

A Review of Next Generation Sequencing Methods and its Applications in Laboratory Diagnosis

Srivalsa Bhaskaran*  and Chithralekha Saikumar 

Department of Microbiology, Sree Balaji Medical College and Hospital, Chromepet, Chennai - 600 044, Tamil Nadu, India.

Abstract

Next-generation sequencing (NGS) is a new technology used to detect the sequence of DNA and RNA and to detect mutations or variations of significance. NGS generates large quantities of sequence data within a short time duration. The various types of sequencing includes Sanger Sequencing, Pyrosequencing, Sequencing by Synthesis (Illumina), Ligation (SoLiD), Single molecule Fluorescent Sequencing (Helicos), Single molecule Real time Sequencing (Pacbio), Semiconductor sequencing (Ion torrent technology), Nanopore sequencing and fourth generation sequencing. These methods of sequencing have been modified and improved over the years such that it has become cost effective and accessible to diagnostic laboratories. Management of Outbreaks, rapid identification of bacteria, molecular case finding, taxonomy, detection of the zoonotic agents and guiding prevention strategies in HIV outbreaks are just a few of the many applications of Next Generation sequencing in clinical microbiology.

Keywords: Next generation sequencing, Sangers, Illumina, SoLiD, Outbreak management

*Correspondence: srivalsa@hotmail.com; +91 9840910646

(Received: November 6, 2021; accepted: May 14, 2022)

Citation: Bhaskaran S, Saikumar C. A Review of Next Generation Sequencing Methods and its Applications in Laboratory Diagnosis. *J Pure Appl Microbiol.* 2022;16(2):825-833. doi: 10.22207/JPAM.16.2.45

© The Author(s) 2022. **Open Access.** This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, sharing, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

INTRODUCTION

Next-generation sequencing (NGS) is a new technology used to detect the sequence of DNA and RNA and mutations or variations of significance. NGS generates large quantities of sequence data within a shorter time duration as compared to conventional Sanger's method of sequencing. This novel technique uses different chemistries, matrices and bioinformatics technologies which can be used to sequence entire genomes or different lengths of DNA and RNA sequences in shorter time periods.⁴ The first Next Generation Sequencing technique to be commercially available was the massively parallel pyrosequencing platform in 2005.⁴⁶

DNA sequencing is done various steps which includes :

1. DNA fragmentation
2. Gene Library
3. Sequencing
4. Data analysis

DNA Fragmentation

Targeted DNA is broken into several small segments using different methods like sonication, enzymatic digestion. The required short segments are isolated using different methods such as Hybridisation Capture Assay, Amplicon Assay

Genomic Library

An organism's genome gets broken down into smaller pieces. Each piece is cloned into a unique vector, which is then carried by a unique microbial cell. A collection of recombinant vectors represents the organism's whole genome. This helps to analyse isolated groups of genes, hence understand their expression and function. It is this library that is sequenced using the various DNA sequencing methods.

All the methods of DNA sequencing starts by attaching short oligonucleotide sequences to the ends of fragmented short strands of DNA. These short oligonucleotide sequences are called adapters. The 5' and 3' ends of fragmented or amplicon DNA are annealed with certain DNA adaptor sequences. Double-stranded DNA adapters are 20–40bp segments with known sequences. On both the 5' and 3' ends of the fragmented DNA, there are two distinct adaptor sequences that can anneal to the DNA fragments.

Each DNA fragment has an adapter on one end that connects it to a solid substrate such

as beads or flow cells, and another adapter on the other end that anneals to a primer that starts the polymerase chain reaction (PCR). PCR produces several copies of the same fragment, which are sequenced at the same time. As a result, these techniques are sometimes referred to as massively parallel sequencing techniques.

DNA Sequencing

An NGS sequencer is used to perform massive parallel sequencing. In a specific sequencer, the library is uploaded onto a sequencing matrix. The platform on which the sequencing takes place is known as a sequencing matrix. Sequencing matrices differ depending on the sequencer. For example, the Illumina NGS sequencer uses flow cells, while the Ion Torrent NGS sequencer uses sequencing chips.

Several generations of sequencing methods have been developed.

