Nanopore Sequencing Technology

Technology Principle

- DNA/RNA passes through protein nanopore
- Changes in electrical current identify bases
- Real-time sequencing - no synthesis required

Read Length

Ultra-long reads: up to 2 Mb

Average: 10-100 Kb

Accuracy

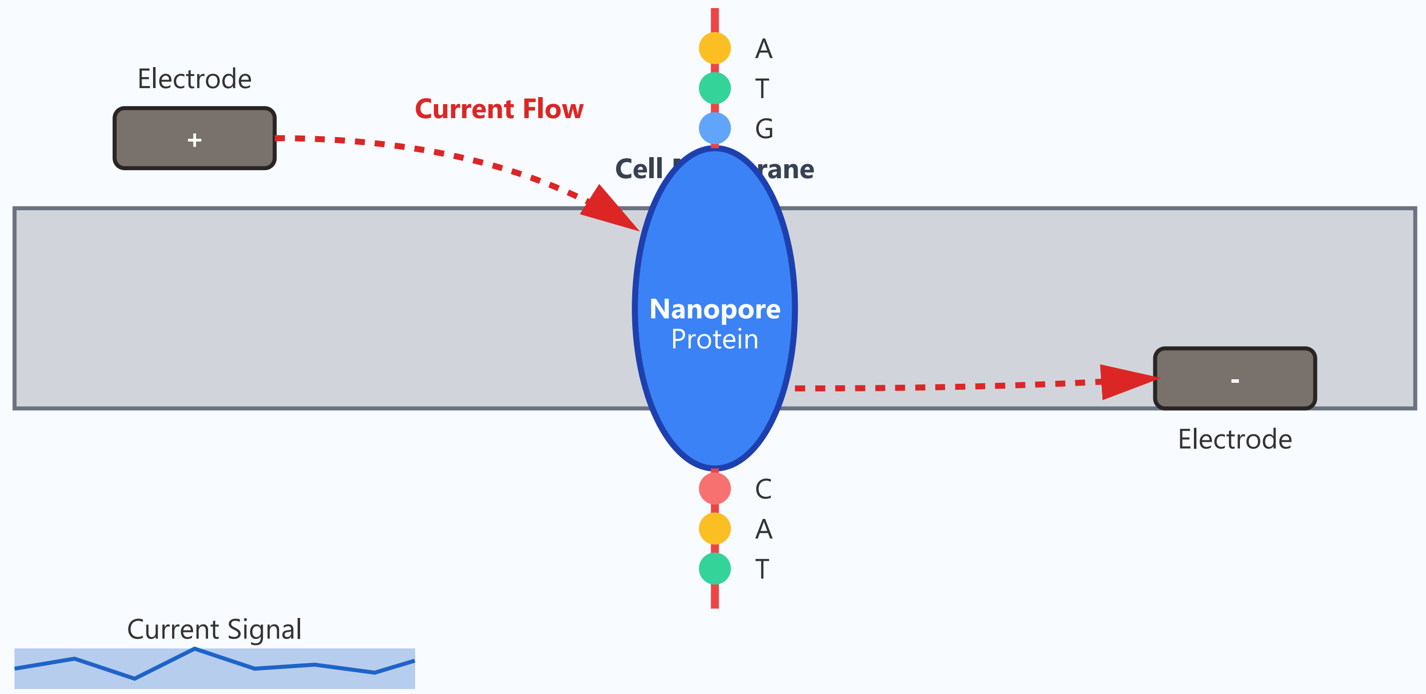
Raw: ~95%

With consensus: >99%

Key Features

- Portable device (MinION USB sequencer)
- Real-time data analysis
- Direct RNA sequencing without reverse transcription
- Detect base modifications natively
- Rapid sequencing for outbreak response

1. TECHNOLOGY PRINCIPLE - DETAILED EXPLANATION



How It Works

Nanopore sequencing is a revolutionary single-molecule sequencing technology that reads DNA or RNA sequences directly by measuring changes in electrical current as nucleic acids pass through a protein nanopore.

