***Review Article***

**Next Generation Sequencing in Parasitic Disease Research: Applications and Advances**

### ****Abstract****

The first complete genome of a free-living organism, Mycoplasma genitalium, was sequenced in 1995 using whole-genome shotgun sequencing, which eliminated the need for initial mapping. Next Generation Sequencing (NGS) has transformed Parasitology by enabling rapid, high-throughput analysis of DNA and RNA. NGS has become not only a cornerstone of modern molecular biology but also an indispensable tool in advancing the understanding and control of parasitic diseases. This review outlined the evolution of sequencing technologies and highlights the broad applications of NGS in parasitic disease research. Unlike traditional sequencing methods, NGS offers massive parallel sequencing, allowing for in-depth studies of parasite genomes, gene expression, and drug resistance. With platforms like Illumina, SMRT, and Oxford Nanopore, NGS supports whole-genome and transcriptome sequencing, enhancing diagnostics, vaccine development, and surveillance of parasitic diseases. The insights of the study lay a robust foundation for advanced molecular explorations in parasite biology. It represents a transformative tool in the prevention, diagnosis, and treatment of parasitic infections in both human and veterinary medicine.

**Keywords:** Next Generation Sequencing (NGS), DNA Sequencing Technologies, Genomic Analysis, Transcriptomics, Single Molecule Sequencing

1. **Introduction**

 DNA sequencing is a fundamental technique used to determine the precise order of nucleotides within a DNA molecule, offering critical insights into the genotype, which in turn determines the phenotype. DNA sequencing has undergone three generations of major evolution. Each generation has its own specifications that are mentioned briefly. Among these generations, nanopore sequencing has its own exciting characteristics that have been given more attention here (Mohammadi & Bavi, 2022). Since its introduction in 1977, the most widely adopted sequencing method has been the Sanger sequencing technique, developed by Fred Sanger. This chain-termination method revolutionised genomics and became the cornerstone of molecular biology for decades (Sanger *et al.,* 1977). It was instrumental in the sequencing of the human genome, a monumental achievement completed in 2001 using automated Sanger platforms (International Human Genome Consortium, 2004). Alongside Sanger’s method, an alternative approach was introduced by Maxam and Gilbert in 1977, which was based on the chemical degradation of DNA. However, it was less widely adopted due to its complex protocol and the use of hazardous radioactive materials, although it represented an important milestone in the history of DNA sequencing (Maxam & Gilbert, 1977).

 The journey toward high-throughput sequencing accelerated with the successful sequencing of complete genomes. The first complete genome of a free-living organism, Mycoplasma genitalium, was sequenced in 1995 using whole-genome shotgun sequencing, which eliminated the need for initial mapping (Fraser *et al.,* 1995). This was followed by the sequencing of Haemophilus influenzae (Fleischmann *et al.,* 1995), the first eukaryotic genome of Saccharomyces cerevisiae (Galibert *et al.,* 1996), and the multicellular organism Caenorhabditis elegans (*C. elegans* Sequencing Consortium, 1998). These achievements laid the foundation for the Human Genome Project, initiated in 1990 by the U.S. Department of Energy and the National Institutes of Health. With a budget of $3 billion over 15 years, the project culminated in the publication of the first draft of the human genome in 2001, covering over 90% of the sequence (Venter *et al.,* 2001). The next significant leap occurred in 2005 with the advent of Next Generation Sequencing (NGS), beginning with platforms based on pyrosequencing (Ronaghi *et al.,* 1996; Ronaghi *et al.,* 1998). Unlike traditional methods, NGS introduced massively parallel sequencing, which enabled millions of DNA fragments to be sequenced simultaneously. This innovation led to a dramatic reduction in both the cost and time per base, with sequencing speeds estimated to be 100 times faster than the Sanger method (Ronaghi, 2001). Complete sequencing can be achieved by whole genome sequencing (WGS), while enrichment enables exome, targeted, RNA, and methylation sequencing. NGS chemistry differs between platforms (Singh et al., 2024). For instance, Illumina uses clonal array formation and a proprietary reversible terminator for large-scale sequencing (Hassan et al., 2023).