1. First Generation sequencing- Sanger Sequencing
2. Second Generation Sequencing
 - a. Pyrosequencing
 - b. Sequencing by Synthesis (Illumina)
 - c. Ligation (SoLID)
3. Third Generation Sequencing
 - a. Single molecule Fluorescent Sequencing (Helicos)
 - b. Single molecule Real time Sequencing (Pacbio)
 - c. Semiconductor sequencing (Ion torrent technology)
 - d. Nanopore sequencing
4. Fourth generation sequencing

Sanger Sequencing

The first method used for sequencing was Sanger's DNA sequencing. Dr. Frederick Sanger's did research to design a DNA sequencing method in the 1970s. He had already published methods for RNA sequencing in the late 1960s. Sanger sequencing was the most commonly used tool in genomic research that led to extraordinary accomplishments. For example, the high-quality, reference sequence of the human genome under the Human Genome Project (HGP) was a product of Sangers sequencing.

The synthesis of DNA occurs in the 5' to 3' direction. The DNA double helix unwinds to expose single strands of the DNA which are replicated by adding on complementary nucleotides (dNTPs). Several enzymes are involved in the

replication of DNA including DNA polymerase, helicase, topoisomerase and DNA gyrase. The replication of DNA usually stops when the replication fork ends or upon reaching termination sequences in the template DNA strand.

Here the DNA to be sequenced is used as a template. The reaction mixture consists of primer (a short sequence of DNA complementary to the region to be sequenced), DNA polymerase, the four deoxynucleoside triphosphates (dNTPs), and dideoxynucleoside triphosphates (ddNTPs). The single stranded template DNA is mixed with this reaction mixture. The synthesis of DNA is carried out as usual with dNTPs till a ddNTP is added to the complementary strand. The ddNTP lacks a hydroxyl group at the 3' end, hence no interaction occurs with the 5'PO₄ of the next dNTP. The elongation of the complementary chain of DNA stops with the ddNTP. Hence, Sanger's sequencing is also called the Chain Termination DNA sequencing method.

Several such strands of DNA are obtained and the various locations of the particular nucleotide on the complementary strand is determined by performing electrophoresis on thin slab polyacrylamide gel. Individual synthesis reactions are prepared for each ddNTP. In the initial models, the position of the different fragments were identified using P³² which was used to label the dATP molecules. This was observed on an X ray film.⁷

Later use of fluorescent labels simplified the process of reading the electrophoresis plates. In 1986, Applied Biosystems introduced a fluorescent DNA sequencing instrument, which used fluorescently labelled primers. The primer for specific nucleotide reaction was labelled with a specific fluorochrome and a scanning laser beam emitting specific wavelengths scanned the surface of the gel. Different excitation wavelengths were emitted from the fluorescently labelled primers which were detected during the separation of fragments electrophoretically. Initially, the fluorochromes were used to label the primer sequences for identification. The further sequences and position of nucleotides were identified by the length of the electrophoretically separated fragments. Improvements were made in the technology and fluorescently labelled dideoxynucleotides started to be used known as terminators. This allowed all the four

dideoxynucleotides to get labelled with different fluorochromes and the same reaction vessel could be used. This helped in lowering the cost of the run.

Automated methods for sequencing became popular with the introduction of fluorescent labelling. ddNTPs are fluorescently labelled and all the four reaction mixtures are loaded into a single lane of a gel. These fragments are separated using electrophoresis. The position of each ddNTP is detected by laser beams when the ddNTPs leave the gel. A chromatogram documenting the fragment order is generated in which the amplitude of each spike represents the fluorescent intensity of each fragment.

Further improvement in the technology was made by introduction of Capillary gel electrophoresis instead of slab gel electrophoresis. This involved directly injecting a polymeric separation matrix into capillaries leading to single-nucleotide resolution.⁷

The sample containing the fragmented and amplified DNA is loaded into the capillaries. Electrical current pulses were passed through the loaded capillary gel thereby separating the DNA fragments to obtain single nucleotide resolution through a process known as electrokinetic injection.

