Review

Nanopore Technology and Its Applications in Gene Sequencing

Bo Lin [1,2], Jianan Hui [1] and Hongju Mao [1,2,*]

[1] State Key Laboratory of Transducer Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China; linbo201@mails.ucas.ac.cn (B.L.); jiananhui2@mail.sim.ac.cn (J.H.)

[2] Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

[*] Correspondence: hjmao@mail.sim.ac.cn; Tel.: +86-21-62511070-8707

Abstract: In recent years, nanopore technology has become increasingly important in the field of life science and biomedical research. By embedding a nano-scale hole in a thin membrane and measuring the electrochemical signal, nanopore technology can be used to investigate the nucleic acids and other biomacromolecules. One of the most successful applications of nanopore technology, the Oxford Nanopore Technology, marks the beginning of the fourth generation of gene sequencing technology. In this review, the operational principle and the technology for signal processing of the nanopore gene sequencing are documented. Moreover, this review focuses on the applications using nanopore gene sequencing technology, including the diagnosis of cancer, detection of viruses and other microbes, and the assembly of genomes. These applications show that nanopore technology is promising in the field of biological and biomedical sensing.

Keywords: nanopore technology; nanopore sequencing; diagnosis of cancer; biosensor

1. Introduction

Nanopore technology refers to nano-scale holes embedded in a thin membrane structure to detect the potential change when charged biological molecules smaller than nanopore pass through the hole [1]. Therefore, nanopore technology has the potential to sense and analyze single-molecule amino acid, DNA, RNA, etc. [2,3]. In this review, we will focus on the applications of nanopore technology in gene sequencing.

Nucleic acid is an important genetic material for most of the living body, and accurate sequencing of the nucleic acids is important for biomedical research, which would be useful for diagnosing human diseases and providing personalized medicine [4]. Since the last century, gene sequencing technology has developed dramatically, and now the nanopore technology has taken a leading role in the area of gene sequencing.

Figure 1 shows the development of gene sequencing technology. Since the 1950s, the determination of DNA double helix structure has preluded the research of genes. With the development of sequencing technology, the milestones have been made in gene sequencing. The whole genome of phage φ174 DNA was sequenced in 1977 [5], which marked the completion of the first genome sequencing in history. The Human Genome Project [6], proposed in 1990, laid a solid foundation for the further study of the human genome. The sequencing of Saccharomyces cerevisiae was completed in 1996, which completed the eukaryotic genome for the first time. The gene sequencing technology has been constantly innovating and evolving in the direction of lower cost, higher throughput, and faster speed [7–13].
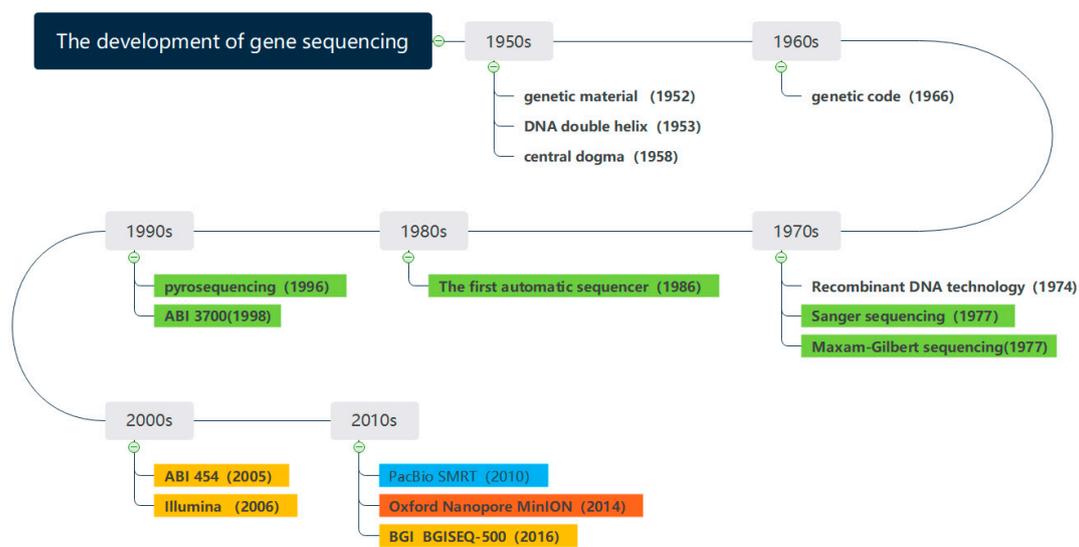
Figure 1. The development of gene sequencing technology.

The first-generation sequencing technology is represented by Sanger sequencing, which mainly utilizes chain-terminating method and gel electrophoresis technology. However, the downside of the Sanger sequencing equipment was costly and time consuming. As the Human Genome Project cost $3 billion and took 13 years to complete [6], further application of first-generation gene sequencing technology was limited.

The second- and third-generation sequencing technologies are characterized by high throughput and are also known as Next Generation Sequencing Technologies (NGS) [14]. Roche's 454 Sequencing Platform [15] and Illumina's Solexa Sequencing System [16] are the representative. The BGISEQ-500 [17], a second-generation sequencing platform later developed by China's BGI, has also taken some domestic market shares. Among the third-generation sequencing platforms, PacBio's SMRT technology has become the backbone with its high throughput and high accuracy.

The fourth generation of gene sequencing technology was the combination between gene engineering technology and computer-aided technology. Oxford Nanopore sequencing technology (ONT) innovatively used the nanopore technology to determine the base sequences from the current distortion when DNA passing through the nanopore [18]. ONT has become one of the most powerful sequencing technologies since its born. It has the advantages of the whole-genome sequencing and disease diagnosis with fast speed and cost-effective performance. Using ONT, the sequencing length of the largest DNA strand is continuously increasing with reported read length of 1 M base pairs [19]. Table 1 shows the comparison of four kind sequencing devices of different sequencing generation.

There are several kinds of nanopore sequencing platforms, including ONT, NabSys, and Sequenom [20]. ONT uses biological nanopore to get the sequence of DNA or RNA. Different from ONT, NabSys uses solid-state nanopore to sequence and it can also reconstruct the DNA map by using the short DNA probe. As for nanopore sequencing technology of Sequenom, it uses a readout system technology for single DNA molecules based on simultaneous optical probing of multiple nanopore. This paper mainly takes ONT as an example to introduce the principle of nanopore sequencing and its applications.

Table 1. The comparison of different generation of gene sequencing.

	Reading Length (kb) N50	Estimated Cost per Gb (US $)	Throughput per Flow Cell (Gb)	Read Accuracy (%)
Sanger(1st)	<1 kb	13,000 [d]	/	>99.9
Illumina(2nd)	0.075–0.15 [a]	50–63	16–30	>99.9
PacBio(3rd)	10–20 [b]	43–86	15–30	>99
ONT(4th)	10–60 [c]	21–42	50–100	87–98

The estimated cost in a, b, c excludes the cost for labor, instrumentation, maintenance, and computer resources. [a] Illumina NovaSeq 550 works in single-end mode. [b] PacBio Sequel II works in HiFi mode. [c] ONT PromethION works in long mode. [d] The cost was calculated by the HCV genome from [21] and the cost includes the operating costs. Data in the table comes from [21–23].

2. Principle of Nanopore Technology

Nanopore-based technologies originated from the Coulter and ion channels [24], which could be traced back to the 1990s [1]. Nanopore technology is done by applying a cathode and anode to the solution on the forward and reverse side of the membrane, respectively. Negatively charged biomolecules, such as DNA, can be placed on the forward side and these molecules can pass through pores in the membrane under the electrophoretic force with applied voltage. When different molecules translocate through the pores, the current level can be captured and be further used for calculation in computer-aided tools.

The category of nanopore used in nanopore technology can be divided into two parts, solid-state nanopore and biological nanopore. We first provide a brief overview of these nanopores in the following part.

2.1. Solid-State Nanopore

Solid-state nanopore can be fabricated using different methods, such as controlled breakdown, electrochemical reactions, laser etching and laser-assisted controlled breakdown [25]. It breaks the limit of natural occurring nanopores and has many advantages, such as very high chemical stability, control of diameter and channel length, adjustable surface properties, and the potential for integration into devices and arrays [26,27].

Si_3N_4 and SiO_2 nanopores are one of the most widely used nanopores and their fabrication is compatible with the complementary metal oxide semiconductor industrial integrated circuit processes [28]. These nanopores can be ion etched in free-standing Si_3N_4 and SiO_2 films, using argon ion beam or electron beam [29].

2.2. Biological Nanopore

Biological nanopore comes from natural protein molecules or artificial nanopores generated by genetic engineering [30]. However, the biological nanopores are frail with features such as short lifetime, intrinsic instability, and strict requirement of a specific environment, which are not able to support a biosensor's long-term operations.

