



The application of next generation sequencing technology in medical diagnostics: a perspective

Anirban Bhar¹

Received: 10 June 2022 / Accepted: 9 August 2022 / Published online: 22 August 2022
© Indian National Science Academy 2022

Abstract

Rapid isolation, characterization, and identification are prerequisites of any successful medical intervention to infectious disease treatment. This is a real challenge to the scientific as well as a medical community to find out a proper and robust method of pathogen detection. Classical cultural, as well as biochemical test-based identification, has its own limitations to their time-consuming and ineffectiveness for closely related pathogens. Molecular diagnostics became a popular alternative to classical techniques for the past couple of decades but it required some prior information to detect the pathogen successfully. Recently, with the advent of next-generation sequencing (NGS) technology identification, and characterization of almost all the pathogenic bacteria become possible without any information a priori. Metagenomic next generation sequencing is another specialized type of NGS that is profoundly utilized in medical biotechnology and diagnostics now a days. Therefore, the present review is focused on a brief introduction to NGS technology, its application in medical microbiology, and possible future aspects for the development of medical sciences.

Keywords Next generation sequencing · Pathogen identification · Whole genome sequencing · Medical microbiology · Molecular diagnostics

Introduction

Medical microbiology is an important field of research since the past, which deals with the identification, characterization, and prevention of any disease-causing microorganisms. The infectious disease can be caused by a virus, bacteria, fungi, algae as well as many other eukaryotic microscopic organisms too (Daszak et al. 2000). Initial diagnosis of any infectious disease largely depends upon successful, proper as well as quick detection of the pathogen (Stafford et al. 2020). The identification of the pathogen is solely dependent on culture methods in diagnostic laboratories. These culture methods are time-consuming as different disease-causing microorganisms demand varied incubation times for their required growth. Many molecular and biochemical test kits are also available in the market for the quick identification of the pathogen.

Although kit-based methods are quick, it demands more accuracy in many cases. The pathogenic variability and co-evolutionary modifications of the microorganisms evade many pathogens from successful identification in many cases (Deurenberg et al. 2017). For this problem, molecular diagnostics become increasingly popular day by day (Taylor et al. 2015). During the last two decades, molecular characterization of the pathogen is customarily done through ribosomal DNA (rDNA) sequencing methods. The prokaryotic microorganisms are characterized through 16 s rDNA sequencing and characterization of eukaryotic microorganisms are done through 18 s rDNA sequencing technology (Deurenberg et al. 2017). The sequence data is then matched with available rDNA sequence data using a different molecular analysis tool for the accurate detection of the pathogen (John et al. 2021). This technique can identify the pathogen (both culturable and non-culturable) correctly but it needs some specific background information to perform the experiment. Any novel pathogenic species is not only very much difficult to identify through this technique but characterization, as well as probable prevention strategy determination, also become a big problem. The repetition of sequence experiment with each sample

✉ Anirban Bhar
anirbanbhar@rkmcvcrachara.ac.in; anirbanbhar06@gmail.com

¹ Post Graduate Department of Botany, Ramakrishna Mission Vivekananda Centenary College, Rahara, Kolkata 700118, India

and performing molecular analyses every time also needs some expertise from the laboratory personnel. On the other hand, generation of molecular signatures for the pathogens will probably be the answer for every problem. The production of molecular signature tools can only be possible after quick and complete molecular characterization of any pathogen. Next-generation sequencing technology brings the opportunity for robust and efficient characterization of the whole genome of any pathogenic organism at any point in time (Deurenberg et al. 2017). NGS can allow the rapid characterization of any unknown pathogenic microorganisms too (Li et al. 2020). This information will not only be utilized for the identification of the pathogen but the flood of sequence information can also help to study detailed molecular host–pathogen interaction dynamics for a particular pathosystem (Leekitcharoenphon et al. 2014). This technology will solve many complex puzzles in medical sciences to promote prevention and drug design attempts to the new direction in the future. The rapid and accurate detection of pathogens by NGS help to identify widespread pathogen during any disease outbreak and help to provide “point of care” to the individual patients (Gwinn et al., 2019). In this view, the present review focuses on a brief introduction to next-generation sequencing techniques and the application of the same in the identification of different pathogenic microorganisms. The possibility of this revolutionary molecular biological technique in future diagnostics has also discussed with its challenges and many pertinent questions.

Next-generation sequencing (NGS) technology, an introduction

Next-generation sequencing is named as it is the generation next to the universal Sanger sequencing techniques. Although the basic principle of sequencing is the same, NGS provides wide possibilities concerning time, size of information, and accuracy. Most importantly NGS doesn't require any previous genetic information to sequence a genome (Behjati and Tarpey 2013). A reference genome sequence can be generated from the de novo sequencing data of any genome of interest and the gene sequences are aligned over that reference sequence (Serwecińska 2020). Many sequencing platforms are available recently to perform the analysis according to the requirement (Table 1). The NGS sequencing can be applied to the genome (DNA Seq) or transcriptome (RNA Seq) to decipher the sequence information of the nucleotides (Ansorge 2009; Wilson et al. 2014). The DNA and/or RNA are sheared at first into smaller fragments then universal adaptor primers are ligated and polymerase chain reaction (PCR) based multiplication is performed, which is termed as clonal amplification (Behjati and Tarpey 2013). The sequencing of each of these fragments is then performed i.e., multiple sequence information is obtained for a single fragment of nucleotide (deep sequencing techniques). This method increases the chance of gaining insight into the low abundance as well as rare transcripts in the sample (Serwecińska 2020). The information is quantified in terms of fluorescence value for each