Sanger's Chain termination method is very time consuming and expensive method of gene sequencing. The human genome project which took 10 years to complete used Sanger's Chain termination DNA sequencing method. Newer methods of DNA sequencing which does not require construction of Genomic Library have been developed.

Pyrosequencing / 454 Sequencing

A succession of enzymatic reactions are used which leads to generation of visible light. Each plastic bead has one DNA fragment attached which is amplified by PCR within an oil-water emulsion. The final product has is a about one million copies of the DNA fragment covering the bead. The PCR products are denatured and each bead is deposited into picoliter sized wells. Three important enzyme reactions take place in these wells. In the first step, DNA synthesis occurs using a primer and single type of unlabelled dNTP such as d ATP, dGTP, dCTP, or dTTP which is catalysed by DNA polymerase. With each nucleotide added,

a pyrophosphate (PPi) molecule is released. In the second step, the enzyme ATP sulfurylase catalyses the conversion of the PPi to ATP.

In the third step, this ATP is converted to light using the enzyme firefly luciferase. A flash of light is hence generated for each nucleotide added and the intensity of the flash depends on how many nucleotides were added. The flashes generated from each well is correlated with the nucleotide added in the particular well. A Computer software is used to monitor the growth of the DNA chain synthesised, one nucleotide at a time.

ABISOLID

This method is called sequencing by ligation or SOLiD technology. Amplification of genomic fragments is similar to pyrosequencing. The amplified fragments are moved onto a glass support surface where a primer is hybridised to the adapter. Following denaturation of amplified fragments, a primer with complementary nucleotide sequence to the adapter is added. Eight based oligonucleotides (octamers) are attached through hybridization to each fragment and this is yet again attached to the primer by the enzyme ligase. The bases in the fourth or fifth position of the octamer are labelled with fluorescent markers. To identify the bases, laser light is used to stimulate the fluorescent labels. After that, the fluorescent label is removed by cleaving the ligated octamer after the fifth base. The ligation and cleavage process is repeated, and the sequence is identified based on the detected fluorescence.

Whole genome resequencing, targeted resequencing, transcriptome research (including gene expression profiling, small RNA analysis, and whole transcriptome analysis), and epigenome research are among applications of SOLiD.

Solexa/Illumina Sequencing

Solexa/Illumina sequencing is sequencing by synthesis and reversible dye terminators which help to identify each nucleotide added onto the DNA strand during the sequencing process. This type of sequencing uses a glass slide coated with a lawn oligonucleotides in 8 channels known as the flow cell. The sequences of these oligonucleotides complement the adapters on both the 3' and 5' ends of the DNA strand, respectively. Hence the fragments of DNA with the adapters at both ends attach to the glass slide. This forms the

solid surface where the sequencing occurs. The adsorbed DNA strand bind on the 3' and 5' ends to complementary oligonucleotides in the flow cell. Using the Template DNA strand linked to the flow cell via adapters, a polymerase enzyme synthesises the complementary DNA strand. The denatured double-stranded molecule is formed and the original template is washed away. Bridge PCR is used for clonal amplification. Bridge PCR is a technique in which the DNA strand linked to the flow cell on the 3' end folds over, allowing the adaptor on the 5' end to hybridise with the flow cell's matching oligonucleotide. A DNA polymerase then generates complementary strands, resulting in the formation of a double-stranded bridge. This double stranded bridge is denatured resulting in 2 single stranded copies of the molecule attached to the flow cell. Ultimately, many copies of a single DNA template are present as clusters in bundles at their respective locations. Once cluster generation is complete, the templates are ready for sequencing. Cluster generation is essential for achieving the signal intensity required for sequencing.

Pacific Biosciences Single Molecule Real Time Reads

This technology applies the Real time principle to sequencing. A SMRT bell library is created by ligating adapters to both ends of the denatured dsDNA creating a circular template. The platform for sequencing here is a SMRT cell which contains millions of tiny wells called Zero mode Waveguides (ZMW). A single circular DNA molecule is incorporated in a single ZMW. Primers along with DNA polymerases are added which adds labelled nucleotides. With every nucleotide added light is discharged. The nucleotide addition is hence, observed in Real time.