 NGS operates by fragmenting DNA into smaller segments, attaching sequencing adapters, and reading sequences in parallel. While early NGS platforms had shorter read lengths compared to Sanger sequencing, recent technologies such as Single Molecule Real-Time (SMRT) sequencing and Oxford Nanopore have overcome this limitation, offering longer reads and real-time sequencing capabilities (Eid *et al.,* 2009; Oxford Nanopore Technologies). The use of NGS has additional advantages. It is easier to include a disorder in screening by adding a gene to a genetic panel than by adding a validated biochemical test to a biochemical panel. Moreover, adding genes to a virtual panel (i.e., one created by digitally filtering whole exome data for a set of regions of interest) for the analysis of WES/whole-genome sequencing (WGS) data is easier than adding them to a targeted panel (Veldman et al., 2022). Today, platforms like Illumina, SMRT, and Oxford Nanopore have revolutionised genetic research. Their applications in Parasitology have enabled comprehensive studies in whole-genome sequencing, transcriptomics, drug resistance profiling, vaccine candidate identification, and disease surveillance. NGS has become not only a cornerstone of modern molecular biology but also an indispensable tool in advancing the understanding and control of parasitic diseases.

**2. Evolution of Sequencing**

**2.1 First-Generation (Traditional) Sequencing**

 The first human reference genome was completed in 13 years using the chain termination method, commonly known as the Sanger sequencing method (Sanger et al., 1977). In the approach used for sequencing the entire human genome, more than 20,000 artificial bacterial clones were created, each carrying a chromosomal fragment of approximately 100 kb in length.

**2.1.1** **Sanger Sequencing**:

 This sequencing technique operates through the addition of 2’-deoxynucleotides (dNTPs) and 2’,3’-dideoxynucleotides (ddNTPs) during the DNA synthesis reaction, mediated by DNA polymerase. Whenever a ddNTP, which lacks a 3’-OH group, is incorporated, the chain elongation is terminated. The resulting DNA fragments of various lengths are separated via capillary electrophoresis. DNA sequences of up to 700–800 base pairs can typically be read with high accuracy.

 One of the major advantages of Sanger sequencing lies in its low error rate, which makes it an important tool for validating genetic variants identified by Next Generation Sequencing (NGS). The method also involves the use of radio-labelled fragments, which are partially digested and then separated through two-dimensional fractionation (Sanger *et al.,* 1965). A key benefit of this method was its suitability for automation, which significantly enhanced its applicability in large-scale sequencing projects.

**2.1.2** **Maxam & Gilbert Method**

###  An alternative approach developed during the same period was the Maxam and Gilbert method, which relied on chemical degradation of DNA strands (Maxam and Gilbert, 1977). However, this technique did not gain widespread use due to several limitations, including the need for hazardous radioactive reagents, procedural complexity, and difficulty in interpreting the sequencing results

### ****2.2 Second-Generation (Next Generation) Sequencing****

 Second-generation sequencing is based on the sequencing of amplified DNA molecules. The main advantages of Next Generation Sequencing (NGS) include its high throughput, the generation of vast amounts of data, high precision, broad applicability, and high-quality reads (Reis-Filho *et al.,* 2009). This method incorporates polymerase chain reaction (PCR), the creation of a genomic library, immobilisation of DNA fragments on a solid substrate, and fluorescence-based detection of incorporated nucleotides.

 Unlike Sanger sequencing, which requires bacterial cloning, the DNA library in NGS is prepared through nebulization or sonication, which fragments the DNA using sound waves or gas pressure into pieces of similar length. Short adapter sequences are then ligated to the ends of the fragments to facilitate replication and sequencing. The prepared library is then amplified and sequenced, depending on the specific platform. The fundamental principle of NGS remains similar to that of Sanger sequencing; it relies on the detection of signals emitted by nucleotide incorporation during DNA synthesis. A critical step in this workflow is the proper preparation of the DNA library (Mardis, 2008). NGS consists of three main stages: Isolation and preparation of the DNA library, Amplification of the DNA templates and Massive parallel sequencing.