Table 1 Principle NGS platforms used in medical microbiology

Next generation sequencing platform	Sequencing chemistry	Read length	Relevant URL
Roche 454-titanium FLX	Pyrosequencing	400 bp	https://www.roche.com/media/releases/med-cor-2011-06-28t.htm
Roche 454-GS Junior	Pyrosequencing	400 bp, 50 MB per run	https://www.roche.com/media/releases/med-cor-2011-06-28t.htm
Illumina/Solexa HiSeq 2000	Reverse Terminator Chemistry	600 bp per run	https://www.illumina.com/documents/products/datasheets/datasheet_hiseq2000.pdf
Illumina/Solexa MiSeq	Reverse Terminator Chemistry	2 × 150 bp 300 bp	https://www.illumina.com/systems/sequencing-platforms/miseq.html
Helicos Biotechnologies	Reverse Terminator Chemistry	25–55 bp, 28 GB per run	www.helicosbio.com http://seqll.com/
ABI/ Life Technology SOLiD 5550XL	Ligation chemistry	2 × 60 bp, 15 GB per day	http://www.columbia.edu/cu/biology/courses/w3034/Dan/readings/SOLiD_System_Brochure.pdf
ABI/ Life technology/Ion torrent	H + ion sensitive transistor	320 MB per run	https://www.thermofisher.com/in/en/home/brands/ion-torrent.html
Oxford Nanopore MinION	Flow cells with nanopores on electro-resistant membrane. Electrophoretic movement of nucleic acids generates characteristic ‘squiggle’, decoded by base call algorithms	~ 10 KB to 150 KB 1D read length 300 KB; 2D read length 60 KB Realtime extended reads	https://nanoporetech.com/products/minion



base and the intensity calculation will ultimately determine the abundance of the nucleotide fragment. In the case of transcriptome sequencing, differential expression of genes (DEGs) can also be deduced from the base mean values of each transcript (Gupta et al. 2017). In the case of DNA seq, the different fragments are aligned and mapped against the published genomic sequence or de novo reference sequence of the genome (Ansorge 2009). This analysis required state of art bioinformatics, statistical as well as computation facilities. The huge data generated through this NGS is very much difficult to handle and required expertise to extract valuable information out of that flood of information. The storage of the raw data for future analyses is another challenge but this information, if properly managed, has huge possibilities in pathogen detection as well as drug target determination in clinical microbiology.

Metagenomic next-generation sequencing (mNGS)

The NGS technology gradually replaces other molecular techniques to identify pathogenic microorganisms due to its rapid and flexible analytical capabilities. This technique can simultaneously and continuously sequence billions of nucleotide fragments. The transcripts originated due to individual sequencing events then overlapped to generate contigs. Finally, these contigs are mapped against a reference genome as discussed earlier. Metagenomic NGS (mNGS) is a specialized form of NGS technology, where a mixed population of nucleotides can be successfully sequenced and matched with the cognate genome to identify the types and approximate numbers of the microorganisms present in a given sample. So, mNGS can able to identify nucleotides from microorganisms that belong to diverse taxonomic groups. This method has tremendous application in the identification of poly pathogenic organisms, cause of any outbreak, mutation analysis, disease tracking, etc. Different clinal samples e.g., blood, body fluids, ocular fluid, respiratory samples, cerebrospinal fluid, alveolar fluids, etc. can be used as samples for mNGS analysis (<https://asm.org/Articles/2019/November/Metagenomic-Next-Generation-Sequencing-How-Does-It> accessed on 1st January 2022). Targeted PCR-based identification of polymicrobial infection has been performed using specific primers designed from conserved 16S ribosomal RNA (16S rRNA) and internal transcribed spacer (ITS) sequences. Although, this broad PCR technique has long been used to identify pathogens, due to the use of specific primers set they are considered to be biased as compared to mNGS. Multiple base calls originated from 16S rRNA sequencing of polymicrobial samples develop mixed nucleotide chromatogram is very common to constrain in microbial identification (Lecuit and Eloit 2014). Sometimes

deconvolutional computation methods are used to counteract this problem, mNGS can efficiently minimize such limitations to rapidly identify mixed microbial clinical pathogens (Han et al. 2019).

Application of next-generation sequencing technology in medical microbiology

NGS technology has an enormous application on medical microbiology and the prognosis of infectious diseases. In 2014, the Centers for Disease Control and Prevention (CDC) established Advanced Molecular Detection (AMD) program that integrates pathogen genome sequencing by NGS with disease surveillance mechanisms in US public health system (Gwinn et al. 2019). NGS is proved to be an expedient approach for whole-genome sequencing of many microorganisms which helps the scientist to find out many questions regarding pathogens e.g., genome information, the virulence of the pathogen, antibiotic sensitivity, etc. (Didelot et al. 2012). Not only restricted to the bacteriology virus identification and viral transcriptomics is also revolutionized on the advent of next-generation sequencing platforms (Sabapathypillai et al. 2021). Many infectious viral particles e.g., influenza virus, human immunodeficiency virus (HIV), human hepatitis C virus quasi-species, etc. are characterized and host modifications are identified through this technique (Barzon et al. 2011). The NGS-based diagnostics has now widely been used to identify different pathogenic microorganisms from critical clinical samples (Table 2).

NGS in molecular diagnostics

Molecular diagnostic is an emerging field in medical microbiology and gaining continuous popularity with the advent of NGS technology. Molecular diagnostic is a wide subject and recent developments generate huge numbers of examples. With the limited scope of this article, some of the recent interesting findings in molecular diagnostics which are related to NGS technology are discussed here. The microbiome present in bronchoalveolar lavage (BAL) samples collected from the patients with Clinical Pulmonary Infection Score (CPIS) ≥ 6 points admitted in a surgical intensive care unit (SICU) were analyzed through NGS (Janda and Abbott 2007). Recently, an interesting finding was observed in BAL fluid (BALF) diagnostics performed with 35 critically ill immunocompromised or non-immunocompromised patients. It had been observed that metagenomic next-generation sequencing (mNGS) is more sensitive and specific in pathogenic detection than PCR-based techniques. Nonetheless, 13 samples were found positive by mNGS based diagnostics which were previously found negative by culture/smear and PCR-based technique (Liu et al. 2021). Besides, microbial