Nanopore DNA Sequencing

Nanopore sequencing devices uses flow cells as their platform for amplification. These flow cells contain numerous minute holes called as Nanopores which are embedded in an electro-resistant membrane. minute holes are lined by electrically conducting molecules such as iron molecule. As a result, the Nanopore has an electric field and acts on its own as an electrode, which is linked to a channel and sensor chip that gauges the electric current flowing through it.⁸ The DNA fragment to be sequenced is passed through the

nanopore such that at a time only one nucleotide can pass through. The different nucleotides produce a change in the magnitude of current which is detected in real time by the sensor chip. The alteration of the electric field of the nanopore is different for each base which is used for detecting the nucleotide base. The characteristic squiggle in the electric field produced by the passage of each nucleic acid through the nanopore allows the DNA to be sequenced. This method can be used for sequencing of DNA, RNA and proteins. The currently in use nanopores are single channel nanopores like biological nanopores and solid state nanopores. Recent research has shown promise for multichannel nanopore arrays where the fluidic chip has multiple pores set in a parallel sequence facing one common chamber where the DNA is added. This in turn helps to achieve better and faster results.²⁹

Helicos tSMS

It is the first commercial NGS platform using the principle of single molecule fluorescent sequencing. DNA samples are first fragmented to small size molecules of 100-200 nucleotides in length. A universal poly A sequence along with a fluorescent adenosine nucleotide is attached to the 3' end of each strand which acts as the template. A Helicos flow cell is the surface where hybridization occurs. The flow cell has several oligoT oligomers immobilised on its surface which act as capture sites for the polyA tagged DNA fragments. The flow cell that contains the hybridised fragments is loaded into the instrument. Each fluorescently labelled template is detected by a laser that illuminates the flow cell. On the flow cell surface, a CCD camera creates a map of the template. The fluorescent template label is cleaved and cleared away once the templates have been recorded. A DNA polymerase is added which catalyses the addition of single type of complementary fluorescently labelled nucleotides one at a time to the primers. A wash step removes the unhybridized nucleotides and DNA polymerase. A laser beam is used again to detect and image the fluorescence from the added nucleotides at the specific locations on the flow cell surface. The fluorescent label is cleaved yet again, and the procedure is repeated with each of the other nucleotide bases until the DNA is

sequenced to the appropriate length. Imaging is done with each addition of the nucleotides. This method can sequence a billion bases per hour. Hence, the time required to sequence complete genomes is considerably reduced.

Fourth Generation Sequencing

These methods are used especially in histology specimens where the sequencing is done in situ in the cells, preserving the spatial coordinates of RNA and DNA. This ultimately enables to map the sequenced data back to the histological context. This enables to identify tumor microenvironments to identify new targets for therapy. Examples of fourth generation sequencing includes single cell RNA sequencing (scRNA-seq) technology, in situ sequencing (ISS).

Comparison Of Various Sequencing Methods

Next generation sequencing was a great advancement over the Sanger's sequencing and helped ensure that this method could be commercially available for use. Some advantages over Sanger's sequencing include

1. Sequencing library construction and clonal amplification of DNA as part of the sequencing procedure
2. Array based sequencing in which DNA can be multiplexed and larger throughput achieved
3. Immobilisation of DNA on solid phase as the platform

The first NGS sequencer to be commercially available was the pyrosequencer. It had several advantages over Sanger's method in having a higher sensitivity, it was faster and more cost effective. The disadvantages of pyrosequencing includes:-

- a. Laborious sample preparation.²⁶
- b. Only short sequences can be sequenced.⁶
- c. Prone to produce errors as the reading was through homopolymeric sequences.⁶
- d. Expensive instruments.²⁶

The Illumina sequencer though offers a high throughput of data, requires very expensive instruments.²⁶

ABI/SOLiD has very high throughput and lesser cost of reagents but it takes a long duration of time to give results and the cost of the instrument is high.²⁶

The short read length and resequencing is a key disadvantage of the Sequencing by Ligation

technique. The Nanopore sequencing method has several advantages over the earlier sequencers given that it can read long lengths of fragments, is fluorescent tag-free,⁷ and that use of enzymes is remarkably low thereby reducing the need for stringent temperature controls.⁸ The problem with the Nanopore sequencer and the Pacific Biosciences Single Molecular Real time Reads is that they are expensive, and low through put.