Step-by-Step Process:

1. A motor protein unwinds the double-stranded DNA and feeds single-stranded DNA through the nanopore at a controlled speed (approximately 450 bases per second).

2. An ionic current is applied across the membrane. As each nucleotide passes through the nanopore, it causes a characteristic disruption to the current.

3. Each of the four DNA bases (A, T, G, C) has a unique chemical structure that blocks the current differently, creating a distinctive electrical signature.

4. Advanced algorithms analyze these current changes in real-time to identify the sequence of bases passing through the pore.

Key Advantages of the Principle

→ No optical detection or fluorescent labels required

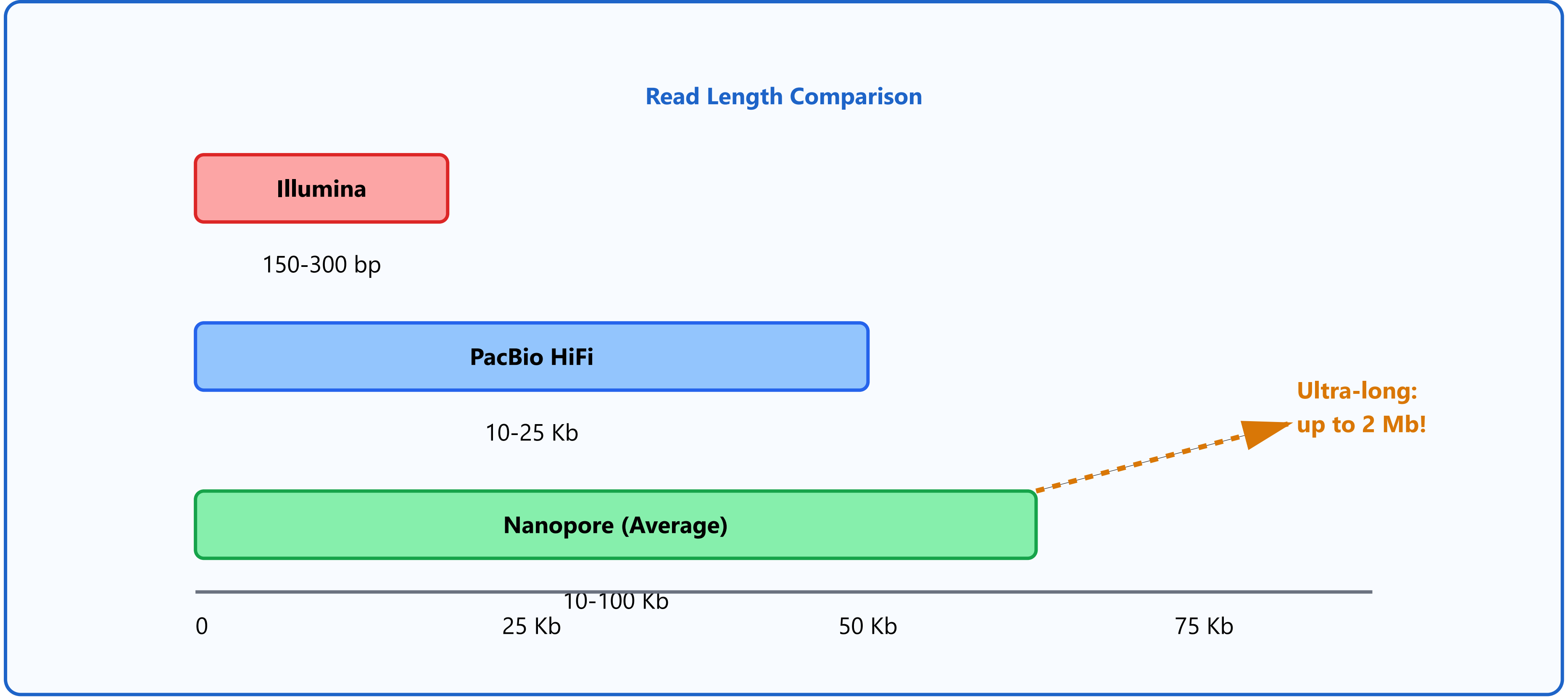
→ No amplification step needed - sequences native DNA/RNA directly

→ No theoretical limit on read length - depends only on DNA fragment length

→ Can detect modified bases (methylation, etc.) in their native form

→ Real-time data streaming enables immediate analysis

2. READ LENGTH CAPABILITIES



Why Read Length Matters

Read length is one of the most critical parameters in DNA sequencing, and nanopore technology excels in this area with dramatically longer reads compared to traditional methods.