Table 2 Application of next generation sequencing (NGS) technology in identifying pathogenic microorganisms

Platform	Sample	Pathogen	Reference
NGS (Sequence-based ultra-rapid pathogen identification, SUPI)	Clinical samples	Mixed population	Naccache et al. (2014)
NGS	Cerebrospinal fluid	<i>Streptococcus pneumoniae</i>	Guo et al. (2019)
NGS (BGISEQ-50 platform)	Cerebrospinal fluid	<i>Toxoplasma gondii</i>	Hu et al. (2018)
NGS (Illumina, Nanopore)	Body fluids	Mixed population	Gu et al. (2020)
NGS (Nanopore, MinION)	Blood	Bacteremia (Sepsis)	Grumaz et al. (2020)
NGS	Blood serum	Murine typhus (<i>Rickettsia typhi</i>)	Stafford et al. (2020)
mNGS	Ophthalmic fluid	Corneal disease, Conjunctivitis e.g., <i>Vittaforma corneae</i> , <i>Capnocytophaga keratitis</i> etc	Sabapathypillai et al. (2021)
NGS along with pan fungal PCR	BALF	<i>Scedosporium apiospermum</i> , <i>Fusarium</i> spp., <i>Aspergillus</i> spp., <i>Candida</i> spp., <i>Fusarium</i> spp.	Kidd et al. (2020)
NGS (Nanopore sequencing)	Respiratory samples	Fungal and Bacterial pathogens	Chan et al. (2020)
NGS	Blood samples	Mixed population	Geng et al. (2021)
mNGS	BALF	<i>Mycobacterium tuberculosis</i>	Liu et al. (2021)
mNGS	Clinical sample from transplant recipients	cytomegalovirus, Epstein–Barr virus, BK virus, human adenovirus, JC virus, herpes simplex virus 1 and 2, varicella zoster virus, human herpesvirus 6A and 6B, and parvovirus B19 et	Sam et al. (2021)
NGS (Poly-dynamic modelling)	Clinical blood samples	COVID 19	John et al. (2021)

detection NGS also has tremendous application in cancer biology and diagnostics. Recently, The European Society for Medical Oncology (ESMO) has approved the application of NGS in some oncogenic detection e.g., advanced non-squamous non-small-cell lung cancer (NSCLC), prostate cancers, ovarian cancers, and cholangiocarcinoma (Naccache et al. 2014). Besides, primary immunodeficiency diseases (PID) could also be detected early and molecular diagnostics provide immense opportunity in the determination of medical strategies (Tang et al. 1997).

Application of NGS in the identification of bacterial pathogens

In May 2011 O104:H4 strain of *Escherichia coli* caused an outbreak and many deaths in Germany due to hemolytic uremic syndrome (HUS). The cause of the outbreak remained unidentified as the particular serotype was rarely detected within animals. The whole-genome sequencing and the characterization of *E. coli* O104:H4 strain were performed using NGS technology in a record minimum time (Mosele et al. 2020). Later on, an outbreak of methicillin-resistant *Staphylococcus aureus* (MRSA) in a neonatal intensive care unit has been detected through NGS and prevented promptly after determination (Boheemen et al. 2020). NGS-based mutagenicity assay has been performed to identify genome-wide chemically induced mutations in *Salmonella typhimurium* LA 2 strains (TA98 and TA100)

(McInerney et al. 2017). The multidrug-resistant *Salmonella typhimurium* ST313 strain is an emerging threat to mankind and it necessitates the whole genome analysis of all the typhoidal as well as non-typhoidal *Salmonella* sp. in detail (Wang et al. 2009). *Salmonella enterica* has different close serovars that cause gastrointestinal infections in massive ways. The closely related serovars are very much hard to detect properly because of their genomic similarities. Recently, WG-NGS associated with single nucleotide polymorphism (SNP) can able to detect *Salmonella enterica* serovar Enteritidis outbreak rapidly in an efficient fashion (Tettelin et al. 2005). Another detailed analysis has been performed to compare the whole genome of *Salmonella typhimurium*, *S. enteric*, and *S. derby* to determine the SNP tree for *Salmonella* (Lefterova et al. 2015). Adult respiratory distress syndrome caused by *Chlamydia psittaci* is a severe but rare disease and is very difficult to identify in its early stage of infection. It causes advanced infection associated with non-responsiveness of antibiotics and septic shock to the patients. A very efficient mNGS based detection tool has been developed recently to diagnose psittacosis pneumonia that reduces the time between infection and detection and hence the severity of the disease (Chen et al. 2020). Thus, NGS technology has an immense impact on medical microbiology and infectious disease control. Presently this technology is well tested and waiting for the widespread application for the betterment of the society globally.



Application of NGS in the identification of pathogenic virus

Viral diseases are always challenging due to the unavailability of effective anti-viral drugs and generally high rate of transmission of the disease. On the other hand, the virus is capable of frequent mutations, which exhibit altered symptoms, transmissibility and make existing medication strategies ineffective. The development of a proper identification strategy of the virus, as well as mutant variants, would provide the opportunity to contain the spread of the disease as well as antiviral therapy. Meningoencephalitis is very much difficult to identify by routine laboratory tests as many causal organisms are producing similar types of symptoms but the NGS technique can identify the causal organism from different source samples easily (Gwinn et al. 2019). In another case, unbiased metagenomic next-generation sequencing has revealed Astrovirus infection in an immunosuppressed adult due to allelogenic bone marrow transplantation with acute encephalitis (Radford et al. 2012). Recently, metagenomic next-generation sequencing (mNGS) has been standardized to identify viral meningoencephalitis in a cost-effective and resource-limited situation (Hong et al. 2020). Among many other viral diseases, flu caused by the influenza virus is the reason for many deaths across the world. Despite routine vaccination, several mutant strains are generated which can easily be detected using NGS techniques (Wain et al. 2013). Most respiratory infections are caused by viral infections. Sometimes mixed population of virus make the situation complicated to detect through a classical mode of detection. Such kind of “pan-pathogenic” detection e.g., simultaneous detection of DNA and RNA virus is possible by mNGS technique (Poelvoorde et al. 2020). Blood and BALF are the common sampling method for most NGS-based diagnostics of virus. In a study with 109 patient groups different other samples were also used e.g., tissue, sputum, pleural effusion, cerebrospinal fluid (CSF), pus, bone marrow, and nasal swab, etc. above-mentioned samples were analyzed using classical as well as mNGS based detection tools in different groups. In all the cases mNGS was found to be more effective in detection efficiency as sensitivity (Duan et al. 2021). The recent pandemic outbreak caused by severe acute respiratory syndrome coronavirus 2 (SARS CoV 2) costs huge loss across the globe. Human civilization has faced continuous challenges due to the rapid mutation of this virus and subsequent waves of infections (Bhar 2021). Different platforms of NGS had been tested to detect mutational variations and to decipher genomic complexities (Chiara et al. 2021). This rapid sequencing capacity, generation of huge data, and greater sensitivity towards single nucleotide polymorphism (SNP) detection capacitate generation of a huge genomic information repository in a record minimum time. This data can be utilized in anti-viral drug discovery