Clinical Microbiology Applications of Next Generation Sequencing

Management of Outbreaks, identification and surveillance of pathogens, rapid identification of bacteria using the 16S-23S rRNA region, molecular case finding, taxonomy, metagenomics approaches on clinical samples, and the determination of the transmission of zoonotic microorganisms from animals to humans are just a few of the many applications of Next Generation sequencing in clinical microbiology.

Outbreak Management

Whole genome sequencing and NGS can be used to detect and monitor Outbreaks in hospital settings, in public health and epidemiological studies.²⁸ Apart from Outbreak tracing and characterisation, NGS can be used to implement control measures to prevent the spread of the infective agent.²⁵ This is particularly useful in outbreaks caused by Multidrug Resistant organisms^{30,37,31} and highly virulent strains of bacteria. This can be applied in the high risk areas initially such as transplant units, neonatal ICU'S etc. to promptly curtail any outbreaks.^{22,33,34} The main challenge here is the lack of user-friendly platforms in the bioinformatics software for the interpretation of data by diagnostic microbiologists.²² A global initiative to sequence and store genetic information about all known pathogens would help to curb outbreaks of emerging pathogens in the initial phase of community transmission.⁴⁷

Molecular Case Finding

A molecular case definition is given to the isolates producing the outbreaks and are further used in clinical and epidemiological investigations. These databases are used retrospectively in outbreaks later on to form a conclusive diagnosis. As a result, cases that would have been missed by standard epidemiological investigations can now be tracked down.⁴¹

Characterisation of Pathogens

NGS can be used to characterise the pathogens with respect to virulence characteristics and detection of novel resistance genes. This is of special significance in organisms such as *Mycobacterium tuberculosis*, where Next Generation sequencing methods can provide data regarding resistance against anti tubercular drugs precipitately within 8 – 9 days as it can be done directly from the specimen. This in turn helps the physician to administer well targeted therapy.²³ The epidemiology and evolution of pathogens such as *Yersinia pestis*, *Vibrio cholerae*, Methicillin resistant *Staphylococcus aureus* etc. has been studied using these sequencing technologies aiding in our understanding of the emergence of these epidemic clones and ultimately helping in preventing such epidemics.^{32,35}

Influenza Virus Vaccine Development

Influenza virus undergoes genetic reassortment, antigenic shift and antigenic drift thereby requiring active surveillance to detect the strain in circulation in the particular region and device vaccines according to the circulating strains. The evolution of human and animal influenza viruses are monitored all around the year to select the appropriate strains for vaccine development.²¹

Guiding Prevention Strategies in HIV

In certain circumstances where a sudden increase in the number of reported cases of HIV occurs, it is important to trace the source and detect HIV genomic sequences and analyse them with epidemiological data to characterise the outbreak and establish the transmission dynamics. This can be done accurately in a shorter duration of time using Next generation sequencing.²¹

Massively parallel sequencing (MPS) and Pyrosequencing can be used to detect single clones of coreceptor (CXCR 4) tropism of the HIV virus, thereby enabling the prediction of drug efficacy while using entry inhibitors.⁴²⁻⁴⁴

Detection of Emerging Pathogens

Use of sequencing is pivotal for diagnosis of emerging infections using genomic and metagenomic analysis.² The detection of potential bioterrorism agents can be done at a faster pace using Next Generation sequencing. Strain specific genotyping helps in tracing the contacts in such suspected cases.⁴⁵

Bacterial Taxonomy

Whole Genome sequencing and 16s rRNA gene are used to identify the taxonomic trees of bacteria.

Metagenomics

Metagenomic analysis of the organisms in a particular environment such as gut microbiome has vast implications. Research has been done to find the relationship between the gut microbiome and many illnesses such as diabetes, Irritable bowel syndrome, obesity etc. Targetted Next Generation Sequencing of the 16s-23s rRNA cluster region for characterisation of bacteria from clinical specimens can be used where the entire microbial population in a clinical sample is detected and analysed.³⁶

Detecting Zoonotic Transmission

Human and animal pathogens can be well distinguished by NGS with more conviction. Early and specific detection of such zoonotic transmission helps in precise and directed treatment targeting the right organism. It is also very important in curtailing community spread of these zoonosis. The virulence factors allowing such transmission also can be studied using NGS.