Biological nanopores are usually produced by selected bacteria, such as α-hemolysin pore protein [31], MspA from Mycobacterium smegmatis [32], and Phi29 from Bacillus subtilis [33]. These biological nanopores are currently used for disease diagnosis [34], gene sequencing [35], and protein sequencing [36].

3. Nanopore Sequencing Technology

ONT is a single molecule sequencing technology based on nanopores. The first prototype, MinION, was released in 2014 [37]. The updated platform, PromethION, was released in 2015 with improved throughput. Two versions of PromethION, naming ProtheION 24 and 48, integrate 24 and 48 flow tanks, respectively. With the booming numbers of flow tanks compared to MinION, the PromethION system could output up to 7.6 Tb data while MinION could only generate 50 Gb within 72-h operation.

There are three forms of nanopore sequencing, 1D, 2D, and $1D^2$. 1D uses nanopore where only one strand of DNA is sequenced. 2D kit was first utilized in ONT. A hairpin structure was used at one end of the double-stranded DNA to connect two strands. After

completing the sequencing of one strand, the sequencing of the other strand begins immediately. In this way, it is equivalent to repeat the sequencing twice, which can be used for base correction. $1D^2$ is similar to 2D, whereas it doesn't need hairpins to physically bind the two strands of DNA together.

The reaction system for nanopore sequencing is carried out in a flow cell, in which two ionic solution-filled compartments were separated by membranes containing either 2048 (MinION) or 12,000 (PromethION) nanopores. The process of nanopore gene sequencing can be divided into three parts, library preparation, sequencing process, and basecaller.

3.1. Library Preparation

The preparation of the library is crucial for the subsequent work of nanopore sequencing. The DNA fragments should be repaired whether it has been sheared or not. When repairing, the repaired connector is a DNA-protein complex with a polymerase or helicase on the complex.

3.2. Sequencing Process

Figure 2 shows the schematic process of sequencing. The DNA strand to be sequenced is mixed with copies of the processive enzyme. When the DNA-protein complex approaches the nanopore, the enzyme binds to a single stranded leader at the end of the double stranded DNA template, unzips the double strand, and feeds a single strand through the nanopore. A single molecule with high specificity can interfere with the current when the unzipped DNA long strand passing through the nanopore one base at a time. These current signals can be used to determine the type of base.
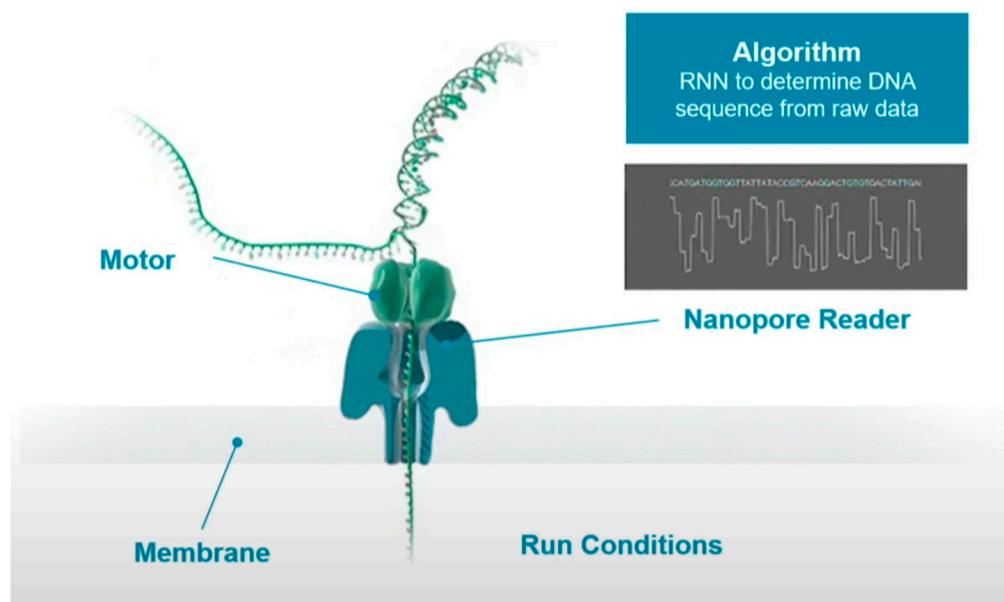
Figure 2. The process of nanopore sequencing [38].

3.3. Basecaller

In the process of base readout, owing to the difference in the charge and structure of nucleotides when they translocate through the nanopore, the measured current would cause small disturbance. These electrical signals can then be translated into DNA sequences with the deep learning algorithms. However, the readout signals are noisy and random as these signals are originated from more than one molecule in the nanopores, which is difficult for the basecaller.

In addition, the resistance of the hole is determined by the bases of multiple nucleotides located at the narrowest point of the hole [39]. Being the last step for the interpre-

tation of the entire DNA sequence, the data analysis using deep learning is challenging, which requires efficient algorithms and large amount of data for computational training.

4. Optimization Ways in ONT

Traditional nanopore sequencing technology is not able to replace the previous generation of sequencing devices owing to the high error rates, up to 5–15% [40,41]. To improve the sequencing accuracy of nanopore sequencing technology, it is necessary to analyze from the perspective of error source. Firstly, the Signal Noise Ratio (SNR) is affected by the inherent characteristics, including the structural similarity of nucleotides, the constant signal in the homopolymer, and the non-uniform velocity of nucleotide through the pore. The SNR needs to be optimized during the preparation of DNA library. Moreover, the process of translating the read current signal into base sequence is limited by the type of basecaller program, which needs to be optimized in the software and algorithms.

Also, the structure of nanopore structure can be upgraded to enhance the performance of the sequencing technology [42]. In 1D sequencing mode, R9 nanopore with 1 reader head can achieve the accuracy of 95% and consistent accuracy of Q44 (99.996%). However, one reader head cannot accurately identify base with multiple identical bases in a row. Hence, besides combining with the latest basecalling software, the number of reader head could increase to be dual headed. Also, the combination of the chips with one and dual reader heads can obtain higher consistent accuracy. In this review, the detailed ways for the optimization of nanopore sequencing are listed as follows:

4.1. Optimization in Library Preparation

This review mainly reports an optimization method called intramolecular-ligated nanopore consensus sequencing (INC-Seq), developed by the Singapore Genome Research Institute [43]. Based on the concept of PacBio's CCS, the INC-seq repeatedly sequences the same molecule and hence a consensus sequence can be calculated to improve accuracy [44].

As the linear DNA template is needed in nanopore sequencing, the researchers adopted a novel library preparation scheme. As shown in Figure 3a, the template DNA molecule is first cycled and then amplified using roll-circle amplification to generate circular molecules. The circular molecules consist of multiple repeating units, which are then broken into linear DNA strands before sequencing on a nanopore sequencing platform.
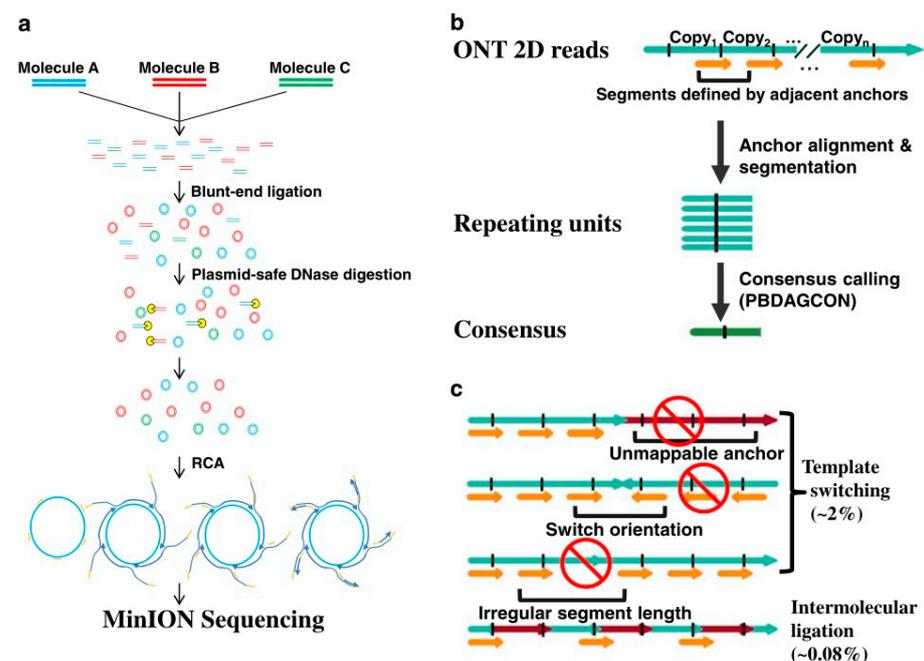
Figure 3. The schematics of INC-seq sequencing technology show (**a**) the process of making RCA, (**b**) the process of locating the repeat sequence with anchor, and (**c**) the sources of chimeras [43].