Applications Enabled by Long Reads:

Application	Why Long Reads Help	Nanopore Advantage
De novo Assembly	Span repetitive regions that are longer than short reads	Can assemble entire bacterial genomes in single contigs

Structural Variants	Detect large deletions, insertions, inversions	Reads can span entire variant, making detection unambiguous
Phasing	Connect distant variants on same chromosome	Ultra-long reads can phase across megabases
Isoform Detection	Sequence full-length transcripts	Direct RNA sequencing of complete mRNA molecules
Repeat Analysis	Resolve complete repetitive regions	Can sequence through centromeres and telomeres

Record-Breaking Achievements

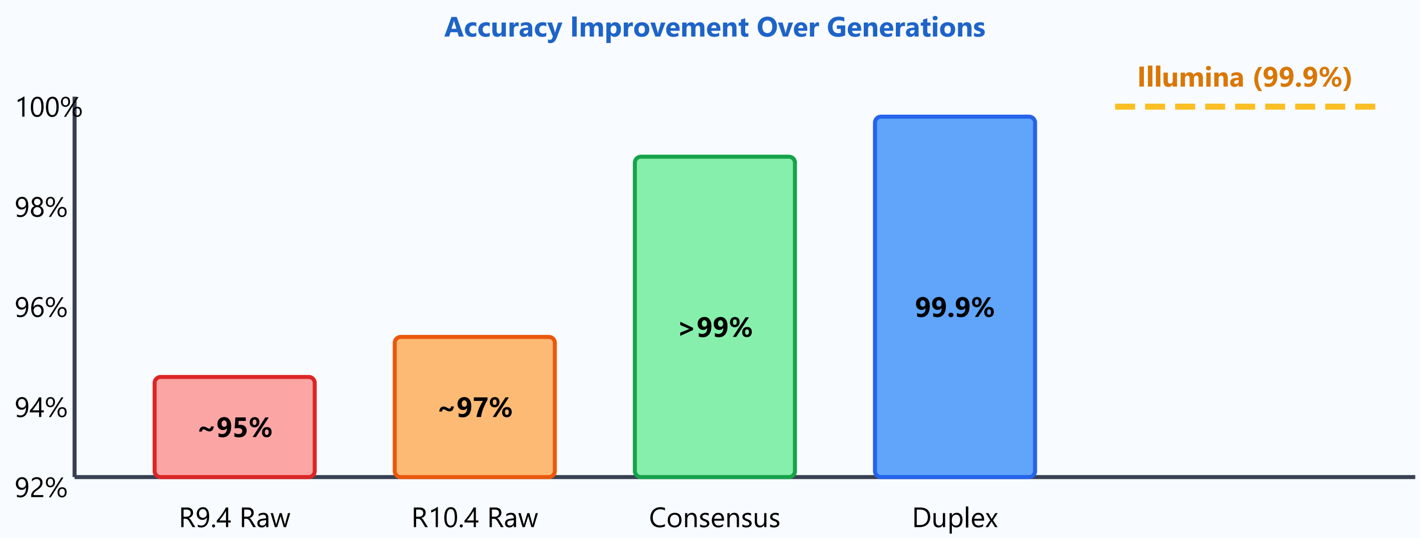
→ Longest single read: >4 Mb (megabases) reported

→ N50 values routinely exceed 50 Kb with proper DNA preparation

→ Ultra-long read protocols can achieve >100 Kb average read length

→ Enabled complete gapless assemblies of human chromosomes

3. SEQUENCING ACCURACY



Understanding Accuracy Metrics

Nanopore sequencing accuracy has dramatically improved over the years through better chemistry, improved nanopores, and advanced basecalling algorithms powered by deep learning.