Diagnosis of Genetic Disorders by Next Generation Sequencing

In genetic disorders where the clinical diagnosis is unclear, genetic testing can help with precise diagnosis. Three methods of sequencing for diagnosis are broadly gene panel, exon and genome sequencing

Control of Antimicrobial Resistance

Whole genome sequencing has immense potential to help in curtailing antimicrobial resistance. The detection of specific genetic changes associated with the development of resistance,²⁷ to the antibiotic as well as the identification of newer targets of action for drugs can be identified which opens the possibilities of novel antibiotic development.²⁴

CONCLUSION

Next generation sequencing methods have evolved through several years to reach a point where it can be used in clinical diagnostic laboratories and research labs to create quality work. It helps to determine new targets for therapy

and diagnosis which in turn broadens the horizon for patient care.

ACKNOWLEDGMENTS

We would like to thank the Department of Microbiology, Sree Balaji Medical College and Hospital for their support.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest.

AUTHORS' CONTRIBUTION

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

None.

DATA AVAILABILITY

All datasets generated or analysed during this study are included in the manuscript.

ETHICS STATEMENT

This article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCES

1. Frey KG, Herrera-Galeano JE, Redden CL, et al. Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. *BMC Genomics*. 2014;15:96. doi: 10.1186/1471-2164-15-96
2. Ekblom R, Galindo J. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*. 2011;107(1):1-15. doi: 10.1038/hdy.2010.152
3. Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet*. 2019;20(6):341-355. doi: 10.1038/s41576-019-0113-7
4. Qin D. Next-generation sequencing and its clinical application. *Cancer Biol Med*. 2019;16(1):4-10. doi: 10.20892/j.issn.2095-3941.2018.0055
5. Bertelli C, Greub G. Rapid bacterial genome sequencing: methods and applications in clinical microbiology. *Clin Microbiol Infect*. 2013;19(9):803-813. doi: 10.1111/1469-0691.12217
6. Lam HYK, Clark MJ, Chen R, et al. Performance comparison of whole-genome sequencing platforms. *Nat Biotechnol*. 2012;30:78-82. doi: 10.1038/nbt.2065
7. Branton D, Deamer DW, Marziali A, et al. The