Second-generation sequencing platforms enable the parallel sequencing of millions of amplified DNA fragments. The DNA is fragmented, adapters are ligated, and PCR amplification is carried out. Major platforms include:

* + 1. **Roche/454 Pyrosequencing**:

 The first commercially available NGS platform, developed in 2005, was the Roche/454 GS FLX Titanium system (Branford, USA), which uses the pyrosequencing method originally introduced by Ronaghi *et al.* (1998). Pyrosequencing is based on the detection of pyrophosphate released during nucleotide incorporation. DNA fragments are attached to beads coated with oligonucleotides complementary to adapter sequences. Amplification occurs via emulsion PCR inside micelles that contain DNA polymerase, primers, and nucleotides (Tawfik and Griffiths, 1998).

 Beads carrying multiple copies of individual DNA fragments are then placed into a plate with millions of wells, each containing enzymes needed for chemiluminescent detection. Sequencing is initiated by adding DNA polymerase and nucleotides. When a nucleotide is incorporated into a growing DNA strand, pyrophosphate is released and converted to ATP by ATP sulfurylase. Luciferase then uses this ATP to emit light, the intensity of which is proportional to the number of incorporated nucleotides, thus enabling base calling (Margulies *et al.,* 2005).

**2.2.2 Illumina/Solexa Genome Analyser:**

 In the platform called Genome Analyser (Illumina/Solexa; San Diego, USA), issued in 2006, a bridge PCR method is used during the amplification phase (Bentley *et al*., 2006 and Bentley *et al*., 2008). DNA fragments with adapters on both ends hybridise to a glass surface coated with complementary primers. During amplification, a looped structure forms, and the free end hybridises to a neighbouring primer, creating a bridge.

 Following each amplification cycle, DNA polymerase synthesises the new strand, and denaturation occurs. Sequencing is then carried out via the cyclic reversible termination method. In this process, DNA polymerase, primers, and fluorescently labelled terminator nucleotides are added. Each time a nucleotide is incorporated, synthesis is temporarily halted, and fluorescence is recorded. After the chemical group on the nucleotide is removed, the cycle continues. Although the Genome Analyser offers high throughput, its read length (typically 2×100 bp) is shorter than that of 454 technology.

* + 1. **Applied Biosystems or Solid and Ion Torrent:**

 The SOLiD (Supported Oligonucleotide Ligation and Detection) system was released by Life Technologies in 2006 (Carlsbad, USA). In this method, DNA fragments are amplified by emulsion PCR and sequenced via ligation rather than synthesis. Short oligonucleotide probes, known as interrogation probes, consist of two defined bases and six degenerate bases. The 5' end of each probe is labelled with a fluorescent dye (Shendure *et al.,* 2005).

 The sequencing reaction uses 16 possible probe combinations, with fluorescence recorded upon successful hybridisation. After cleavage and removal of the fluorescent label, the next probe is ligated. The process is repeated, starting one base downstream, allowing each base to be interrogated twice, resulting in a low error rate. Ion Torrent, another platform from Life Technologies, uses a different principle. It does not rely on fluorescence but instead detects hydrogen ions released during nucleotide incorporation, which causes a change in pH. This change is detected by an ion sensor and recorded (Rothberg *et al.,* 2011).

* 1. **Third-Generation (Next-Next Generation) Sequencing**

 Although various next-generation sequencing platforms have been developed based on different biochemical principles, the core methodologies share certain similarities (Mardis *et al.,* 2008). A common feature among these technologies is the requirement to prepare a library of single-stranded DNA. While sequencing steps vary between platforms, the overall process typically involves three key stages: (1) isolation and creation of the DNA library, (2) template amplification, and (3) massively parallel sequencing. In contrast to previous generations, **third-generation sequencing** enables **single-molecule sequencing** without the need for pre-amplification. This advance allows for more direct, real-time sequencing with reduced biases and shorter processing times.

