Grain Boundary-Mediated Nanopores in Molybdenum Disulfide Grown by Chemical Vapor Deposition. *Nanoscale* **2017**, *9*, 1591–1598, doi:10.1039/C6NR08958E.

62. Asatekin, A.; Gleason, K.K. Polymeric Nanopore Membranes for Hydrophobicity-Based Separations by Conformal Initiated Chemical Vapor Deposition. *Nano Lett.* **2011**, *11*, 677–686, doi:10.1021/nl103799d.

63. Lepoittevin, M.; Ma, T.; Bechelany, M.; Janot, J.-M.; Balme, S. Functionalization of Single Solid State Nanopores to Mimic Biological Ion Channels: A Review. *Adv. Colloid Interface Sci.* **2017**, *250*, 195–213, doi:10.1016/j.cis.2017.09.001.

64. Wang, H.; Tang, H.; Yang, C.; Li, Y. Selective Single Molecule Nanopore Sensing of MicroRNA Using PNA Functionalized Magnetic Core–Shell  $\text{Fe}_3\text{O}_4$ –Au Nanoparticles. *Anal. Chem.* **2019**, *91*, 7965–7970, doi:10.1021/acs.analchem.9b02025.

65. Grünberger, F.; Ferreira-Cerca, S.; Grohmann, D. Nanopore Sequencing of RNA and cDNA Molecules in *Escherichia Coli*. *RNA* **2022**, *28*, 400–417, doi:10.1261/rna.078937.121.

66. DNA and RNA Sequencing Kits Available online: <https://nanoporetech.com/products/kits> (accessed on 14 January 2023).

67. Wahl, A.; Huptas, C.; Neuhaus, K. Comparison of rRNA Depletion Methods for Efficient Bacterial mRNA Sequencing. *Sci. Rep.* **2022**, *12*, 5765, doi:10.1038/s41598-022-09710-y.

68. Kraus, A.J.; Brink, B.G.; Siegel, T.N. Efficient and Specific Oligo-Based Depletion of rRNA. *Sci. Rep.* **2019**, *9*, 12281, doi:10.1038/s41598-019-48692-2.

69. McDonald, T.L.; Zhou, W.; Castro, C.P.; Mumm, C.; Switzenberg, J.A.; Mills, R.E.; Boyle, A.P. Cas9 Targeted Enrichment of Mobile Elements Using Nanopore Sequencing. *Nat. Commun.* **2021**, *12*, 3586, doi:10.1038/s41467-021-23918-y.

70. Stangl, C.; de Blank, S.; Renkens, I.; Westera, L.; Verbeek, T.; Valle-Inclan, J.E.; González, R.C.; Henssen, A.G.; van Roosmalen, M.J.; Stam, R.W.; et al. Partner Independent Fusion Gene Detection by Multiplexed CRISPR-Cas9 Enrichment and Long Read Nanopore Sequencing. *Nat. Commun.* **2020**, *11*, 2861, doi:10.1038/s41467-020-16641-7.

71. Giraldez, M.D.; Spengler, R.M.; Etheridge, A.; Godoy, P.M.; Barczak, A.J.; Srinivasan, S.; De Hoff, P.L.; Tanriverdi, K.; Courtright, A.; Lu, S.; et al. Comprehensive Multi-Center Assessment of Small RNA-Seq Methods for Quantitative MiRNA Profiling. *Nat. Biotechnol.* **2018**, *36*, 746–757, doi:10.1038/nbt.4183.

72. Feng, Y.; Cohen, S.N. Unpaired Terminal Nucleotides and 5' Monophosphorylation Govern 3' Polyadenylation by *Escherichia Coli* Poly(A) Polymerase I. *Proc. Natl. Acad. Sci.* **2000**, *97*, 6415–6420, doi:10.1073/pnas.120173797.

73. MacKenzie, M.; Tigert, S.; Lovato, D.; Mir, H.; Zahedi, K.; Barone, S.L.; Brooks, M.; Soleimani, M.; Argyropoulos, C. *To Make a Short Story Long: Simultaneous Short and Long RNA Profiling on Nanopore Devices*; Molecular Biology, 2022;

74. Gleeson, J.; Leger, A.; Prawer, Y.D.J.; Lane, T.A.; Harrison, P.J.; Haerty, W.; Clark, M.B. Accurate Expression Quantification from Nanopore Direct RNA Sequencing with NanoCount. *Nucleic Acids Res.* **2022**, *50*, e19–e19, doi:10.1093/nar/gkab1129.