- potential and challenges of nanopore sequencing. *Nat Biotechnol.* 2008;26(10):1146-1153. doi: 10.1038/nbt.1495
8. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE. Structure of staphylococcal α -hemolysin, a heptameric transmembrane pore. *Science.* 1996;274(5294):1859-1865. doi: 10.1126/science.274.5294.1859
 9. Chang-Seok K. Recent Advances in the Clinical Application of Next-Generation Sequencing. *Pediatr Gastroenterol Hepatol Nutr.* 2021;24(1):1-6. doi: 10.5223/pghn.2021.24.1.1
 10. Levy SE, Myers RM. Advancements in Next-Generation Sequencing. *Annu Rev Genom Hum Genet.* 2016;17(1):95-115. doi: 10.1146/annurev-genom-083115-022413
 11. Schuster SC. Next-generation sequencing transforms today's biology. *Nat Methods.* 2008;5(1):16-18. doi: 10.1038/nmeth1156
 12. Shendure J, Porreca GJ, Reppas NB, et al Accurate multiplex polony sequencing of an evolved bacterial genome. *Science.* 2005;309:1728-1732. doi: 10.1126/science.1117389
 13. Gargis AS, Kalman L, Berry MW, et al. Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat Biotechnol.* 2012;30(11):1033-1036. doi: 10.1038/nbt.2403
 14. Moore RA, Warren RL, Freeman JD, et al. The sensitivity of massively parallel sequencing for detecting candidate infectious agents associated with human tissue. *PLoS One.* 2011;6(5):e19838. doi: 10.1371/journal.pone.0019838
 15. Tellez-Sosa J, Rodriguez MH, Gomez-Barreto RE, et al. Using high-throughput sequencing to leverage surveillance of genetic diversity and oseltamivir resistance: a pilot study during the 2009 influenza A(H1N1) pandemic. *PLoS One.* 2013;8(7):e67010. doi: 10.1371/journal.pone.0067010
 16. Loman NJ, Misra RV, Dallman TJ, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol.* 2012;30(5):434-439. doi: 10.1038/nbt.2198
 17. Quail MA, Smith M, Coupland P, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics.* 2012;13:341. doi: 10.1186/1471-2164-13-341
 18. Merriman B, Ion Torrent R, Team D, Rothberg JM. Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis.* 2012;33(23):3397-3417. doi: 10.1002/elps.201200424
 19. Mende DR, Waller AS, Sunagawa S, et al. Assessment of metagenomic assembly using simulated next generation sequencing data. *PLoS One.* 2012;7(2):e31386. doi: 10.1371/journal.pone.0031386
 20. Applications of Clinical Microbial Next-Generation Sequencing: Report on an American Academy of Microbiology Colloquium held in Washington, DC, in April 2015. Washington (DC): American Society for Microbiology. doi: 10.1128/AAMCol.Apr.2015
 21. Gwinn M, MacCannell DR, Khabbaz RF. Integrating Advanced Molecular Technologies into Public Health. *J Clin Microbiol.* 2017;55(3):703-714. doi: 10.1128/JCM.01967-16
 22. Sherry NL, Porter JL, Seemann T, Watkins A, Stinear TP, Howden BP. Outbreak investigation using high-throughput genome sequencing within a diagnostic microbiology laboratory. *J Clin Microbiol.* 2013;51(5):1396-1401. doi: 10.1128/JCM.03332-12
 23. Maladan Y, Krismawati H, Wahyuni T, et al. The whole-genome sequencing in predicting Mycobacterium tuberculosis drug susceptibility and resistance in Papua, Indonesia. *BMC Genomics.* 2021;22(1):844. doi: 10.1186/s12864-021-08139-3
 24. Koser CU, Ellington MJ, Peacock SJ. Whole-genome sequencing to control antimicrobial resistance. *Trends Genet.* 2014;30(9):401-407. doi: 10.1016/j.tig.2014.07.003
 25. Koser CU, Holden MT, Ellington MJ, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med.* 2012;366(24):2267-2275. doi: 10.1056/NEJMoa1109910
 26. Niedringhaus TP, Milanova D, Kerby MB, Snyder MP, Barron AE. Landscape of next-generation sequencing technologies. *Anal Chem.* 2011;83(12):4327-4341. doi: 10.1021/ac2010857
 27. Livermore DM, Wain J. Revolutionising bacteriology to improve treatment outcomes and antibiotic stewardship. *Infect Chemother.* 2013;45(1):1-10. doi: 10.3947/ic.2013.45.1.1
 28. Koser CU, Ellington MJ, Cartwright EJ, et al. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathog.* 2012;8:e1002824. doi: 10.1371/journal.ppat.1002824
 29. Maitra RD, Kim J, Dunbar WB. Recent advances in nanopore sequencing. *Electrophoresis.* 2012;33(23):3418-3428. doi: 10.1002/elps.201200272
 30. Cartwright EJ, Koser CU, Peacock SJ. Microbial sequences benefit health now. *Nature.* 2011;471(7340):578. doi: 10.1038/471578d
 31. Harris SR, Cartwright EJ, Torok ME, et al. Whole-genome sequencing for analysis of an outbreak of methicillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect Dis.* 2013;13(2):130-136. doi: 10.1016/S1473-3099(12)70268-2
 32. Parkhill J, Wren BW. Bacterial epidemiology and biology—lessons from genome sequencing. *Genome Biol.* 2011;12(10):230. doi: 10.1186/gb-2011-12-10-230
 33. Shea PR, Beres SB, Flores AR, et al. Distinct signatures of diversifying selection revealed by genome analysis of respiratory tract and invasive bacterial populations. *Proc Natl Acad Sci U S A.* 2011;108(12):5039-5044. doi: 10.1073/pnas.1016282108
 34. Gardy JL, Johnston JC, Ho Sui SJ, et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med.* 2011;364(8):730-739. doi: 10.1056/NEJMoa1003176
 35. Holt KE, Parkhill J, Mazzoni CJ, et al. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella Typhi*. *Nat Genet.* 2008;40(8):987-993. doi: 10.1038/ng.195
 36. Hall N. Advanced sequencing technologies and their wider impact in microbiology. *J Exp Biol.* 2007;210(Pt