**2.3.1 Single-Molecule DNA Template Sequencing Technologies**

 Traditional NGS platforms rely on amplified signals for detection, thus requiring DNA templates to undergo amplification. However, third-generation technologies have introduced sequencing platforms that **do not require amplification**, allowing for direct sequencing of **single DNA molecules** with enhanced sensitivity to low-level fluorescent signals. One of the earliest platforms in this category was the **HeliScope** by **Helicos Biosciences** (Cambridge, USA) (Braslavsky *et al.,* 2003). In this approach, the 3’ ends of genomic DNA fragments are labelled with fluorescently tagged adenosines and are attached to **poly-A oligonucleotides**. These fragments hybridise to **poly-T primers** that are immobilised on a flow cell surface. DNA polymerase then synthesises the complementary strand using fluorescently labelled dNTPs, with the signal from each incorporated nucleotide being recorded before the label is cleaved and removed. Another major advancement is the **Single Molecule Real-Time (SMRT) sequencing** developed by **Pacific Biosciences** (Menlo Park, USA). In this system, **DNA polymerase enzymes** are immobilised within nanostructured wells (Zero-Mode Waveguides), and real-time fluorescent signals are detected as individual nucleotides are incorporated into the growing DNA strand. This platform is capable of capturing **millisecond-resolution signals**, enabling fast, real-time sequencing of long reads (Eid *et al.,* 2009).

**2.3.2** **Nanopore Sequencing (Oxford Nanopore Technologies)**:

 Another innovative real-time sequencing platform is the **GridION** device, developed by **Oxford Nanopore Technologies** (Oxford, UK). This method does not rely on fluorescence. Instead, **single-stranded DNA molecules** are passed through **nanopores,** tiny holes with nanometre-scale diameters across which an **electric current** is applied. As each nucleotide passes through the pore, it causes a characteristic change in the current, allowing for **real-time detection** of the sequence based on electrical signal shifts. The **main advantage** of nanopore and other single-molecule sequencing technologies is the **elimination of the amplification phase**, significantly reducing the amount of starting DNA required, often less than **1 µg.** However, one limitation is that **the sequencing cannot be repeated** for the same molecule (unlike platforms that use clonal amplification), leading to a potentially **higher error rate** (Pop and Salzberg, 2008).

1. **Next-Generation Sequencing Data Analysis**

 The massive volume of data generated through Next-Generation Sequencing (NGS) requires robust and systematic analysis pipelines (Su *et al.,* 2011). NGS data analysis typically involves several key steps (Trapnell *et al*., 2009). The first step is **base calling,** where recorded fluorescent or light signal images from the sequencing experiment are converted into nucleotide sequences. Most NGS platforms employ proprietary algorithms that assign quality scores to each read to assess confidence levels. The second step involves **alignment** of short sequence reads to a reference genome, if available. In cases where no reference genome exists, **de novo assembly** is used. Genetic variations, including single-nucleotide polymorphisms (SNPs), insertions, and deletions, are identified by comparing reads with the reference genome, such as GRCh37. An essential factor here is **coverage depth;** the greater the overlap of reads in a genomic region, the more accurate the sequence data. However, short-read lengths present a challenge in regions with high repeat content, pseudogenes, or homologous genes, leading to difficulties in correct assembly (Voelkerding *et al.,* 2009). Another limitation is the detection of large structural variants, such as **copy number variations (CNVs)** spanning more than 1 kb, though newer computational tools have improved CNV detection (Duan *et al.,* 2013).

1. **Applications of Next Generation Sequencing**

 Different applications are possible on a single NGS, which spans across the central dogma of molecular biology (DNA makes RNA makes protein)

* 1. **Whole Genome Sequencing (WGS)-** Provides comprehensive insight into the entire genomic content, facilitating large-scale biological analyses (Nusbaum *et al.,* 2009).
	2. **Whole Exome Sequencing**- Targets protein-coding regions to explore the functional impact of genetic variants (Laurent *et al.,* 2014).
	3. **Targeted Sequence**- Focuses on specific genomic regions of interest, enhancing efficiency in variant detection for diagnostic purposes (Altmuller *et al.,* 2014).
	4. **Metagenomic-** Enables the study of microbial communities and parasitic populations in ecological niches without culturing
	5. **Amplicon Sequencing**- Deep sequencing of PCR-amplified regions allows precise detection of genetic mutations or variants in genes of interest.
	6. **Transcriptome Analysis**- Examines gene expression and splicing events to understand cellular function and regulation (Wang et al., 2009).
	7. **Chromatin Immunoprecipitation Analysis (Chip)**- Investigates interactions between DNA and proteins such as transcription factors, revealing mechanisms of gene regulation (Park *et al.,* 2009).
1. **Application in Parasitic Disease Research**