as well as effective vaccine hunts. Besides detection, NGS can also be used as a tool for surveying epidemiology of the infectious disease like COVID 19. The huge data originating using this technology may also be utilized to evaluate the possible evolutionary origin of this new pathogenic virus. Several “poly-dynamic modeling” based approaches in monitoring SARS CoV 2 had been proposed recently (Kidd et al. 2020). Although NGS has enormous possibilities in molecular diagnostics particularly monitoring pandemic situations, massive operational cost per sample and specialized skill to perform and analysis are the key constraints.

“Pan-genome analysis”, the future of molecular diagnostics?

The word “pan” originated from a Greek word that means “whole” or “entire”. The term “pan-genome” was first used by Tettelin et al., 2005 while describing genome analyses of different isolates of *Streptococcus agalactiae* (Ugolotti et al. 2017). Identification of polymicrobial samples is a challenging task. Besides microbial identification, pan-genome analysis has wide application in therapeutic strategies, gene manipulative techniques, development of antiphase defense, analysis of antimicrobial resistance response, etc. (Kommedal et al. 2008). The uncontrolled and non-specific use of many antimicrobials especially, antibiotics tend to generate many antibiotic-resistant bacteria (Socransky and Haffajee 1992). These concerning pathogens have been categorized as “priority pathogens for R&D of new antibiotics” by World Health Organizations (WHO, <http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/>). On the other hand, different “emerging infectious pathogens” have been categorized by the National Institute of Allergy and Infectious Diseases (NIAID, <https://www.niaid.nih.gov/research/emerging-infectious-diseases-pathogens>), which may consider being threats of many future outbreaks (Kommedal et al. 2008). The pan-genome analysis develops the “core genome” and “accessory genome” of a given clade of microorganisms. Core genomes are conserved and common among the members of that particular clade, hence, constituting essential genes for the growth, development, and vitality of that particular clade. Whereas, accessory genomes are the assemblage of genes that are present in a single or some members of that clade (Mellmann et al. 2011). This data is not only been used for the identification of pathogenic microorganisms but entails evolutionary trajectories of those pathogens (Kommedal et al. 2008). Besides, this pan-genomic data can be successfully used to develop antimicrobial drugs targeting a group of pathogenic organisms belonging to a particular clade. Antibiotic resistance response can also be analyzed easily with this data and possible strategies will be adopted to counteract the same.