- 9):1518-1525. doi: 10.1242/jeb.001370
37. Brzuszkiewicz E, Thurmer A, Schuldes J, et al. Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Arch Microbiol.* 2011;193(12):883-891. doi: 10.1007/s00203-011-0725-6
 38. Chin CS, Sorenson J, Harris JB, et al. The origin of the Haitian cholera outbreak strain. *N Engl J Med.* 2011;364(1):33-42. doi: 10.1056/NEJMoa1012928
 39. Harris SR, Clarke IN, Seth-Smith HMB, et al. Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. *Nat Genet.* 2012;44(4):413-419. doi: 10.1038/ng.2214
 40. Eyre DW, Golubchik T, Gordon NC, et al. A pilot study of rapid benchtop sequencing of *Staphylococcus aureus* and *Clostridium difficile* for outbreak detection and surveillance. *BMJ Open* 2012;2(3):e001124. doi: 10.1136/bmjopen-2012-001124
 41. McAdam PR, Templeton KE, Edwards GF, et al. Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. *Proc Natl Acad Sci U S A.* 2012;109(23):9107-9112. doi: 10.1073/pnas.1202869109
 42. Daumer M, Kaiser R, Klein R, Lengauer T, Thiele B, Thielen A. Genotypic tropism testing by massively parallel sequencing: qualitative and quantitative analysis. *BMC Med Inform Decis Mak.* 2011;11:30. doi: 10.1186/1472-6947-11-30
 43. Saliou A, Delobel P, Dubois M, et al. Concordance between two phenotypic assays and ultradeep pyrosequencing for determining HIV-1 tropism. *Antimicrob Agents Chemother.* 2011;55(6):2831-2836. doi: 10.1128/AAC.00091-11
 44. Swenson LC, Mo T, Dong WWY, et al. Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients. *J Infect Dis.* 2011;203(2):237-245. doi: 10.1093/infdis/jiq030
 45. Kuroda M, Sekizuka T, Shinya F, et al. Detection of a possible bioterrorism agent, *Francisella* sp., in a clinical specimen by use of next-generation direct DNA sequencing. *J Clin Microbiol.* 2012;50(5):1810-1812. doi: 10.1128/JCM.06715-11
 46. Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. *Clin Chem.* 2009;55(4):641-658. doi: 10.1373/clinchem.2008.112789
 47. Gupta R, Michalski MH, Rijsberman FR. Can an infectious disease genomics project predict and prevent the next pandemic? *PLoS Biol.* 2009;7(10):e1000219. doi: 10.1371/journal.pbio.1000219
 48. Chain PSG, Grafham DV, Fulton RS, et al. Genomics: Genome project standards in a new era of sequencing. *Science.* 2009;326(5950):236-237. doi: 10.1126/science.1180614
 49. Forgetta V, Oughton MT, Marquis P, et al. Fourteen-genome comparison identifies DNA markers for severe-disease-associated strains of *Clostridium difficile*. *J Clin Microbiol.* 2011;49(6):2230-2238. doi: 10.1128/JCM.00391-11
 50. Jeffrey BM, Suchland RJ, Quinn KL, Davidson JR, Stamm WE, Rockey DD. Genome sequencing of recent clinical *Chlamydia trachomatis* strains identifies loci associated with tissue tropism and regions of apparent recombination. *Infect Immun.* 2010;78(6):2544-2553. doi: 10.1128/IAI.01324-09
 51. Lieberman TD, Michel JB, Aingaran M, et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat Genet.* 2011;43(12):1275-1280. doi: 10.1038/ng.997