 Next Generation Sequencing (NGS) has significantly enhanced our understanding of parasitic diseases by enabling detailed genomic and transcriptomic analyses. Almeida *et al.* (2007) utilised massively parallel sequencing and bioinformatics to analyse the transcriptome of Schistosoma mansoni, revealing molecular mechanisms of pathogenesis and identifying novel vaccine candidates. Similarly, Oliveira *et al.* (2007) sequenced Schistosoma japonicum for gene discovery and genome annotation, aiding the identification of potential vaccine targets. Ojopi *et al.* (2007) further contributed by revealing gene expression frequencies in the S. mansoni transcriptome. Liu *et al*. (2008), using Illumina sequencing, uncovered differentially expressed genes in S. japonicum, providing fundamental insights into schistosome biology, evolution, and host-parasite interactions. Young *et al.* (2010) applied high-throughput sequencing to study the transcriptomes of the carcinogenic liver flukes Clonorchis sinensis and Opisthorchis viverrini, successfully predicting gene content and expression profiles. In a comparative study, Young *et al.* (2011) used Illumina sequencing to analyse the transcriptome of Fasciola gigantica alongside other trematodes such as F. hepatica. Yoo *et al.* (2013) focused again on C. sinensis, predicting differentially expressed functional genes across developmental stages. Cinzia *et al.* (2015) examined the transcriptome and secreted proteome of adult Fascioloides magna using Illumina platforms, contributing to proteomic profiling of this giant liver fluke. Protasio *et al. (*2016) conducted a comprehensive transcriptomic study of S. mansoni using Illumina sequencing to predict genes and profile their expression. Xuhang *et al.* (2017) studied the transcriptome of Taenia multiceps, generating a foundational dataset for understanding the parasite’s biology. Finally, Biswal *et al.* (2018) sequenced the mitochondrial genome of Fasciolopsis buski using both Ion Torrent and Illumina technologies, developing an integrated bioinformatics pipeline as a valuable resource for comparative mitochondrial genomics.

1. **Conclusion**

 The advent of Next-Generation Sequencing platforms, coupled with bioinformatics and subsystem-based annotations, has revolutionised parasitic disease research. Massively parallel sequencing technologies now enable detailed genomic, transcriptomic, and proteomic investigations, enhancing our understanding of gene homology, ontologies, and metabolic pathways. These insights lay a robust foundation for advanced molecular explorations in parasite biology. NGS has proven instrumental in identifying drug targets, vaccine candidates, and key mechanisms underlying host-parasite interactions. With its ability to process large-scale data rapidly and accurately, NGS represents a transformative tool in the prevention, diagnosis, and treatment of parasitic infections in both human and veterinary medicine.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