After obtaining the rolling-circle DNA sequencing results, the sub-read from the original nanopore is used as anchors to scan the entire reading to find the location of duplicated units. The sequence results can be read from the consensus sequence by comparing between the start point and the end point of an adjacent anchor. In addition, in the process of library preparation, each read in INC-Seq is required to be tandem repeats of the same template for the calibration protocol, as shown in Figure 3b.

The main problem faced by this method in library preparation is chimerism, which can lead to a poor or chimeric consensus during sequence interpretation. As shown in Figure 3c, two types of patterns are exhibited in the chimeras of genes, one from template switching and the other from intermolecular ligations. The one from template could induce incompatible mappings of anchor points, such as unmappable anchor points and orientation switch, or irregular anchor distance.

To solve these two problems, researchers designed a bioinformatics pipeline that extracts repeating sequence segments and corrects these segments by constructing a consensus sequence, also called anchor-based consensus construction. Instead of detecting the chimeras from intermolecular ligations during the transmission process, the detection is made by detecting whether the consensus sequences are more than twice as long as expected. In the conditions of the INC-Seq experiment, the observation of chimeras from intermolecular ligations was very rare, which is only about 0.08%.

The INC-seq enables accurate species-level classification and identification at 0.1% abundance. It could robustly quantify the relative abundances on the MinION system. Moreover, by applying INC-Seq for 16S rRNA-based bacterial profiling, it could generate full-length amplicon sequences with a median accuracy >97% [43].

4.2. Optimization of Basecalling Programs

With the development of GPU, the traditional machine learning method, Hidden Markov models, has been replaced by deep learning methods, which enables to train more complicated neural network. Many different aspects could affect the predicted results, by the neural network, including the construction method of hidden layer, the architecture of the neural network for modeling, the species of biological genome corresponding to the data set in the training process, etc. [39]. Table 2 concludes the normal machine learning algorithms and deep learning networks that could be used to construct the training program. In addition, larger neural networks can improve the accuracy of basecalling and consensus sequences at the expense of sequencing speed [39].

Table 2. Neural network algorithms used in nanopore basecaller.

Algorithms	Descriptions
Hidden Markov models [45]	A stochastic model that makes predictions based only on the previous event and a series of observations.
Recurrent Neural Network (RNN) [46]	A model that allows networks with periodic connections to learn complex tasks that require information to be maintained for fixed or indeterminate periods of time.
Long-short-term memory (LSTM) [47]	A type of recurrent neural network that can be used as a component of a larger network. It has specific input, output, and forgetting gates, which can be implemented to retain or discard information passed from the previous state.
Convolutional Neural Network (CNN) [48]	A neural network model for image classification, which extracts input features through a convolution algorithm.
Connectionist Temporal Classification (CTC) [49]	A convolutional neural network for marking neural network output and scoring of sequence data. It does not require pre-segmented training data and post-processed output

Figure 4 shows the base model of convolutional neural networks and recurrent neural network. The convolutional neural network (ConvNet), including CNN and CTC, has the ability to extract the features in the input training data. The pre-processing required in a ConvNet is much less compared to other classifications. With enough training, ConvNets is able to learn different features. The recurrent neural network, including RNN and LSTM, is another kind of artificial neural network where the connections between nodes in the network form a directed graph along a temporal sequence. This allows it to exhibit a temporal dynamic behavior. Deriving from feed forward neural networks, the recurrent neural networks can use their internal state (memory) to process inputs with variable length sequences [50].
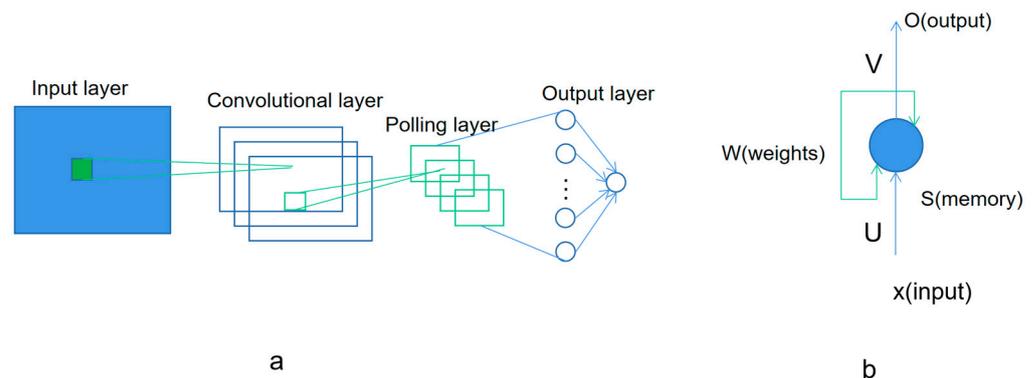
Figure 4. The base model of (**a**) the convolutional neural network and (**b**) the recurrent neural networks.

Many studies have focused on improving the basecaller by developing an appropriate algorithm. When the MinION sequencing platform was launched in 2014, the basecaller was called Metrichor, a cloud-based EPI2ME platform that uses Hidden Markov models for basecalling. However, this service was discontinued in March 2017. In August 2016, the Minknow software began supporting basecalling, which can be downloaded to a PC for monitoring and controlling MinION sequencing. ONT also offers several other basecalling programs, including the command-line basecalling program Albacore, the programs naming Nanonet and Scrappie for research and testing [41]. In addition, many researchers have also developed several basecalling programs based on various deep learning algorithms, including Nanocall [51], SACall [52], DeepNano [53], Chiron [54], and BasecRAWller [55]. Nanocall is the first open basecalling source that could be downloaded free and operated offline. The Nanocall and SACall are reviewed as follows.

4.2.1. Nanocall Basecalling Program

Nanocall is a program realizing deep learning algorithms using C when Metrichor was still servicing [51]. The advantage of Nanocall is that it offers offline operation when running basecalling program while Metrichor must be connected to the internet. Also, Nanocall uses Expectation maximization (EM) algorithm to perform several rounds to identify bases more accurately.

However, Nanocall has disadvantages compared to Metrichor, as it performs basecalling in the 1D form and, thus, it cannot integrate information from complementary chains in the 2D sequencing form. When it comes to 2D sequencing form, it uses a special signal when the hairpin structure passes through the nanopore to realize the separation of the two chains. The signal cannot be integrated due to the lack of base in the DNA backbone.

4.2.2. SACall Basecalling Program

SACall is an end-to-end basecaller composed of the convolution layer, transformer self-attention layer, and a CTC decoder. It is a program to convert the current signal to base sequence by using neural network. In SACall, the convolutional layer is used to subsample

the signal to capture the local information. To realize the upstream and downstream correlation of signals, the self-concern layer is used to calculate the similarity of every two positions in the original signal sequence. The CTC decoder uses a beam search algorithm to obtain the base sequence.

Huang et al. used a benchmark set consisting of nine independent genomes to compare the performance of SACall with the official test programs, including Albacore and Guppy, for nanopore sequencing [52]. SACall performs better than the other two programs in the readout accuracy, assembly quality, and consensus accuracy. SACall is also an open-source program with the advantage of wider spreading among researchers.

4.3. Optimization of Nanopore Types

Currently, nanopore R9.4.1 has been used in multiple areas for rapid sequencing and showed a good performance, including cancer research [56], human genetics [57] and microbiology [58]. However, as R9.4.1 has only one reader head, it cannot accurately identify bases in some special sequences. For example, in the homopolymeric regions with multiple continuous identical bases, the deletion of bases would occur when using R9.4.1 with one reader head.

In 2019, R10 was introduced with dual reader head, which can better recognize and correct the results when detecting DNA with identical bases. When DNA molecules pass through the nanopores of R10 chip, the dual reader would detect the sequence twice as high as that of R9 to determine the bases. Hence, it could better identify low complexity sequences, minimize errors, and improve consistent accuracy. Since then, the Oxford Nanopore R&D teams have been working to tune the nanopore to increase the throughput and accuracy. Similar to R10, R10.3 was introduced with dual reader header while the input amounts and raw accuracy were improved. The improvement of read accuracy and processing speed can be found in Figure 5.
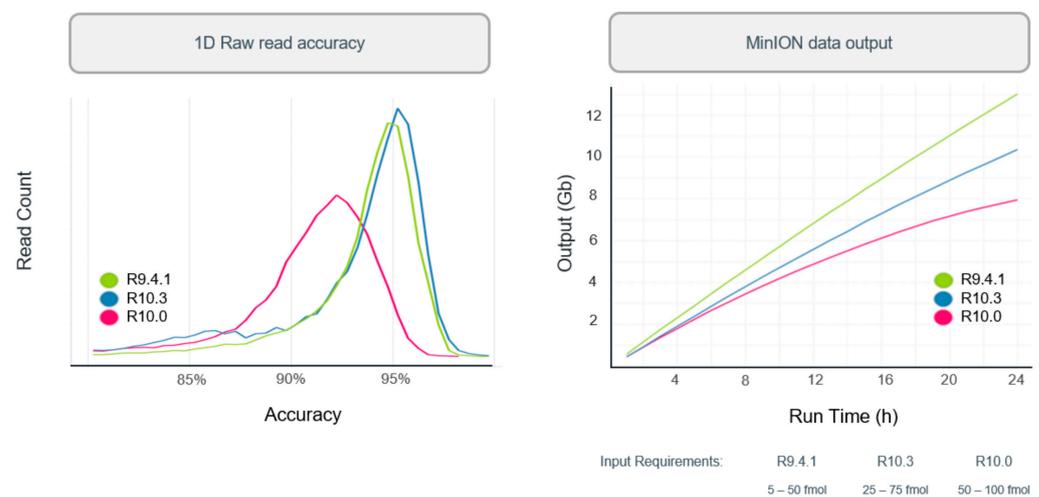
Figure 5. The comparison of performance using different types of nanopore. Oxford Nanopore Technologies Inc., Oxford, UK, 2020 [59].