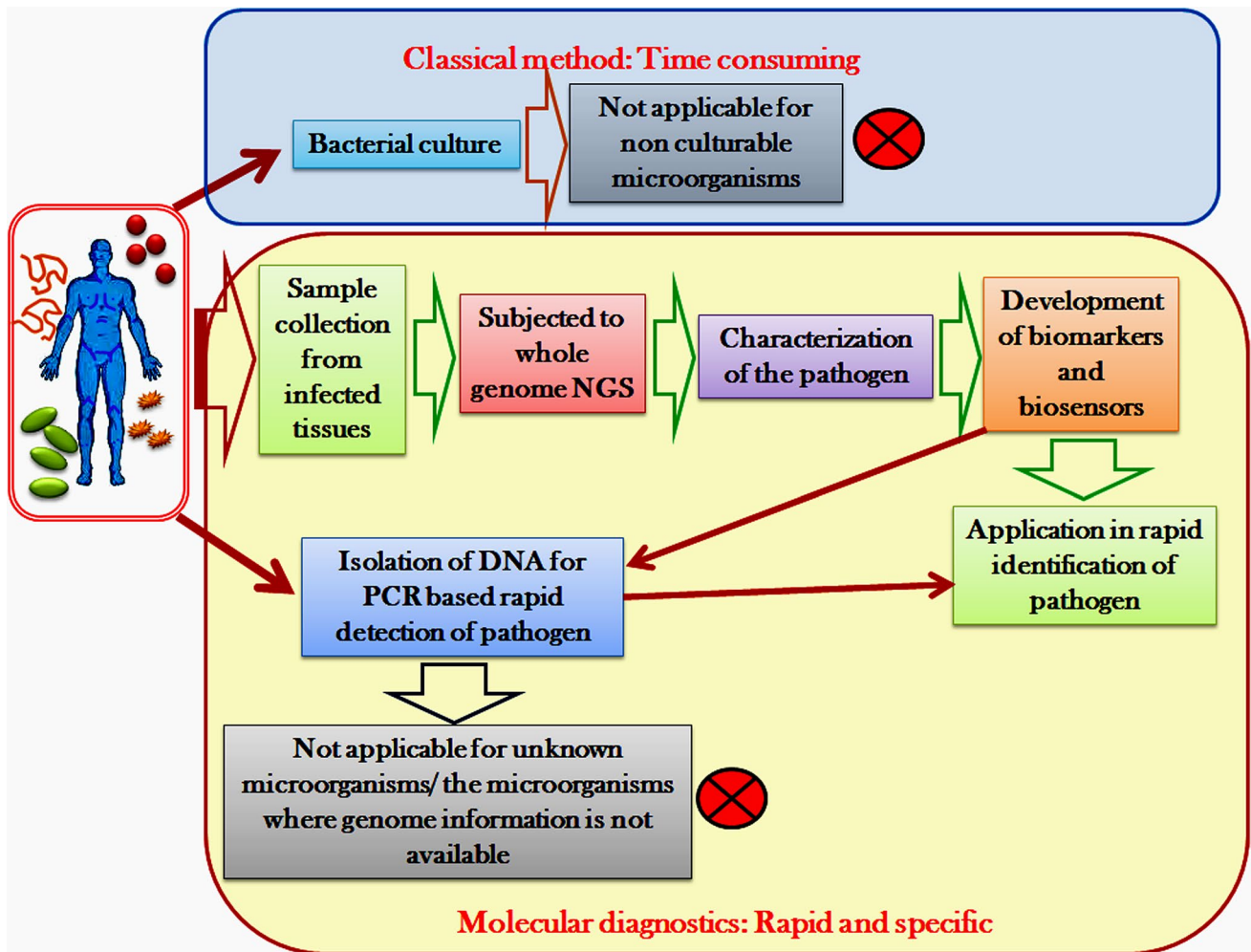


Fig. 1 The schematic representation of different approaches for the identification of pathogenic microorganisms from clinical samples

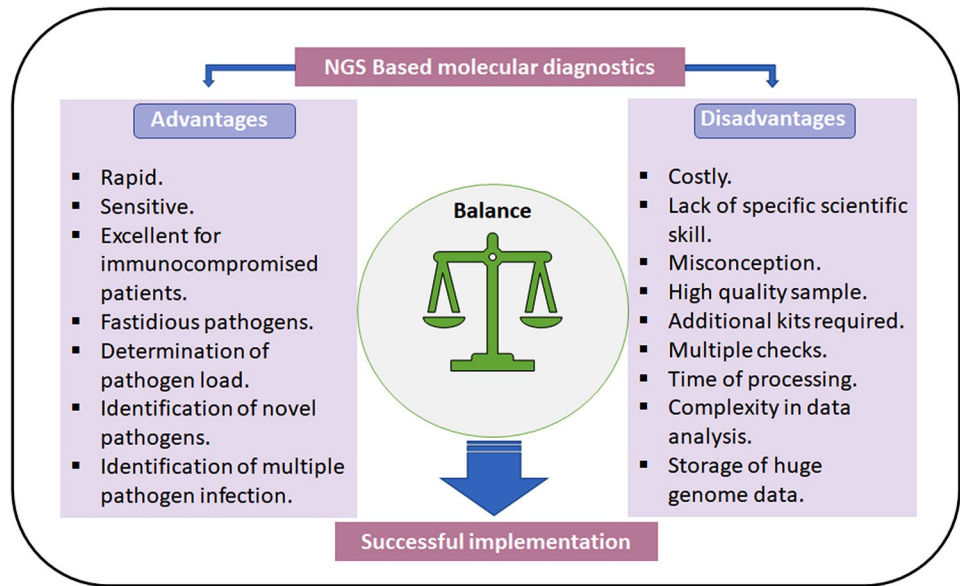
Conclusion and future questions

The utmost priority is always given to the medical sciences and allied research globally. The application of cutting-edge technologies in the identification of pathogenic microorganisms is always necessary for proper diagnostics. Culture-based identification of the pathogen is widely used but demands huge time and is not applicable for non-culturable microorganisms. Molecular diagnostics is presently appreciated for its efficiency as well as rapid identification of the pathogen. Presently many molecular diagnostics tool/ kit is available in the market for pathological screening but lack of sensitivity to discriminate two closely related serovars is a major problem. Polymerase chain reaction (PCR) based methods are found to be efficient in many cases but lack of genome information for many pathogenic organisms restricts its full utilization in laboratories. Recently, next-generation sequencing technology revolutionize medical microbiology dramatically. Although, NGS-based whole

genome sequencing and characterization of many known as well as unknown microorganisms have already been started, a huge database is required to address the future challenge in this field. The storage of huge data and specific scientific expertise is a big challenge for its application in medical sciences. On the other hand, the ability to detection of pathogens from metagenomic samples largely depends on sample quality, complexity, and range of pathogen diversity of the said sample. Less complex samples have a higher chance of detection of pathogens, even organism-specific microorganisms can be detected faster. Besides, intrinsic characteristics of the pathogen e.g., excessively small genome size may interfere with proper identification by NGS methods; in such cases, PCR-based methods are best suited (Frey and Bishop-Lilly 2015). Judicial utilization and an organized scientific approach can meet the target without any difficulties. Along with the sequencing identification of proper markers for the pathogen and the development of sequence-based biosensors will resolve the crisis in the future (Fig. 1). Besides, the



Fig. 2 The schematic representation of different advantages and disadvantages of NGS-based molecular diagnostics upon successful implementation in the health care system



cost of NGS as compared to the other classical diagnostics is another constraint for its utilization in full potential in developing countries. Although a long way to go but NGS based techniques will occupy the position of an important molecular diagnostics tool in near future (Fig. 2).

Funding No funding has received to perform this work.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Conflicts of interest Not applicable.