**Reference**

1. Almeida, G. T., et al. "Expression profiling of Schistosoma mansoni in response to praziquantel treatment using oligonucleotide microarrays." Molecular and Biochemical Parasitology, vol. 154, no. 1, 2007, pp. 119–129.
2. Bentley, David R. "Whole-genome re-sequencing." Current Opinion in Genetics & Development, vol. 16, no. 6, 2006, pp. 545–552.
3. Bentley, David R., et al. "Accurate whole human genome sequencing using reversible terminator chemistry." Nature, vol. 456, no. 7218, 2008, pp. 53–59.
4. Biswal, D. K., et al. "Sequencing and analysis of the complete mitochondrial genome of Fasciolopsis buski." Acta Tropica, vol. 180, 2018, pp. 106–112.
5. Braslavsky, I., et al. "Sequence information can be obtained from single DNA molecules." Proceedings of the National Academy of Sciences USA, vol. 100, no. 7, 2003, pp. 3960–3964.
6. Cinzia, Cantacessi, et al. "Transcriptome and proteome analysis of the liver fluke Fascioloides magna." Molecular and Biochemical Parasitology, vol. 201, no. 2, 2015, pp. 66–73.
7. Duan, J., et al. "Comparative studies of copy number variation detection methods for next-generation sequencing technologies." PLOS ONE, vol. 8, no. 3, 2013, e59128.
8. Eid, J., et al. "Real-time DNA sequencing from single polymerase molecules." Science, vol. 323, no. 5910, 2009, pp. 133–138.
9. Fleischmann, R. D., et al. "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd." Science, vol. 269, no. 5223, 1995, pp. 496–512.
10. Fraser, C. M., et al. "The minimal gene complement of Mycoplasma genitalium." Science, vol. 270, no. 5235, 1995, pp. 397–403.
11. Galibert, F., et al. "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome X." Nature, vol. 369, 1996, pp. 586–588.
12. International Human Genome Consortium. "Finishing the euchromatic sequence of the human genome." Nature, vol. 431, no. 7011, 2004, pp. 931–945.
13. Laurent, L. C., et al. "Whole exome sequencing: Current applications and future perspectives." Genomics Proteomics Bioinformatics, vol. 12, no. 5, 2014, pp. 223–235.
14. Liu, F., et al. "New perspectives on host–parasite interactions from the Schistosoma japonicum genome." PLOS Pathogens, vol. 4, no. 1, 2008, e1000002.
15. Mardis, Elaine R. "Next-generation DNA sequencing methods." Annual Review of Genomics and Human Genetics, vol. 9, 2008, pp. 387–402.
16. Margulies, M., et al. "Genome sequencing in microfabricated high-density picolitre reactors." Nature, vol. 437, no. 7057, 2005, pp. 376–380.
17. Massere, P., et al. "Metagenomic analysis of gut microbial communities reveals microbial markers for obesity." PLOS ONE, vol. 10, no. 7, 2015, e0126920.
18. Maxam, A. M., and W. Gilbert. "A new method for sequencing DNA." Proceedings of the National Academy of Sciences of the United States of America, vol. 74, no. 2, 1977, pp. 560–564.
19. Nusbaum, C., et al. "Sensitive, specific polymorphism discovery in bacteria using massively parallel sequencing." Nature Methods, vol. 6, no. 1, 2009, pp. 67–69.
20. Ojopi, E. P., et al. "Transcriptome analysis of the human pathogen Schistosoma mansoni." Genome Biology, vol. 8, no. 4, 2007, R57.
21. Oliveira, G., et al. "Gene expression analysis of the Schistosoma mansoni transcriptome reveals novel genes with potential functions in host–parasite interaction." BMC Genomics, vol. 9, 2007, p. 88.
22. Park, P. J. "ChIP-seq: Advantages and challenges of a maturing technology." Nature Reviews Genetics, vol. 10, no. 10, 2009, pp. 669–680.
23. Pop, Mihai, and Steven L. Salzberg. "Bioinformatics challenges of new sequencing technology." Trends in Genetics, vol. 24, no. 3, 2008, pp. 142–149.
24. Protasio, A. V., et al. "A systematically improved high-quality genome and transcriptome of the human blood fluke Schistosoma mansoni." PLOS Neglected Tropical Diseases, vol. 6, no. 1, 2012, e1455.
25. Reis-Filho, Jorge S., et al. "Molecular profiling: moving away from histology?" Pathology, vol. 41, no. 1, 2009, pp. 20–27. (year not match)