Types of Accuracy:

1. Raw Read Accuracy (~95-97%): Single-pass sequencing of one DNA strand. The R10.4 chemistry with improved nanopores and basecallers like Dorado achieve ~97% raw accuracy.

2. Consensus Accuracy (>99%): Multiple reads of the same DNA molecule are combined using consensus algorithms. This requires higher coverage (typically 20-50×) but achieves accuracy comparable to short-read platforms.

3. Duplex Accuracy (99.9%): Both strands of the same DNA molecule are sequenced and compared. This provides the highest accuracy, matching Illumina's gold standard, while maintaining long read lengths.

Factors Affecting Accuracy

→ **Pore Chemistry:** R10.4 pores provide better resolution than earlier R9.4 versions

→ **Basecalling Algorithm:** Deep learning models (Guppy, Dorado) significantly improve accuracy

→ **Sequencing Speed:** Slower translocation gives more data points per base

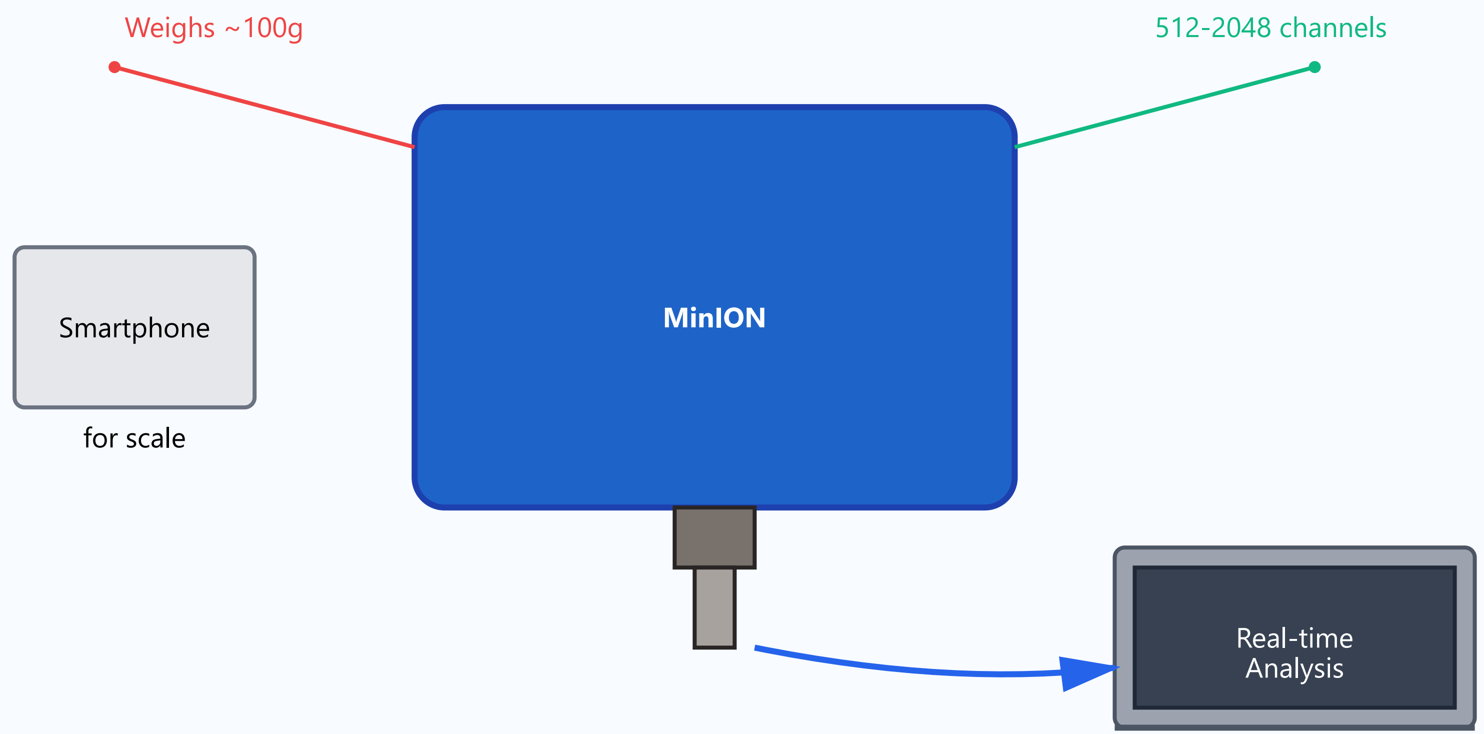
→ **DNA Quality:** High molecular weight, pure DNA yields better results