75. Li, Y.; Gutiérrez Moreno, J.J.; Song, Z.; Liu, D.; Wang, M.; Ramiere, A.; Feng, Z.; Niu, Q.J.; Sasaki, T.; Cai, X. Controlled Synthesis of Perforated Oxide Nanosheets with High Density Nanopores Showing Superior Water Purification Performance. *ACS Appl. Mater. Interfaces* **2022**, *14*, 18513–18524, doi:10.1021/acsami.2c01474.

76. Zhang, S.; Huang, G.; Versloot, R.C.A.; Bruininks, B.M.H.; de Souza, P.C.T.; Marrink, S.-J.; Maglia, G. Bottom-up Fabrication of a Proteasome–Nanopore That Unravels and Processes Single Proteins. *Nat. Chem.* **2021**, *13*, 1192–1199, doi:10.1038/s41557-021-00824-w.

77. Chen, K.; Kong, J.; Zhu, J.; Ermann, N.; Predki, P.; Keyser, U.F. Digital Data Storage Using DNA Nanostructures and Solid-State Nanopores. *Nano Lett.* **2019**, *19*, 1210–1215, doi:10.1021/acs.nanolett.8b04715.

78. Ying, Y.-L.; Hu, Z.-L.; Zhang, S.; Qing, Y.; Fragasso, A.; Maglia, G.; Meller, A.; Bayley, H.; Dekker, C.; Long, Y.-T. Nanopore-Based Technologies beyond DNA Sequencing. *Nat. Nanotechnol.* **2022**, *17*, 1136–1146, doi:10.1038/s41565-022-01193-2.

79. Jain, M.; Koren, S.; Miga, K.H.; Quick, J.; Rand, A.C.; Sasani, T.A.; Tyson, J.R.; Beggs, A.D.; Dilthey, A.T.; Fiddes, I.T.; et al. Nanopore Sequencing and Assembly of a Human Genome with Ultra-Long Reads. *Nat. Biotechnol.* **2018**, *36*, 338–345, doi:10.1038/nbt.4060.

80. Hu, L.; Ru, K.; Zhang, L.; Huang, Y.; Zhu, X.; Liu, H.; Zetterberg, A.; Cheng, T.; Miao, W. Fluorescence in Situ Hybridization (FISH): An Increasingly Demanded Tool for Biomarker Research and Personalized Medicine. *Biomark. Res.* **2014**, *2*, 3, doi:10.1186/2050-7771-2-3.

81. Jeck, W.R.; Lee, J.; Robinson, H.; Le, L.P.; Iafrate, A.J.; Nardi, V. A Nanopore Sequencing-Based Assay for Rapid Detection of Gene Fusions. *J. Mol. Diagn.* **2019**, *21*, 58–69, doi:10.1016/j.jmoldx.2018.08.003.

82. Jeck, W.R.; Iafrate, A.J.; Nardi, V. Nanopore Flongle Sequencing as a Rapid, Single-Specimen Clinical Test for Fusion Detection. *J. Mol. Diagn.* **2021**, *23*, 630–636, doi:10.1016/j.jmoldx.2021.02.001.

83. Fang, L.; Liu, Q.; Monteys, A.M.; Gonzalez-Alegre, P.; Davidson, B.L.; Wang, K. DeepRepeat: Direct Quantification of Short Tandem Repeats on Signal Data from Nanopore Sequencing. *Genome Biol.* **2022**, *23*, 108, doi:10.1186/s13059-022-02670-6.

84. Ren, Z.-L.; Zhang, J.-R.; Zhang, X.-M.; Liu, X.; Lin, Y.-F.; Bai, H.; Wang, M.-C.; Cheng, F.; Liu, J.-D.; Li, P.; et al. Forensic Nanopore Sequencing of STRs and SNPs Using Verogen’s ForenSeq DNA Signature Prep Kit and MinION. *Int. J. Legal Med.* **2021**, *135*, 1685–1693, doi:10.1007/s00414-021-02604-0.

85. Tytgat, O.; Škevin, S.; Deforce, D.; Van Nieuwerburgh, F. Nanopore Sequencing of a Forensic Combined STR and SNP Multiplex. *Forensic Sci. Int. Genet.* **2022**, *56*, 102621, doi:10.1016/j.fsigen.2021.102621.

86. Kinkar, L.; Gasser, R.; Webster, B.; Rollinson, D.; Littlewood, D.; Chang, B.; Stroehlein, A.; Korhonen, P.; Young, N. Nanopore Sequencing Resolves Elusive Long Tandem-Repeat Regions in Mitochondrial Genomes. *Int. J. Mol. Sci.* **2021**, *22*, 1811, doi:10.3390/ijms22041811.