References

- Ansorge, W.J.: Next-generation DNA sequencing techniques. *N. Biotechnol.* **25**, 195–203 (2009). <https://doi.org/10.1016/j.nbt.2008.12.009>
- Barzon, L., Lavezzo, E., Militello, V., Toppo, S., Palù, G.: Applications of next-generation sequencing technologies to diagnostic virology. *Int. J. Mol. Sci.* **12**, 7861–7884 (2011). <https://doi.org/10.3390/ijms12117861>
- Behjati, S., Tarpey, P.S.: What is next generation sequencing? *Arch. Dis. Child. Educ. Pract.* **98**, 236–238 (2013). <https://doi.org/10.1136/archdischild-2013-304340>
- Bhar, A.: Is it possible to ensure COVID19 vaccine supply by using plants? *Nucleus* **64**, 137–141 (2021). <https://doi.org/10.1007/s13237-021-00361-4>
- Chan, W.S., Au, C.H., Leung, S.M., Ho, D.N., Wong, E.Y.L., To, M.Y., Ng, M.K., Chan, T.L., Ma, E.S.K., Tang, B.S.F.: Potential utility of targeted Nanopore sequencing for improving etiologic diagnosis of bacterial and fungal respiratory infection. *Diagn Pathol.* (2020). <https://doi.org/10.1186/s13000-020-00960-w>
- Chen, X., Cao, K., Wei, Y., Qian, Y., Liang, J., Dong, D., Tang, J., Zhu, Z., Gu, Q., Yu, W.: Metagenomic next-generation sequencing in the diagnosis of severe pneumonias caused by *Chlamydia psittaci*. *Infection* **48**, 535–542 (2020). <https://doi.org/10.1007/s15010-020-01429-0>
- Chiara, M., D’Erchia, A.M., Gissi, C., Manzari, C., Parisi, A., Resta, N., Zambelli, F., Picardi, E., Pavesi, G., Horner, D.S., Pesole, G.: Next generation sequencing of SARS-CoV-2 genomes: challenges, applications and opportunities. *Brief Bioinform.* **22**, 616–630 (2021). <https://doi.org/10.1093/bib/bbaa297>
- Daszak, P., Cunningham, A.A., Hyatt, A.D.: Emerging infectious diseases of wildlife—threats to biodiversity and human health. *Science* **287**, 443 (2000). <https://doi.org/10.1126/science.287.5452.443>
- Deurenberg, R.H., Bathoorn, E., Chlebowicz, M.A., Couto, N., Ferdous, M., García-Cobos, S., Kooistra-Smid, A.M.D., Raangs, E.C., Rosema, S., Veloo, A.C.M., Zhou, K., Friedrich, A.W., Rossen, J.W.A.: Application of next generation sequencing in clinical microbiology and infection prevention. *J. Biotechnol.* **243**, 16–24 (2017). <https://doi.org/10.1016/j.jbiotec.2016.12.022>
- Didelot, X., Bowden, R., Wilson, D.J., Peto, T.E., Crook, D.W.: Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* **13**, 601–612 (2012). <https://doi.org/10.1038/nrg3226>
- Duan, H., Li, X., Mei, A., Li, P., Liu, Y., Li, X., Li, W., Wang, C., Xie, S.: The diagnostic value of metagenomic next-generation sequencing in infectious diseases. *BMC Infect. Dis.* **21**, 1–13 (2021). <https://doi.org/10.1186/s12879-020-05746-5>
- Frey, K.G., Bishop-Lilly, K.A.: Next-generation sequencing for pathogen detection and identification. *Methods Microbiol.* **42**, 525–554 (2015). <https://doi.org/10.1016/bs.mim.2015.06.004>
- Geng, S., Mei, Q., Zhu, C., Fang, X., Yang, T., Zhang, L., Fan, X., Pan, A.: Metagenomic next-generation sequencing technology for detection of pathogens in blood of critically ill patients. *Int. J. Infect. Dis.* **103**, 81–87 (2021). <https://doi.org/10.1016/j.ijid.2020.11.166>
- Grumaz, C., Hoffmann, A., Vainshtein, Y., Kopp, M., Grumaz, S., Stevens, P., Decker, S.O., Weigand, M.A., Hofer, S., Brenner, T., Sohn, K.: Rapid next-generation sequencing-based diagnostics of bacteremia in septic patients. *J. Mol. Diagn.* **22**, 405–418 (2020). <https://doi.org/10.1016/j.jmoldx.2019.12.006>