26. Ronaghi, Mostafa, et al. "Real-time DNA sequencing using detection of pyrophosphate release." Analytical Biochemistry, vol. 242, no. 1, 1996, pp. 84–89.
27. Ronaghi, Mostafa, Ulf Landegren, and Pål Nyrén. "A sequencing method based on real- time pyrophosphate." Science, vol. 281, no. 5375, 1998, pp. 363–365.
28. Ronaghi, Mostafa. "Pyrosequencing sheds light on DNA sequencing." Genome Research, vol. 11, no. 1, 2001, pp. 3–11.
29. Rothberg, Jonathan M., et al. "An integrated semiconductor device enabling non-optical genome sequencing." Nature, vol. 475, no. 7356, 2011, pp. 348–352.
30. Sanger, Frederick, et al. "DNA sequencing with chain-terminating inhibitors." Proceedings of the National Academy of Sciences of the United States of America, vol. 74, no. 12, 1977, pp. 5463–5467.
31. Sanger, Frederick, et al. "Nucleotide sequence of bacteriophage φX174 DNA." Nature, vol. 251, 1965, pp. 546–551. (Note: Year possibly needs confirmation)
32. Shendure, Jay, et al. "Accurate multiplex polony sequencing of an evolved bacterial genome." Science, vol. 309, no. 5741, 2005, pp. 1728–1732.
33. Su, Z., et al. "Next-generation sequencing and its applications in molecular diagnostics." Expert Review of Molecular Diagnostics, vol. 11, no. 3, 2011, pp. 333–343.
34. Tawfik, Dan S., and Andrew D. Griffiths. "Man-made cell-like compartments for molecular evolution." Nature Biotechnology, vol. 16, no. 7, 1998, pp. 652–656.
35. The C. elegans Sequencing Consortium. "Genome sequence of the nematode C. elegans: a platform for investigating biology." Science, vol. 282, no. 5396, 1998, pp. 2012– 2018.
36. Trapnell, C., and Salzberg, S. L. "How to map billions of short reads onto genomes." Nature Biotechnology, vol. 27, no. 5, 2009, pp. 455–457.
37. Venter, J. Craig, et al. "The sequence of the human genome." Science, vol. 291, no. 5507, 2001, pp. 1304–1351.
38. Voelkerding, K. V., et al. "Next-generation sequencing: From basic research to diagnostics." Clinical Chemistry, vol. 55, no. 4, 2009, pp. 641–658.
39. Wang, Z., Gerstein, M., and Snyder, M. "RNA-Seq: A revolutionary tool for transcriptomics." Nature Reviews Genetics, vol. 10, no. 1, 2009, pp. 57–63.
40. Xuhang, L., et al. "Transcriptomic analysis of the larval stage of the tapeworm Taenia multiceps." Parasitology Research, vol. 116, no. 6, 2017, pp. 1579–1589. (AUTHOR NOT MATCH AND TITLE PROPERLY NOT MATCH)
41. Yoo, W. G., et al. "Genome-wide analysis of differential gene expression in Clonorchis sinensis according to developmental stage." Parasitology Research, vol. 112, no. 3, 2013, pp. 1139–1150.
42. Young, N. D., et al. "Comparative genomics of the major parasitic worms." Nature Genetics, vol. 44, no. 1, 2011, pp. 83–89.
43. Young, N. D., et al. "Whole-genome sequence of the carcinogenic human liver fluke Clonorchis sinensis." Genome Biology, vol. 11, no. 10, 2010, R115.
44. Mohammadi, M. M., & Bavi, O. (2022). DNA sequencing: an overview of solid-state and biological nanopore-based methods. *Biophysical reviews*, *14*(1), 99-110.
45. Hassan, S., Bahar, R., Johan, M. F., Mohamed Hashim, E. K., Abdullah, W. Z., Esa, E., Abdul Hamid, F. S., & Zulkafli, Z. (2023). Next-Generation Sequencing (NGS) and Third-Generation Sequencing (TGS) for the Diagnosis of Thalassemia. *Diagnostics*, *13*(3), 373.
46. Veldman, A., Kiewiet, M. B. G., Heiner-Fokkema, M. R., Nelen, M. R., Sinke, R. J., Sikkema-Raddatz, B., Voorhoeve, E., Westra, D., Dollé, M. E. T., Schielen, P. C. J. I., & van Spronsen, F. J. (2022). Towards Next-Generation Sequencing (NGS)-Based Newborn Screening: A Technical Study to Prepare for the Challenges Ahead. *International Journal of Neonatal Screening*, *8*(1), 17.
47. Singh, L. S., Iqbal, G., Singh, A., Limbola, M., & Kumar, A. V. (2024). The Impact of Next-generation Sequencing on Biotechnology: A Review of Current Applications. *Journal of Scientific Research and Reports*, *30*(11), 91–97.