87. Perešini, P.; Boža, V.; Brejová, B.; Vinař, T. Nanopore Base Calling on the Edge. *Bioinformatics* **2021**, *37*, 4661–4667, doi:10.1093/bioinformatics/btab528.

88. Athreya, N.; Milenkovic, O.; Leburton, J.-P. Interaction Dynamics and Site-Specific Electronic Recognition of DNA-Nicks with 2D Solid-State Nanopores. *Npj 2D Mater. Appl.* **2020**, *4*, 32, doi:10.1038/s41699-020-00166-0.

---

89. Laird, P.W. Principles and Challenges of Genome-Wide DNA Methylation Analysis. *Nat. Rev. Genet.* **2010**, *11*, 191–203, doi:10.1038/nrg2732.

90. Schatz, M.C. Nanopore Sequencing Meets Epigenetics. *Nat. Methods* **2017**, *14*, 347–348, doi:10.1038/nmeth.4240.

91. Davenport, C.F.; Scheithauer, T.; Dunst, A.; Bahr, F.S.; Dorda, M.; Wiehlmann, L.; Tran, D.D.H. Genome-Wide Methylation Mapping Using Nanopore Sequencing Technology Identifies Novel Tumor Suppressor Genes in Hepatocellular Carcinoma. *Int. J. Mol. Sci.* **2021**, *22*, 3937, doi:10.3390/ijms22083937.

92. Rand, A.C.; Jain, M.; Eizenga, J.M.; Musselman-Brown, A.; Olsen, H.E.; Akeson, M.; Paten, B. Mapping DNA Methylation with High-Throughput Nanopore Sequencing. *Nat. Methods* **2017**, *14*, 411–413, doi:10.1038/nmeth.4189.

93. Petersen, L.M.; Martin, I.W.; Moschetti, W.E.; Kershaw, C.M.; Tsongalis, G.J. Third-Generation Sequencing in the Clinical Laboratory: Exploring the Advantages and Challenges of Nanopore Sequencing. *J. Clin. Microbiol.* **2019**, *58*, e01315-19, doi:10.1128/JCM.01315-19.

94. Yakovleva, A.; Kovalenko, G.; Redlinger, M.; Liulchuk, M.G.; Bortz, E.; Zadorozhna, V.I.; Scherbinska, A.M.; Wertheim, J.O.; Goodfellow, I.; Meredith, L.; et al. Tracking SARS-CoV-2 Variants Using Nanopore Sequencing in Ukraine in 2021. *Sci. Rep.* **2022**, *12*, 15749, doi:10.1038/s41598-022-19414-y.

95. Vandenbogaert, M.; Kwasiborski, A.; Gonofio, E.; Descorps-Declère, S.; Selekon, B.; Nkili Meyong, A.A.; Ouilibona, R.S.; Gessain, A.; Manuguerra, J.-C.; Caro, V.; et al. Nanopore Sequencing of a Monkeypox Virus Strain Isolated from a Pustular Lesion in the Central African Republic. *Sci. Rep.* **2022**, *12*, 10768, doi:10.1038/s41598-022-15073-1.

96. Depledge, D.P.; Srinivas, K.P.; Sadaoka, T.; Bready, D.; Mori, Y.; Placantonakis, D.G.; Mohr, I.; Wilson, A.C. Direct RNA Sequencing on Nanopore Arrays Redefines the Transcriptional Complexity of a Viral Pathogen. *Nat. Commun.* **2019**, *10*, 754, doi:10.1038/s41467-019-08734-9.

97. Zhang, J.; Yan, S.; Chang, L.; Guo, W.; Wang, Y.; Wang, Y.; Zhang, P.; Chen, H.-Y.; Huang, S. Direct MicroRNA Sequencing Using Nanopore-Induced Phase-Shift Sequencing. *iScience* **2020**, *23*, 100916, doi:10.1016/j.isci.2020.100916.

98. Liu, Z.; Tao, C.; Li, S.; Du, M.; Bai, Y.; Hu, X.; Li, Y.; Chen, J.; Yang, E. CircFL-Seq Reveals Full-Length Circular RNAs with Rolling Circular Reverse Transcription and Nanopore Sequencing. *eLife* **2021**, *10*, e69457, doi:10.7554/eLife.69457.

99. 1D Squared Kit Available in the Store: Boost Accuracy, Simple Prep Available online: <https://nanoporetech.com/about-us/news/1d-squared-kit-available-store-boost-accuracy-simple-prep> (accessed on 14 January 2023).

99. UCSF Chimera--a visualization system for exploratory research and analysis. Petersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem.* 2004 Oct;25(13):1605-12.