- Gu, W., Deng, X., Lee, M., Sucu, Y.D., Arevalo, S., Stryke, D., Federman, S., Gopez, A., Reyes, K., Zorn, K., Sample, H., Yu, G., Ishpuniani, G., Briggs, B., Chow, E.D., Berger, A., Wilson, M.R., Wang, C., Hsu, E., Miller, S., DeRisi, J.L., Chiu, C.Y.: Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat. Med.* **27**, 115–124 (2020). <https://doi.org/10.1038/s41591-020-1105-z>
- Guo, L., Li, Y., Liu, L., Wu, H., Zhou, J., Zhang, Y., Feng, W., Zhu, L., Hu, B., Hu, H., Chen, T., Guo, X., Chen, H., Yang, Y., Liu, G.: Detection of pediatric bacterial meningitis pathogens from cerebrospinal fluid by next-generation sequencing technology. *J. Infect.* **78**, 323–337 (2019). <https://doi.org/10.1016/j.jinf.2018.12.001>
- Gupta, S., Bhar, A., Chatterjee, M., Ghosh, A., Das, S.: Transcriptomic dissection reveals wide spread differential expression in chickpea during early time points of *Fusarium oxysporum* f. sp. ciceri Race 1 attack. *PLoS ONE* **12**, e0178164 (2017). <https://doi.org/10.1371/journal.pone.0178164>
- Gwinn, M., MacCannell, D., Armstrong, G.L.: Next-generation sequencing of infectious pathogens. *JAMA* **321**, 893–894 (2019). <https://doi.org/10.1001/jama.2018.21669>
- Han, D., Li, Z., Li, R., Tan, P., Zhang, R., Li, J.: mNGS in clinical microbiology laboratories: on the road to maturity. *Crit. Rev. Microbiol.* **45**, 668–685 (2019). <https://doi.org/10.1080/1040841X.2019.1681933>
- Hong, N.T.T., Anh, N.T., Mai, N.T.H., Nghia, H.D.T., Nhu, L.N.T., Thanh, T.T., Phu, N.H., Deng, X., van Doorn, H.R., Chau, N.V.V., Delwart, E.: Performance of metagenomic next-generation sequencing for the diagnosis of viral meningoencephalitis in a resource-limited setting. *Open Forum Infect Dis.* **7**, ofaa046 (2020). <https://doi.org/10.1093/ofid/ofaa046>
- Hu, Z., Weng, X., Xu, C., Lin, Y., Cheng, C., Wei, H., Chen, W.: Metagenomic next-generation sequencing as a diagnostic tool for toxoplasmic encephalitis. *Ann. Clin. Microbiol. Antimicrob.* (2018). <https://doi.org/10.1186/s12941-018-0298-1>
- Huebinger, R.M., Liu, M.-M., Dowd, S.E., Rivera-Chavez, F.A., Boynton, J., Carey, C., Hawkins, K., Minshall, C.T., Wolf, S.E., Minei, J.P., Barber, R.C.: Examination with next-generation sequencing technology of the bacterial microbiota in bronchoalveolar lavage samples after traumatic injury. *Surg. Infect.* **14**, 275–282 (2013). <https://doi.org/10.1089/sur.2012.095>
- Janda, J.M., Abbott, S.L.: 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol.* **45**, 2761–2764 (2007). <https://doi.org/10.1128/JCM.01228-07>
- John, G., Sahajpal, N.S., Mondal, A.K., Ananth, S., Williams, C., Chaubey, A., Rojjani, A.M., Kolhe, R.: Next-generation sequencing (NGS) in COVID-19: a tool for SARS-CoV-2 diagnosis, monitoring new strains and phylogenetic modeling in molecular epidemiology. *Curr. Issues Mol. Biol.* **43**, 845–867 (2021). <https://doi.org/10.3390/cimb43020061>
- Kidd, S.E., Chen, S.C.A., Meyer, W., Halliday, C.L.: A new age in molecular diagnostics for invasive fungal disease: are we ready? *Front. Microbiol.* (2020). <https://doi.org/10.3389/fmicb.2019.02903>
- Kim, Y., Gu, C., Kim, H.U., Lee, S.Y.: Current status of pan-genome analysis for pathogenic bacteria. *Curr. Opin. Biotechnol.* **63**, 54–62 (2020). <https://doi.org/10.1016/j.copbio.2019.12.001>
- Kommedal, Ø., Karlsen, B., Sæbø, Ø.: Analysis of mixed sequencing chromatograms and its application in direct 16S rRNA gene sequencing of polymicrobial samples. *J. Clin. Microbiol.* **46**, 3766–3771 (2008). <https://doi.org/10.1128/JCM.00213-08>
- Lecuit, M., Eloit, M.: The diagnosis of infectious diseases by whole genome next generation sequencing: a new era is opening. *Front. Cell Infect. Microbiol.* **4**, 25 (2014). <https://doi.org/10.3389/fcimb.2014.00025>
- Leekitcharoenphon, P., Nielsen, E.M., Kaas, R.S., Lund, O., Aarestrup, F.M.: Evaluation of whole genome sequencing for outbreak detection of *Salmonella enterica*. *PLoS ONE* **9**, e87991 (2014). <https://doi.org/10.1371/journal.pone.0087991>
- Lefterova, M.I., Suarez, C.J., Banaei, N., Pinsky, B.A.: Next-generation sequencing for infectious disease diagnosis and management: a report of the Association for Molecular Pathology. *J Mol. Diagn.* **17**, 623–634 (2015). <https://doi.org/10.1016/j.jmoldx.2015.07.004>
- Li, Y., Sun, B., Tang, X., Liu, Y.L., He, H.Y., Li, X.Y., Wang, R., Guo, F., Tong, Z.H.: Application of metagenomic next-generation sequencing for bronchoalveolar lavage diagnostics in critically ill patients. *Eur. J. Clin. Microbiol. Infect. Dis.* **39**, 369–374 (2020). <https://doi.org/10.1007/s10096-019-03734-5>
- Liu, X., Chen, Y., Ouyang, H., Liu, J., Luo, X., Huang, Y., Chen, Y., Ma, J., Xia, J., Ding, L.: Tuberculosis diagnosis by metagenomic next-generation sequencing on bronchoalveolar lavage fluid: a cross-sectional analysis. *Int. J. Infect. Dis.* **104**, 50–57 (2021). <https://doi.org/10.1016/j.ijid.2020.12.063>
- Matsumura, S., Ito, Y., Morita, O., Honda, H.: Genome resequencing analysis of *Salmonella typhimurium* LT-2 strains TA98 and TA100 for the establishment of a next-generation sequencing-based mutagenicity assay. *J. Appl. Toxicol.* **37**, 1099–1263 (2017). <https://doi.org/10.1002/jat.3463>
- McInerney, J.O., McNally, A., O'Connell, M.J.: Why prokaryotes have pangenomes. *Nat. Microbiol.* **2**, 1–5 (2017). <https://doi.org/10.1038/nmicrobiol.2017.40>
- Mellmann, A., Harmsen, D., Cummings, C.A., Zentz, E.B., Leopold, S.R., Rico, A., Prior, K., Szczepanowski, R., Ji, Y., Zhang, W., McLaughlin, S.F., Henkhaus, J.K., Leopold, B., Bielaszewska, M., Prager, R., Brzoska, P.M., Moore, R.L., Guenther, S., Rothberg, J.M., Karch, H.: Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104: H4 outbreak by rapid next generation sequencing technology. *PLoS ONE* **6**, e22751 (2011). <https://doi.org/10.1371/journal.pone.0022751>
- Mosele, F., Remon, J., Mateo, J., Westphalen, C.B., Barlesi, F., Lolkema, M.P., Normanno, N., Scarpa, A., Robson, M., Meric-Bernstam, F., Wagle, N.: Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: a report from the ESMO Precision Medicine Working Group. *Ann Oncol.* **31**, 1491–1505 (2020). <https://doi.org/10.1016/j.annonc.2020.07.014>
- Naccache, S.N., Federman, S., Veeraghavan, N., Zaharia, M., Lee, D., Samayoa, E., Bouquet, J., Greninger, A.L., Luk, K.-C., Enge, B., Wadford, D.A., Messenger, S.L., Genrich, G.L., Pellegrino, K., Gard, G., Leroy, E., Schneider, B.S., Fair, J.N., Martínez, M.A., Isa, P., Crump, J.A., DeRisi, J.L., Sittler, T., Hackett, J., Jr., Miller, S., Chiu, C.Y.: A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res.* **24**, 1180–1192 (2014). <https://doi.org/10.1101/gr.171934.113>
- Naccache, S.N., Peggs, K.S., Mattes, F.M., Phadke, R., Garson, J.A., Grant, P., Samayoa, E., Federman, S., Miller, S., Lunn, M.P., Gant, V., Chiu, C.Y.: Diagnosis of neuroinvasive astrovirus infection in an immunocompromised adult with encephalitis by unbiased next-generation sequencing. *Clin Infect Dis.* (2015). <https://doi.org/10.1093/cid/ciu912>
- Radford, A.D., Chapman, D., Dixon, L., Chantrey, J., Darby, A.C., Hall, N.: Application of next-generation sequencing technologies in virology. *J. Gen. Virol.* **93**, 1853–1868 (2012). <https://doi.org/10.1099/vir.0.043182-0>
- Sabapathypillai, S.L., James, H.R., Lyerla, R.R.L., Hassman, L.: The next generation of ocular pathogen detection. *Asia-Pac. J. Ophthalmol.* **10**, 109–113 (2021). <https://doi.org/10.1097/APO.0000000000000366>
- Sam, S.S., Rogers, R., Gillani, F.S., Tsongalis, G.J., Kraft, C.S., Caliando, A.M.: Evaluation of a next-generation sequencing