→ **Homopolymers:** Stretches of same base (e.g., AAAAA) remain challenging

Error Profile

Unlike Illumina sequencing which has systematic errors and GC bias, nanopore errors are more random and can be corrected through consensus. The error rate is higher in homopolymer regions but more uniform across different sequence contexts.

4. PORTABLE DEVICE - MINION USB SEQUENCER



Revolutionary Portability

The MinION is the world's first portable DNA sequencer, roughly the size of a USB thumb drive. This breakthrough in miniaturization has transformed where and how sequencing can be performed.

Device Specifications:

Feature	MinION	Traditional Sequencer
Size	10 cm × 3.2 cm × 2 cm	Desktop to room-sized

Weight	~100 grams	50-500+ kg
Cost	$1,000 device (flow cells $500-900)	$100,000 - $1,000,000+
Power	USB-powered (5W)	Dedicated power (1000W+)
Setup	Plug-and-play, <10 minutes	Specialized facility required
Throughput	Up to 50 Gb per flow cell	100 Gb - 6 Tb per run

Game-Changing Applications

→ **Field Research:** Used in Antarctic, rainforests, and remote locations for real-time pathogen detection

→ **Outbreak Response:** Deployed during Ebola, Zika, and COVID-19 outbreaks for rapid viral genome sequencing

→ **Clinical Settings:** Bedside sequencing for rapid diagnosis in ICUs and emergency situations

→ **Space Research:** First DNA sequencer used on the International Space Station (2016)

→ **Educational Use:** Makes sequencing accessible to teaching labs and small research groups

→ **Point-of-Care:** Enables sequencing in resource-limited settings without specialized infrastructure

Oxford Nanopore Product Line

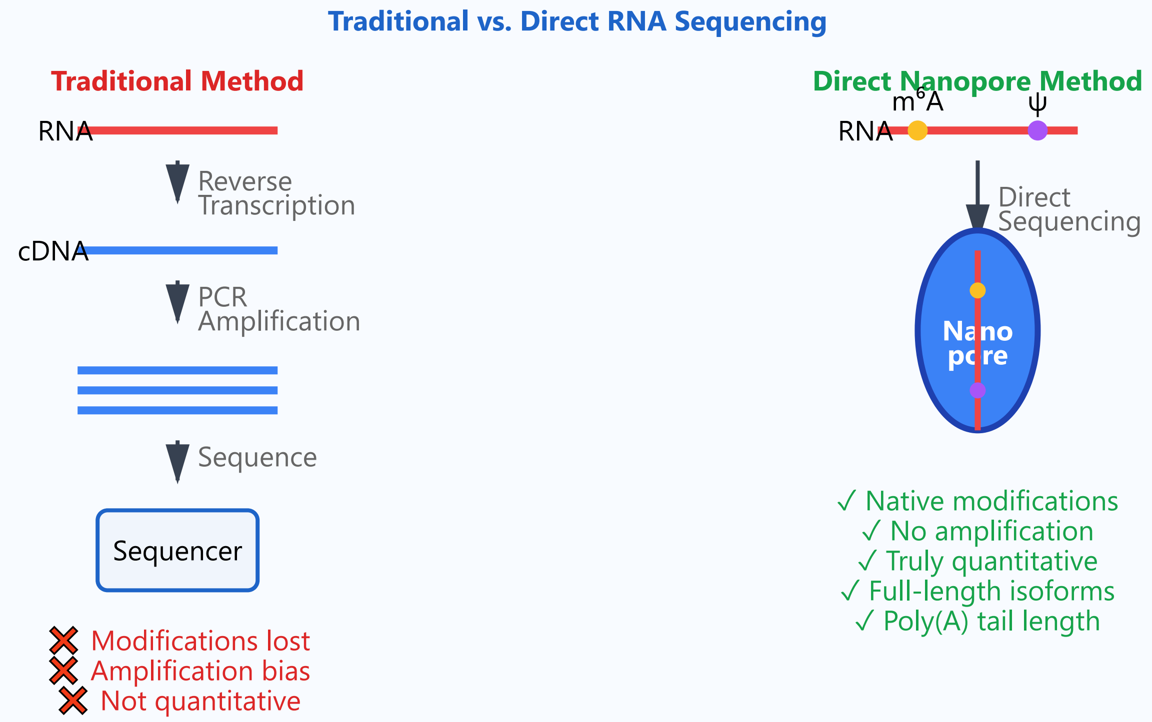
Flongle: Even smaller, single-use adapter with 126 channels for quick tests (~$90)