5. Applications of Nanopore Technology

There are many applications using nanopore sequencing as it enables an easier and more accessible study of genetic and epigenetic modification as well as their role in gene expression [60]. Besides, in the area of biomedical research, nanopore sequencing can provide insights into the mechanism of diseases and solutions to the problems in medical diagnosis and testing from the perspective of gene and chromosome. For detecting the new types of virus and bacteria, nanopore gene sequencing also gives a high sequencing solution and a portable sequencing platform. This review will give representative applications of nanopore technology. Table 3 concludes the applications of the nanopore technology and some of the applications are mentioned particularly in the following part.

Table 3. Applications of Nanopore Technology.

Applications	Descriptions
Clinical research	With nanopore technology, the long read can help researchers to identify and phase genetic variant, and fully characterize novel isoforms and fusion transcripts. Nanopore technology gives a new insight to health and disease, from cancer, immunology, to neuroscience. The representing references are [61–66]
Detection of microbes	The nanopore technology can be used to sequence the DNA or RNA sequence of microbes, and it helps researchers to classify or monitor the microbes. Moreover, it is promising to establish microbe surveillance and response quickly during pandemic, if the nanopore technology can be applied in the area of public health. The representing references are [34,35,67,68]
Assemble genomes	Owing to the long-read sequencing ability, nanopore sequencing technology can overcome the problems that short-read sequence devices meet in area the long-repeated fragments. The representing references are [35,69–71]
Environmental genomics	The portable and affordable nanopore sequencing technology provides a unique tool for environmental research, including biodiversity assessment, pathogen identification and animal conservation. Besides, the real-time data analysis provides immediate access to results, whether in labs or in the field. The representing references are [72–74]

5.1. Diagnosis of Cancer by Gene Modification

5.1.1. Structure Variations

The structural variation in the cancer genome is characterized by a large number of deletions, amplification, inversion, duplication, and translocations. Changes in these sequences usually lead to abnormality in gene regulation such as fusion genes and copy number changes, which may lead to activation or overexpression of oncogenes and inactivation of tumor suppressor genes [75]. Nanopore sequencing technology can overcome the structural variation challenges of short read long sequencing meets when it comes to copy with the long-repeated fragments.

Norris et al. detected the inactivation of tumor suppressor genes (CDKN2A/p16 and Smad4/DPC4) due to structural variation in pancreatic cancer using nanopore sequencing [76]. It shows that nanopore sequencing technology has a certain application prospect in the detection of a series of structural variations with good characteristics. Williams et al. also demonstrated that targeted nanopore sequencing is an effective way to identify ABCB1 structural variation in THP-1AML cells and HGSCs [63].

Nanopore sequencing gives access to genomic regions that may be inaccessible to traditional methods of sequencing. If nanopore technology can be applied in clinical practice, it will become an ideal tool for identification of structural variation and the early diagnosis, treatment, prognosis of cancer, and therapeutic monitoring.

5.1.2. Transcription Factor

Transcription factor (TF) is the main regulation factor of gene expression and signaling pathway in all known biological systems. It is a group of genes that can bind specifically with specific sequences of the upstream 5'- terminal of the gene, and, thus, bind with specific DNA sequences to promote or suppress gene expression.

In cells, the majority of oncogenes and tumor suppressor genes encode TF [65,77]. Its abnormal activity can be used to characterize the presence of cancer clinically. Squires et al. used nanopores to analyze the ion current sublevels generated by the translocation of the ZIF268/DNA complex by single molecular analysis, proving that nanopore technology can be used to distinguish the specific and non-specific binding conformations of ZIF268 [78]. This further demonstrated nanopore to be capable of detecting complex structures and protein conformations.

These studies open the door to new applications of nanopore technology, which can be used to study DNA complex research work and is expected to be targeted to guide the diagnosis and treatment of cancer. Furthermore, nanopore sequencing technology also

gives a chance to investigate molecular mechanisms of the TFs [79]. Also, it could provide an in-depth insight into a broad panel of human cell lines [80].

5.1.3. Telomeres

Telomeres are DNA-protein complexes at the ends of linear chromosomes in eukaryotic cells, which forms a special "cap" structure together with telomere binding proteins. Telomeres, centromeres, and replication origin are the three essential factors for chromosome integrity and stability. Telomeres shorten by an average of 19 bases per year due to aging, oxidation, stress, mitotic activity, and lifestyle habits.

Ding et al. used α-hemolysin as a nanopore sensor to study the folding structure of I-motif in human telomeres at different pH values [81]. The potential of α-hemolysin as part of biosensor development has been demonstrated. It is helpful for us to understand the lifetime and biologically related structure of I-motif of telomere sequences.

If the nanopore technology is proved to be valid to investigate the telomeres, we can better understand the dynamics and mechanisms of telomeres and analyze the relationship between telomere length and environment factors, including oxidation, stress, etc. In addition, the study of telomere G-quadruplexes molecules in the telomere structure of cancer cells is expected to provide guidance for the diagnosis and treatment of cancer in a broader way [82,83]. By studying the telomeres, more binding molecules on G-quadruplexes may be explored to be applicable to the treatment of a wide range of human cancers.

5.2. Diagnosis of Cancer by Epigenetics

5.2.1. DNA Methylation

In the human genome, DNA methylation is an epigenetic modification, including 5-methylcytosine (5-mC), N4-methylcytosine (4-mC), and N6-methyladenine (6-mA) [84]. In mammalian cells, CpG methylation can directly or indirectly suppress gene expression [85]. Moreover, the degree of abnormal methylation (high or low methylation) is associated with some cancers. DNA methylation pattern is one of the earliest and most common molecular changes in human tumors [86]. Studies have shown that almost all cancer types have hundreds of genes with an abnormally increased methylation [87,88]. There is also an increasing number of studies using nanopore sequencing technology to further investigate DNA methylation.

Jiwook et al. [89] proposed a nanopore array method for DNA methylation detection, which bypassed traditional bisulfite conversion [90] and amplification by PCR reaction [91], demonstrating that the use of 10 nm nanopore can distinguish between hypermethylated and unmethylated dsDNA oligonucleotides. The work provides a direct electrical analysis technique to detect various methylation levels on DNA fragments at the single-molecule level, showing the potential of nanopore to identify abnormally methylated DNA in clinical tests aimed at diagnosis of diseases such as cancer.

Lee et al. [60] used nanopore sequencing technology to detect endogenous CpG methylation and at the same time to exogenously label chromatin accessibility sites. By analyzing four human cell lines, the team constructed a human epigenome map that included information on CpG methylation and chromatin accessibility, and revealed differences in methylation and chromatin accessibility between breast cancer cells and non-cancer cells.

The research mentioned above both demonstrated that the use of nanopore technology to detect abnormal DNA methylation may play an important role in cancer therapy and precancerous detection [92,93]. The nanopore-based methylation-sensitive assay provides a more convenient method in studying the role of epigenetics in human disease without bisulfite conversion, fluorescent labeling, and PCR.

5.2.2. MicroRNA

MicroRNAs (miRNAs) are small endogenous biomolecules, with a length of 18–22 bps, which can regulate gene expression and their expression level is correlated with different

diseases [94]. Studies have shown that abnormal expression of miRNAs has been found in different tumor tissues [95,96]. In addition, miRNAs play important roles in embryonic differentiation, hematopoiesis, cardiac hypertrophy, and many cancer-related processes, including proliferation, apoptosis, differentiation, migration, and metabolism [97,98].

Many research groups have used biological and solid-state nanopores to detect miR-NAs in different tissues. For example, Meni et al. [99] used nanopore technology for rapid detection of probe-specific miRNAs (miRNA-122a and miRNA-153) and proved the potential of this approach by detecting liver-specific miRNAs at microgram levels from rat liver. Wang et al. [100] selectively detected single-molecule miRNAs in plasma samples of lung cancer patients using α-hemolysin-based nanopore sensors. The sensor uses programmable oligonucleotide probes to generate target-specific signals that quantify sub-millimolar levels of cancer-associated miRNAs.

These methods have potential application value for quantitative miRNA detection, detection of disease markers, and non-invasive early diagnosis of cancer. Also, nanopore gene sequencing method enables researchers to investigate overexpression mechanism of many kinds of miRNA.

5.3. Detection of Viruses and Bacteria

5.3.1. Monitor Virus Using Nanopore Sequencing Technology

During a pandemic, gene sequencing is an ideal way to determine the source of infection and the rate at which it evolves. Gene sequencing technology can also be used to identify characteristics of host fitness, identify and detect diagnostic targets, and characterize responses to vaccines and treatments [67,101].

Human beings' struggling against viruses is an evolving process. In the face of Ebola virus in West Africa, some scientists proposed to use nanopore sequencing technology to quickly detect virus samples in the environment, so as to help identify virus mutation types. The sequencing device, MinION, was used. In 2016, the Zika virus outbreak was declared an international public health emergency by the World Health Organization. The Zika in Brazil Real-time Analysis project was established to sequence a thousand of genomes from Brazil to monitor the epidemiological information [102]. The whole genome of Zika virus was obtained from clinical samples using the MinION sequencer for bioinformatics analysis [40]. MinION was used under these circumstances as it has many advantages. It weighs less than 100 g and is easy to operate in developing countries where the experimental condition cannot be achieved. Another advantage is that nanopore sequencing does not require precise microscope alignment and repeated calibration in comparison to the Illumina platform [103] and, hence, make the operation of MinION much simpler.

In view of the ongoing global epidemic of the 2019-nCoV virus, many teams used nanopore sequencing technology to explore and detect the novel coronavirus [104–106]. Among them, Wang et al. [105] proposed a nanopore targeted sequencing (NTS) technology to detect SARS-CoV-2 and other respiratory viruses simultaneously within 6–10 h. Researchers proved that NTS can effectively monitor the mutation of nucleic acid sequences, categorize the type of SARS-CoV-2 and other respiratory viruses in the samples.

In the future, with the rapid sequencing devices of nanopore sequencing, it is promising to establish virus surveillance and further establish rapid responsive system to track and monitor the diseases that threaten public health.

5.3.2. Study on Bacteria Using Nanopore Sequencing Technology

Nanopore technology could also be used for investigating the bacteria. The traditional method for bacteria research is to study the 16S rRNA of bacteria [107]. Brandt et al. [68] provide a promising method for monitoring the abundance of community present in microbial anaerobic bacterial communities by using metagenomic reading classification of nanopores, which can be used as an alternative method for bacterial taxonomy using 16S rRNA.

In recent years, the antibiotic resistance of bacteria has become a threat to aquatic and terrestrial biodiversity [108]. Also, it is crucial to investigate the resistance gene in bacteria with the gene sequencing method. Zhang et al. [109] used Nanopore sequencing platform and Illumina sequencing platform to quantitatively study the major antibiotic resistance gene types, showing that nanopore sequencing technology could be used for quantitative study of bacterial resistance. The result is conducive to promoting the research on antibiotic resistance gene transfer in sewage treatment plants and promoting the solution of the hot issue of antibiotic failure [110].

5.4. Assembly of Genome

Due to the characteristics of high throughput and long read-length, nanopore sequencing technology has more advantages in assembling genomes compared to the traditional short-read-long method. It can produce a more continuous genome assembly in the process of whole-genome assembly by crossing highly repetitive regions and regions containing structural variation [111,112].

Figure 6 shows the de novo assembly of chromosomes, which is the process of determining the order of each base in the genome by reconstructing randomly sampled sequence fragments [6]. The first step is to get sequence fragments (reads) by using the nanopore gene sequencing device. Splicing the short read according to the overlap between reads, the longer continuous sequences (contigs) are formed. Then the scaffold can be formed by joining the contigs together. After eliminating the error of scaffolds and gaps in the scaffold, a high-quality whole genome sequence can be obtained.
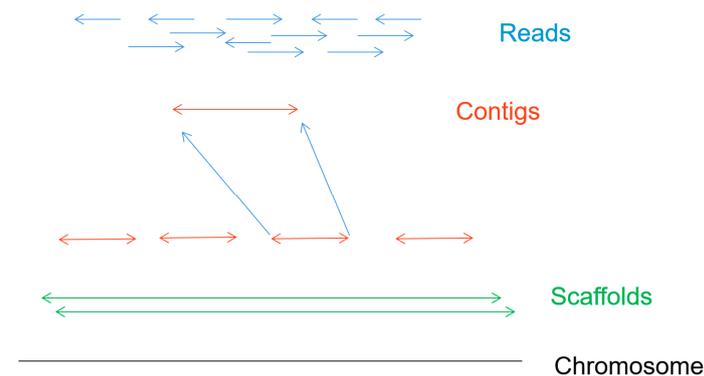
Figure 6. The process of assembly of chromosomes.

Readings from the same region of the genome can be linked together to form overlapping clusters. When using short-read sequencing technology, the repeated lengths exceeded the overlapping lengths can cause ambiguous reconstructions and fragmentation of assembly fragments. In comparison, nanopore sequencing technology can solve this problem by taking advantage of the ultra-long reading length.

To assemble the diploid genomes, such as human genomes, the challenge is to achieve accurate haplotype solutions from telomere to telomere without reference. Many studies have pointed out that short-read sequencing technology is insufficient to traverse most of the repeating structures of the genome [113]. In comparison, nanopore sequencing technology is promising with the ultra-long reading length [6,114]. An example is the sequenced genomes in GenBank. At the beginning of 2015, the major genome data was mainly achieved by short read gene sequencing technique, with only 99 mammalian genome combinations. The average overlap group N50 was only 41 kb [115]. Benefiting from the long-read sequencing technology such as PacBio and ONT by the year 2020, the average overlap group N50, whose length is more than 800 genomes, in the GenBank was greater than 5 Mb [23].

6. Conclusions

Detecting the potential change when charged biological molecules smaller than nanopores pass through the hole, nanopore technology provides a new way to identify and quantify a variety of analytes. Nanopore sequencing technology combines nanopore technology and biosensors, and it will have a tremendous impact in the area of gene sequencing.

Nanopore sequencing technology has broad application prospects in gene modification, epigenetics research, detecting microbes, and other related fields. It also plays an increasingly important role in human cancer diagnosis and other medical tests. Also, if the problem of high error rate can be solved and the ONT devices can reach the same level of accuracy, fourth-generation sequencing technology can take the place of the second-generation technology, which will dominate the market in the near future. In the future, nanopore sequencing technology will be continuously optimized from the structure of nanopore, machine algorithm of basecalling, and new methods of library preparation. It will lead human beings to continuously explore new types of gene variation and help people better solve the problems encountered in modern medicine.

At the same time, in the field of life science research, more reference genomes will be constructed, nanopore technology will be embedded into more kinds of devices and further broaden the understanding of genetics, epigenetics and transcription, variation, and its relationship to human phenotypes.

Author Contributions: Investigation and resources, B.L. and J.H.; writing-original draft preparation, B.L. and J.H.; writing-review and editing, B.L., J.H., and H.M.; visualization, B.L. and J.H.; supervision, H.M.; project administration, H.M.; funding acquisition, H.M. and J.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Deamer, D.W.; Akeson, M. Nanopores and nucleic acids: Prospects for ultrarapid sequencing. *Trends Biotechnol.* **2000**, *18*, 147–151. [CrossRef]
2. Rodriguez-Larrea, D. Single-aminoacid discrimination in proteins with homogeneous nanopore sensors and neural networks. *Biosens. Bioelectron.* **2021**, *180*, 7. [CrossRef] [PubMed]
3. Branton, D.; Deamer, D.W.; Marziali, A.; Bayley, H.; Benner, S.A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X.H.; et al. The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **2008**, *26*, 1146–1153. [CrossRef] [PubMed]
4. Zhu, X.Q.; Li, J.; He, H.P.; Huang, M.; Zhang, X.H.; Wang, S.F. Application of nanomaterials in the bioanalytical detection of disease-related genes. *Biosens. Bioelectron.* **2015**, *74*, 113–133. [CrossRef]
5. Sanger, F.; Air, G.M.; Barrell, B.G.; Brown, N.L.; Coulson, A.R.; Fiddes, J.C.; Hutchison, C.A.; Slocombe, P.M.; Smith, M. Nucleotide-sequence of bacteriophage Phichi174 DNA. *Nature* **1977**, *265*, 687–695. [CrossRef]
6. Lander, E.S.E.A. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [CrossRef]
7. Sanger, F.; Nicklen, S.; Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5463–5467. [CrossRef]
8. Smith, L.; Sanders, J.; Kaiser, R.; Hughes, P.; Dodd, C.; Connell, C.; Heiner, C.; Kent, S.; Hood, L. Fluorescence detection in Automated DNA sequence analysis. *Nature* **1986**, *321*, 674–679. [CrossRef] [PubMed]
9. Connell, C.; Fung, S.; Heiner, C.; Bridgham, J.; Chakerian, V.; Heron, E.; Jones, B.; Menchen, S.; Mordan, W.; Raff, M.; et al. Automated DNA sequence analysis. *BioTechniques* **1987**, *5*, 342–348.
10. Ronaghi, M.; Karamohamed, S.; Pettersson, B.; Uhlen, M.; Nyrén, P. Real-Time DNA sequencing using detection of pyrophosphate release. *Anal. Biochem.* **1996**, *242*, 84–89. [CrossRef]
11. Margulies, M.; Egholm, M.; Altman, W.E.; Attiya, S.; Bader, J.S.; Bemben, L.A.; Berka, J.; Braverman, M.S.; Chen, Y.J.; Chen, Z.; et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **2005**, *437*, 376–380. [CrossRef]