MinION: USB portable device with 512-2048 channels (varies by flow cell)

GridION: Benchtop device that runs 5 MinION flow cells simultaneously

PromethION: High-throughput platform with up to 48 flow cells, generating terabases of data

5. DIRECT RNA SEQUENCING & NATIVE BASE MODIFICATION DETECTION



Direct RNA Sequencing

Nanopore technology is unique in its ability to sequence RNA molecules directly without conversion to cDNA. This preserves critical information about RNA modifications and structure that is lost in traditional RNA-seq methods.

What Can Be Detected:

Feature	Information Preserved	Applications
Full-length Transcripts	Complete mRNA from 5' cap to poly(A) tail	Isoform identification, alternative splicing analysis
RNA Modifications	m^6A, m^5C, pseudouridine (ψ), inosine	Epitranscriptomics, gene regulation studies
Poly(A) Tail Length	Exact length of poly(A) tail (not just presence)	mRNA stability, translation regulation
Base Modifications	Direct detection without chemical treatment	RNA editing, post-transcriptional regulation
True Quantification	No PCR bias, direct molecule counting	Accurate gene expression levels

DNA Base Modification Detection

Beyond RNA, nanopore sequencing can detect modified bases in DNA, including methylation patterns, without bisulfite conversion or other chemical treatments that can damage DNA.

Detectable DNA Modifications:

- 5-methylcytosine (5mC) - most common DNA methylation
- 5-hydroxymethylcytosine (5hmC)
- N6-methyladenine (6mA)
- 4-methylcytosine (4mC) - bacterial restriction modification

Revolutionary Applications

→ **Cancer Epigenetics:** Detect methylation patterns across entire genomes simultaneously with sequence

→ **Epitranscriptomics:** Study of RNA modifications and their roles in gene regulation

→ **Bacterial Typing:** Identify bacteria by their unique methylation patterns

→ **Full Isoform Characterization:** Understand complete transcript structures including novel splice variants

→ **No Bias:** Eliminates PCR and RT biases that affect quantitative accuracy

→ **Developmental Biology:** Track changes in modifications during cell differentiation

The Future of Modification Detection

As machine learning algorithms improve, nanopore sequencing is becoming increasingly capable of detecting more types of modifications with higher accuracy. This opens new frontiers in understanding how chemical modifications regulate gene expression and cellular function beyond the genetic code itself.

CONCLUSION: THE FUTURE OF SEQUENCING

Transforming Genomics

Nanopore sequencing represents a paradigm shift in how we read genetic information. By eliminating the need for synthesis, amplification, and specialized laboratory infrastructure, it has democratized access to sequencing technology and enabled applications that were previously impossible.

The combination of ultra-long reads, portability, real-time analysis, and native modification detection makes nanopore sequencing uniquely suited for addressing some of the most challenging problems in genomics, from completing reference genomes to rapid pathogen surveillance during disease outbreaks.

As the technology continues to improve in accuracy and throughput, nanopore sequencing is positioned to become a primary tool for both research and clinical applications, bringing us closer to the vision of ubiquitous, real-time genomic information.