12. McCarthy, A. Third Generation DNA sequencing: Pacific biosciences' single molecule real time technology. *Chem. Biol.* **2010**, *17*, 675–676. [CrossRef] [PubMed]

13. Maxam, A.M.; Gilbert, W.J.B. A new method for sequencing DNA. 1977. *Biotechnology* **1992**, *24*, 99–103.

14. Mardis, E.R. Next-generation DNA sequencing methods. *Annu. Rev. Genom. Hum. Genet.* **2008**, *9*, 387–402. [CrossRef]

15. Glenn, T.C. Field guide to next-generation DNA sequencers. *Mol. Ecol. Resour.* **2011**, *11*, 759–769. [CrossRef] [PubMed]

16. Huang, W.; Li, L.; Myers, J.R.; Marth, G.T. ART: A next-generation sequencing read simulator. *Bioinformatics* **2012**, *28*, 593–594. [CrossRef] [PubMed]

17. Fang, C.; Zhong, H.; Lin, Y.; Chen, B.; Han, M.; Ren, H.; Lu, H.; Luber, J.M.; Xia, M.; Li, W.; et al. Assessment of the cPAS-based BGISEQ-500 platform for metagenomic sequencing. *Gigascience* **2018**, *7*, 1–8. [CrossRef]

18. Mikheyev, A.S.; Tin, M.M.Y. A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.* **2014**, *14*, 1097–1102. [CrossRef]

19. Miga, K.H.; Koren, S.; Rhie, A.; Vollger, M.R.; Gershman, A.; Bzikadze, A.; Brooks, S.; Howe, E.; Porubsky, D.; Logsdon, G.A.; et al. Telomere-to-telomere assembly of a complete human X chromosome. *Nature* **2020**, *585*, 79–84. [CrossRef]

20. Sun, X.; Song, L.; Yang, W.; Zhang, L.; Liu, M.; Li, X.; Tian, G.; Wang, W. Nanopore sequencing and its clinical applications. *Methods Mol. Biol.* **2020**, *2204*, 13–32. [CrossRef]

21. Patel, N.; Ferns, B.R.; Nastouli, E.; Kozlakidis, Z.; Kellam, P.; Morris, S. Cost analysis of standard Sanger sequencing versus next generation sequencing in the ICONIC study. *Lancet* **2016**, *388*. [CrossRef]

22. Ameur, A.; Kloosterman, W.P.; Hestand, M.S. Single-Molecule Sequencing: Towards Clinical Applications. *Trends Biotechnol.* **2019**, *37*, 72–85. [CrossRef] [PubMed]

23. Logsdon, G.A.; Vollger, M.R.; Eichler, E.E. Long-read human genome sequencing and its applications. *Nat. Rev. Genet.* **2020**, *21*, 597–614. [CrossRef] [PubMed]

24. Feng, Y.; Zhang, Y.; Ying, C.; Wang, D.; Du, C. Nanopore-based fourth-generation DNA sequencing technology. *Genom. Proteom. Bioinform.* **2015**, *13*, 4–16. [CrossRef] [PubMed]

25. 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. [CrossRef]

26. Ma, Q.; Si, Z.X.; Li, Y.; Wang, D.G.; Wu, X.L.; Gao, P.C.; Xia, F. Functional solid-state nanochannels for biochemical sensing. *Trac. Trends Anal. Chem.* **2019**, *115*, 174–186. [CrossRef]

27. Khan, M.S.; Dosoky, N.S.; Berdiev, B.K.; Williams, J.D. Electrochemical impedance spectroscopy for black lipid membranes fused with channel protein supported on solid-state nanopore. *Eur. Biophys. J. Biophys. Lett.* **2016**, *45*, 843–852. [CrossRef]

28. Tang, Z.F.; Zhang, D.H.; Cui, W.W.; Zhang, H.; Pang, W.; Duan, X.X. Fabrications, applications and challenges of solid-state nanopores: A mini review. *Nanomater. Nanotechnol.* **2016**, *6*, 12. [CrossRef]

29. Kim, M.J.; McNally, B.; Murata, K.; Meller, A. Characteristics of solid-state nanometre pores fabricated using a transmission electron microscope. *Nanotechnology* **2007**, *18*, 5. [CrossRef]

30. Mohammad, M.M.; Iyer, R.; Howard, K.R.; McPike, M.P.; Borer, P.N.; Movileanu, L. Engineering a rigid protein tunnel for biomolecular detection. *J. Am. Chem. Soc.* **2012**, *134*, 9521–9531. [CrossRef]

31. Bayley, H.; Cremer, P.S. Stochastic sensors inspired by biology. *Nature* **2001**, *413*, 226–230. [CrossRef] [PubMed]

32. Derrington, I.M.; Butler, T.Z.; Collins, M.D.; Manrao, E.; Pavlenok, M.; Niederweis, M.; Gundlach, J.H. Nanopore DNA sequencing with MspA. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16060–16065. [CrossRef]

33. 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. [CrossRef]

34. Brown, E.; Freimanis, G.; Shaw, A.E.; Horton, D.L.; Gubbins, S.; King, D. Characterising foot-and-mouth disease virus in clinical samples using nanopore sequencing. *Front. Vet. Sci.* **2021**, *8*, 10. [CrossRef]

35. Quick, J.; Loman, N.J.; Duraffour, S.; Simpson, J.T.; Ettore, S.; Cowley, L.; Bore, J.A.; Koundouno, R.; Dudas, G.; Mikhail, A.; et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* **2016**, *530*, 228–232. [CrossRef]

36. Hu, Z.L.; Huo, M.Z.; Ying, Y.L.; Long, Y.T. Biological nanopore approach for single-molecule protein sequencing. *Angew. Chem. Int. Edit.* **2020**, *133*. [CrossRef]

37. Bentley, D.L. Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* **2014**, *15*, 163–175. [CrossRef]

38. Oxford Nanopore Technologies. Introduction to Real Time Analysis. 2020. Available online: https://www.youtube.com/watch?v=8oNEjt5Ov_Q (accessed on 2 April 2021).

39. Wick, R.R.; Judd, L.M.; Holt, K.E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* **2019**. [CrossRef]

40. Kono, N.; Arakawa, K. Nanopore sequencing: Review of potential applications in functional genomics. *Dev. Growth Differ.* **2019**, *61*, 316–326. [CrossRef] [PubMed]

41. Rang, F.J.; Kloosterman, W.P.; de Ridder, J. From squiggle to basepair: Computational approaches for improving nanopore sequencing read accuracy. *Genome Biol.* **2018**, *19*, 11. [CrossRef] [PubMed]

42. Karst, S.M.; Ziels, R.M.; Kirkegaard, R.H.; Sørensen, E.A.; McDonald, D.; Zhu, Q.; Knight, R.; Albertsen, M. High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing. *Nat. Methods* **2021**, *18*, 165–169. [CrossRef] [PubMed]

43. Li, C.; Chng, K.R.; Boey, E.J.; Ng, A.H.; Wilm, A.; Nagarajan, N. INC-Seq: Accurate single molecule reads using nanopore sequencing. *Gigascience* **2016**, *5*, 34. [CrossRef]

44. van Dijk, E.L.; Jaszczyszyn, Y.; Naquin, D.; Thermes, C. The third revolution in sequencing technology. *Trends Genet.* **2018**, *34*, 666–681. [CrossRef]

45. Eddy, S.R. What is a hidden Markov model? *Nat. Biotechnol.* **2004**, *22*, 1315–1316. [CrossRef] [PubMed]

46. Williams, R.J.; Zipser, D. A learning algorithm for continually running fully recurrent neural networks. *Neural Comput.* **1989**, *1*, 270–280. [CrossRef]

47. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]

48. Meng, J.; Victor, B.; He, Z.; Liu, H.; Jiang, T. DeepSSV: Detecting somatic small variants in paired tumor and normal sequencing data with convolutional neural network. *Brief. Bioinform.* **2020**. [CrossRef] [PubMed]

49. Graves, A.; Fernández, S.; Gomez, F.; Schmidhuber, J. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA, 25–29 June 2006; pp. 369–376.

50. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* **2018**, *4*, e00938. [CrossRef]

51. David, M.; Dursi, L.J.; Yao, D.; Boutros, P.C.; Simpson, J.T. Nanocall: An open source basecaller for Oxford Nanopore sequencing data. *Bioinformatics* **2017**, *33*, 49–55. [CrossRef]

52. Huang, N.; Nie, F.; Ni, P.; Luo, F.; Wang, J. SACall: A neural network basecaller for Oxford Nanopore sequencing data based on self-attention mechanism. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2020**. [CrossRef]

53. Boza, V.; Brejova, B.; Vinar, T. DeepNano: Deep recurrent neural networks for base calling in MinION nanopore reads. *PLoS ONE* **2017**, *12*, e0178751. [CrossRef] [PubMed]

54. Teng, H.; Minh Duc, C.; Hall, M.B.; Duarte, T.; Wang, S.; Coin, L.J.M. Chiron: Translating nanopore raw signal directly into nucleotide sequence using deep learning. *Gigascience* **2018**, *7*. [CrossRef] [PubMed]

55. Stoiber, M.; Brown, J. BasecRAWller: Streaming nanopore basecalling directly from raw signal. *BioRxiv* **2017**, 133058. [CrossRef]

56. Chan, W.S.; Chan, T.L.; Au, C.H.; Leung, C.P.; To, M.Y.; Ng, M.K.; Leung, S.M.; Chan, M.K.M.; Ma, E.S.K.; Tang, B.S.F. An economical Nanopore sequencing assay for human papillomavirus (HPV) genotyping. *Diagn. Pathol.* **2020**, *15*, 18. [CrossRef] [PubMed]

57. 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**, 9. [CrossRef]

58. Wu, X.W.; Luo, H.; Xu, F.; Ge, C.T.; Li, S.T.; Deng, X.Y.; Wiedmann, M.; Baker, R.C.; Stevenson, A.; Zhang, G.T.; et al. Evaluation of salmonella serotype prediction with multiplex nanopore sequencing. *Front. Microbiol.* **2021**, *12*, 13. [CrossRef]

59. Oxford Nanopore Technologies. R10.3: The Newest Nanopore for High Accuracy Nanopore Sequencing—Now Available in Store. 2020. Available online: https://nanoporetech.com/about-us/news/r103-newest-nanopore-high-accuracy-nanopore-sequencing-now-available-store (accessed on 13 January 2020).

60. Lee, I.; Razaghi, R.; Gilpatrick, T.; Molnar, M.; Gershman, A.; Sadowski, N.; Sedlazeck, F.J.; Hansen, K.D.; Simpson, J.T.; Timp, W. Simultaneous profiling of chromatin accessibility and methylation on human cell lines with nanopore sequencing. *Nat. Methods* **2020**, *17*, 1191–1199. [CrossRef]

61. Zalvidea, D.; Claverol-Tinture, E. Second Harmonic Generation for time-resolved monitoring of membrane pore dynamics subserving electroporation of neurons. *Biomed. Opt. Express* **2011**, *2*, 305–314. [CrossRef]

62. Zheng, G.X.Y.; Lau, B.T.; Schnall-Levin, M.; Jarosz, M.; Bell, J.M.; Hindson, C.M.; Kyriazopoulou-Panagiotopoulou, S.; Masquelier, D.A.; Merrill, L.; Terry, J.M.; et al. Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. *Nat. Biotechnol.* **2016**, *34*, 303–311. [CrossRef]

63. Williams, M.S.; Basma, N.J.; Amaral, F.M.R.; Williams, G.; Weightman, J.P.; Breitwieser, W.; Nelson, L.; Taylor, S.S.; Wiseman, D.H.; Somervaile, T.C.P. Targeted nanopore sequencing for the identification of ABCB1 promoter translocations in cancer. *BMC Cancer* **2020**, *20*, 9. [CrossRef]

64. Suzuki, A.; Suzuki, M.; Mizushima-Sugano, J.; Frith, M.C.; Makalowski, W.; Kohno, T.; Sugano, S.; Tsuchihara, K.; Suzuki, Y. Sequencing and phasing cancer mutations in lung cancers using a long-read portable sequencer. *DNA Res.* **2017**, *24*, 585–596. [CrossRef]

65. Libermann, T.A.; Zerbini, L.F. Targeting transcription factors for cancer gene therapy. *Curr. Gene Ther.* **2006**, *6*, 17–33. [CrossRef]

66. Ahumada-Garcia, R.; Gonzalez-Puelma, J.; Alvarez-Saravia, D.; Barrientos, R.J.; Uribe-Paredes, R.; Lopez-Cortes, X.A.; Navarrete, M.A. Identification of immunoglobulin gene usage in immune repertoires sequenced by nanopore technology. In *Bioinformatics and Biomedical Engineering, Iwbbio 2019, Pt I*; Rojas, I., Valenzuela, O., Rojas, F., Ortuno, F., Eds.; Lecture Notes in Bioinformatics; Springer International Publishing Ag: Cham, Switzerland, 2019; Volume 11465, pp. 295–306.

67. Tan, S.Y.; Dvorak, C.M.T.; Murtaugh, M.P. Characterization of Emerging swine viral diseases through oxford nanopore sequencing using senecavirus A as a model. *Viruses Basel.* **2020**, *12*, 1136. [CrossRef]

68. Brandt, C.; Bongcam-Rudloff, E.; Muller, B. Abundance tracking by long-read nanopore sequencing of complex microbial communities in samples from 20 different Biogas/Wastewater plants. *Appl. Sci. Basel* **2020**, *10*, 7518. [CrossRef]

69. Porubsky, D.; Ebert, P.; Audano, P.A.; Vollger, M.R.; Harvey, W.T.; Marijon, P.; Ebler, J.; Munson, K.M.; Sorensen, M.; Sulovari, A.; et al. Fully phased human genome assembly without parental data using single-cell strand sequencing and long reads. *Nat. Biotechnol.* **2020**. [CrossRef]

70. Peters, B.A.; Kermani, B.G.; Sparks, A.B.; Alferov, O.; Hong, P.; Alexeev, A.; Jiang, Y.; Dahl, F.; Tang, Y.T.; Haas, J.; et al. Accurate whole-genome sequencing and haplotyping from 10 to 20 human cells. *Nature* **2012**, *487*, 190–195. [CrossRef] [PubMed]

71. Loman, N.J.; Quick, J.; Simpson, J.T. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat. Methods* **2015**, *12*, 733–735. [CrossRef] [PubMed]

72. Reddington, K.; Eccles, D.; O'Grady, J.; Drown, D.M.; Hansen, L.H.; Nielsen, T.K.; Ducluzeau, A.L.; Leggett, R.M.; Heavens, D.; Peel, N.; et al. Metagenomic analysis of planktonic riverine microbial consortia using nanopore sequencing reveals insight into river microbe taxonomy and function. *Gigascience* **2020**, *9*, 12. [CrossRef]

73. Pomerantz, A.; Penafiel, N.; Arteaga, A.; Bustamante, L.; Pichardo, F.; Coloma, L.A.; Barrio-Amoros, C.L.; Salazar-Valenzuela, D.; Prost, S. Real-time DNA barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building. *Gigascience* **2018**, *7*, 14. [CrossRef] [PubMed]

74. Krehenwinkel, H.; Pomerantz, A.; Prost, S. Genetic biomonitoring and biodiversity assessment using portable sequencing technologies: Current uses and future directions. *Genes* **2019**, *10*, 858. [CrossRef] [PubMed]

75. Mijuskovic, M.; Brown, S.M.; Tang, Z.; Lindsay, C.R.; Efstathiadis, E.; Deriano, L.; Roth, D.B. A streamlined method for detecting structural variants in cancer genomes by short read paired-end sequencing. *PLoS ONE* **2012**, *7*, e48314. [CrossRef]

76. Norris, A.L.; Workman, R.E.; Fan, Y.F.; Eshleman, J.R.; Timp, W. Nanopore sequencing detects structural variants in cancer. *Cancer Biol. Ther.* **2016**, *17*, 246–253. [CrossRef] [PubMed]

77. Darnell, J.E. Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2002**, *2*, 740–749. [CrossRef] [PubMed]

78. Squires, A.; Atas, E.; Meller, A. Nanopore sensing of individual transcription factors bound to DNA. *Sci. Rep.* **2015**, *5*. [CrossRef]

79. Hou, C.; Tian, Y.; Wang, Y.; Lian, H.; Liang, D.; Shi, S.; Deng, N.; He, B. Revealing the developmental dynamics in male strobilus transcriptome of Gnetum luofuense using nanopore sequencing technology. *Sci. Rep.* **2021**, *11*, 10516. [CrossRef] [PubMed]

80. Boti, M.A.; Adamopoulos, P.G.; Tsiakanikas, P.; Scorilas, A. Nanopore sequencing unveils diverse transcript variants of the epithelial cell-specific transcription factor Elf-3 in human malignancies. *Genes* **2021**, *12*, 839. [CrossRef]

81. Ding, Y.; Fleming, A.M.; He, L.; Burrows, C.J. Unfolding Kinetics of the Human Telomere i-Motif Under a 10 pN Force Imposed by the alpha-Hemolysin Nanopore Identify Transient Folded-State Lifetimes at Physiological pH. *J. Am. Chem. Soc.* **2015**, *137*, 9053–9060. [CrossRef]

82. Neidle, S.; Read, M.A. G-quadruplexes as therapeutic targets. *Biopolymers* **2000**, *56*, 195–208. [CrossRef]

83. Neidle, S. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J.* **2010**, *277*, 1118–1125. [CrossRef]

84. Sánchez-Romero, M.A.; Cota, I.; Casadesús, J. DNA methylation in bacteria: From the methyl group to the methylome. *Curr. Opin. Microbiol.* **2015**, *25*, 9–16. [CrossRef]

85. Vu, T.; Davidson, S.-L.; Borgesi, J.; Maksudul, M.; Jeon, T.-J.; Shim, J. Piecing together the puzzle: Nanopore technology in detection and quantification of cancer biomarkers. *RSC Adv.* **2017**, *7*, 42653–42666. [CrossRef]

86. Lyko, F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* **2018**, *19*, 81–92. [CrossRef]

87. Baylin, S.B.; Jones, P.A. A decade of exploring the cancer epigenome—Biological and translational implications. *Nat. Rev. Cancer* **2011**, *11*, 726–734. [CrossRef]

88. Jones, P.A.; Baylin, S.B. The epigenomics of cancer. *Cell* **2007**, *128*, 683–692. [CrossRef] [PubMed]

89. Shim, J.; Kim, Y.; Humphreys, G.I.; Nardulli, A.M.; Kosari, F.; Vasmatzis, G.; Taylor, W.R.; Ahlquist, D.A.; Myong, S.; Bashir, R. Nanopore-based assay for detection of methylation in double-stranded DNA fragments. *Acs. Nano* **2015**, *9*, 290–300. [CrossRef] [PubMed]

90. Murrell, A.; Rakyan, V.K.; Beck, S. From genome to epigenome. *Hum. Mol. Genet.* **2005**, *14*, R3–R10. [CrossRef] [PubMed]

91. Trinh, B.N.; Long, T.I.; Laird, P.W. DNA Methylation analysis by MethyLight technology. *Methods* **2001**, *25*, 456–462. [CrossRef]

92. Shim, J.; Humphreys, G.I.; Venkatesan, B.M.; Munz, J.M.; Zou, X.; Sathe, C.; Schulten, K.; Kosari, F.; Nardulli, A.M.; Vasmatzis, G.; et al. Detection and quantification of methylation in DNA using solid-state nanopores. *Sci. Rep.* **2013**, *3*. [CrossRef] [PubMed]

93. Rauf, S.; Zhang, L.; Ali, A.; Ahmad, J.; Liu, Y.; Li, J.H. Nanopore-Based, Label-Free, and Real-Time Monitoring Assay for DNA Methyltransferase Activity and Inhibition. *Anal. Chem.* **2017**, *89*, 13252–13260. [CrossRef] [PubMed]

94. Martinez, B.; Peplow, P.V. Altered microRNA expression in animal models of Huntington's disease and potential therapeutic strategies. *Neural Regen. Res.* **2021**, *16*, 2159–2169. [CrossRef] [PubMed]

95. Garzon, R.; Calin, G.A.; Croce, C.M. MicroRNAs in Cancer. *Annu. Rev. Med.* **2009**, *60*, 167–179. [CrossRef] [PubMed]

96. Cecile, O.; Marie-Pierre, P.; Marius, I.; Pascal, B.; Bernard, M.; Paul, H. MicroRNAs and lung cancer: New oncogenes and tumor suppressors, new prognostic factors and potential therapeutic targets. *Curr. Med. Chem.* **2009**, *16*, 1047–1061.

97. Winter, J.; Jung, S.; Keller, S.; Gregory, R.I.; Diederichs, S. Many roads to maturity: MicroRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **2009**, *11*, 228–234. [CrossRef] [PubMed]

98. Carthew, R.W.; Sontheimer, E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–655. [CrossRef] [PubMed]

99. Wanunu, M.; Dadosh, T.; Ray, V.; Jin, J.; McReynolds, L.; Drndic, M. Rapid electronic detection of probe-specific microRNAs using thin nanopore sensors. *Nat. Nanotechnol.* **2010**, *5*, 807–814. [CrossRef]

100. Wang, Y.; Zheng, D.L.; Tan, Q.L.; Wang, M.X.; Gu, L.Q. Nanopore-based detection of circulating microRNAs in lung cancer patients. *Nat. Nanotechnol.* **2011**, *6*, 668–674. [CrossRef] [PubMed]
101. Kwak, D.K.; Kim, J.S.; Lee, M.K.; Ryu, K.S.; Chi, S.W. Probing the neuraminidase activity of influenza virus using a cytolysin a protein nanopore. *Anal. Chem.* **2020**, *92*, 14303–14308. [CrossRef]
102. De Jesus, J.G.; Giovanetti, M.; Faria, N.R.; Alcantara, L.C. Acute vector-borne viral infection:Zika and MinION surveillance. *Microbiol. Spectr.* **2019**, *7*, 11. [CrossRef]
103. Kugelman, J.R.; Wiley, M.R.; Mate, S.; Ladner, J.T.; Beitzel, B.; Fakoli, L.; Taweh, F.; Prieto, K.; Diclaro, J.W.; Minogue, T.; et al. Monitoring of ebola virus makona evolution through establishment of advanced genomic capability in liberia. *Emerg. Infect. Diseases* **2015**, *21*, 1135–1143. [CrossRef]
104. Chan, W.M.; Ip, J.D.; Chu, A.W.H.; Yip, C.C.Y.; Lo, L.S.; Chan, K.H.; Ng, A.C.K.; Poon, R.W.S.; To, W.K.; Tsang, O.T.Y.; et al. Identification of nsp1 gene as the target of SARS-CoV-2 real-time RT-PCR using nanopore whole-genome sequencing. *J. Med. Virol.* **2020**, *92*, 2725–2734. [CrossRef]
105. Wang, M.; Fu, A.S.; Hu, B.; Tong, Y.Q.; Liu, R.; Liu, Z.; Gu, J.S.; Xiang, B.; Liu, J.H.; Jiang, W.; et al. Nanopore targeted sequencing for the accurate and comprehensive detection of SARS-CoV-2 and other respiratory viruses. *Small* **2020**, *16*, 15. [CrossRef]
106. Shepherd, B.A.; Tanjil, M.R.E.; Jeong, Y.J.; Baloglu, B.; Liao, J.Q.; Wang, M.C. Angstrom ngstrom- and Nano-scale pore-based nucleic acid sequencing of current and emergent pathogens. *MRS Adv.* **2020**, *5*, 2889–2906. [CrossRef]
107. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072. [CrossRef]
108. Fiaz, A.; Zhu, D.C.; Sun, J.Z. Environmental fate of tetracycline antibiotics: Degradation pathway mechanisms, challenges, and perspectives. *Environ. Sci. Eur.* **2021**, *33*, 17. [CrossRef]
109. Che, Y.; Xia, Y.; Liu, L.; Li, A.D.; Yang, Y.; Zhang, T. Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing. *Microbiome* **2019**, *7*, 13. [CrossRef] [PubMed]
110. Pehrsson, E.C.; Tsukayama, P.; Patel, S.; Mejía-Bautista, M.; Sosa-Soto, G.; Navarrete, K.M.; Calderon, M.; Cabrera, L.; Hoyos-Arango, W.; Bertoli, M.T.; et al. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* **2016**, *533*, 212–216. [CrossRef]
111. Michael, T.P.; Jupe, F.; Bemm, F.; Motley, S.T.; Sandoval, J.P.; Lanz, C.; Loudet, O.; Weigel, D.; Ecker, J.R. High contiguity Arabidopsis thaliana genome assembly with a single nanopore flow cell. *Nat. Commun.* **2018**, *9*, 541. [CrossRef] [PubMed]
112. Zhou, A.; Lin, T.; Xing, J. Evaluating nanopore sequencing data processing pipelines for structural variation identification. *Genome. Biol.* **2019**, *20*, 237. [CrossRef] [PubMed]
113. Kronenberg, Z.N.; Fiddes, I.T.; Gordon, D.; Murali, S.; Cantsilieris, S.; Meyerson, O.S.; Underwood, J.G.; Nelson, B.J.; Chaisson, M.J.P.; Dougherty, M.L.; et al. High-resolution comparative analysis of great ape genomes. *Science* **2018**, *360*, 1085. [CrossRef]
114. 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. [CrossRef]
115. Chaisson, M.J.P.; Wilson, R.K.; Eichler, E.E. Genetic variation and the de novo assembly of human genomes. *Nat. Rev. Genet.* **2015**, *16*, 627–640. [CrossRef] [PubMed]