- metagenomics assay to detect and quantify DNA viruses in plasma from transplant recipients. *J. Mol. Diagn.* **23**, 719–731 (2021). <https://doi.org/10.1016/j.jmoldx.2021.02.008>
- Schuster, S.C.: Next-generation sequencing transforms today's biology. *Nat. Methods.* **5**, 16 (2007). <https://doi.org/10.1038/nmeth1156>
- Serwecińska, L.: Antimicrobials and antibiotic-resistant bacteria: a risk to the environment and to public health. *Water* **12**, 3313 (2020). <https://doi.org/10.3390/w12123313>
- Socransky, S.S., Haffajee, A.D.: The bacterial etiology of destructive periodontal disease: current concepts. *J. Periodontol.* **63**, 322–331 (1992). <https://doi.org/10.1902/jop.1992.63.4s.322>
- Stafford, I.A., Centeno, F.H., Al Mohajer, M., Parkerson, G., Woc-Colburn, L., Burgos-Lee, A.J., Rac, M., Dunn, J., Muldrew, K.: Successful detection of unrecognized rickettsia typhi in pregnancy using cell-free next-generation sequencing. *Case Rep. Obstetr. Gynecol.* (2020). <https://doi.org/10.1155/2020/6767351>
- Sun, J., Yang, L., Lu, Y., Wang, H., Peng, X., Dong, X., Cheng, G., Cao, Y., Wu, B., Wang, X., Zhou, W.: Screening for primary immunodeficiency diseases by next-generation sequencing in early life. *Clin. Transl. Immunol.* **9**, e1138 (2020). <https://doi.org/10.1002/cti2.1138>
- Tang, Y.W., Procop, G.W., Persing, D.H.: Molecular diagnostics of infectious diseases. *Clin. Chem.* **43**, 2021–2038 (1997). <https://doi.org/10.1093/clinchem/43.11.2021>
- Taylor, A.J., Lappi, V., Wolfgang, W.J., Lapierre, P., Palumbo, M.J., Medus, C., Boxrud, D.: Characterization of foodborne outbreaks of *Salmonella enterica* serovar Enteritidis with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection. *J. Clin. Microbiol.* **53**, 3334–3340 (2015). <https://doi.org/10.1128/JCM.01280-15>
- Tettelin, H., Massignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., DeBoy, R.T.: Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome.” *Proc. Natl. Acad. Sci.* **102**, 13950–13955 (2005). <https://doi.org/10.1073/pnas.0506758102>
- Ugolotti, E., Di Marco, E., Bandettini, R., Biassoni, R.: Genomic characterization of a paediatric MRSA outbreak by next-generation sequencing. *J. Hosp. Infect.* **98**, 155–160 (2017). <https://doi.org/10.1016/j.jhin.2017.08.009>
- van Boheemen, S., van Rijn, A.L., Pappas, N., Carbo, E.C., Vorderman, R.H., Sidorov, I., vant Hof, P.J., Mei, H., Claas, E.C., Kroes, A.C., de Vries, J.J.: Retrospective validation of a metagenomic sequencing protocol for combined detection of RNA and DNA viruses using respiratory samples from pediatric patients. *J. Mol. Diagn.* **22**, 196–207 (2020). <https://doi.org/10.1016/j.jmoldx.2019.10.007>
- Van Poelvoorde, L.A., Saelens, X., Thomas, I., Roosens, N.H.: Next-generation sequencing: an eye-opener for the surveillance of antiviral resistance in influenza. *Trends Biotechnol.* **38**, 360–367 (2020). <https://doi.org/10.1016/j.tibtech.2019.09.009>
- Wain, J., Keddy, K.H., Hendriksen, R.S., Rubino, S.: Using next generation sequencing to tackle non-typhoidal Salmonella infections. *J. Infect. Dev. Ctries.* **7**, 001–005 (2013). <https://doi.org/10.3855/jidc.3080>
- Wang, Z., Gerstein, M., Snyder, M.: RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009). <https://doi.org/10.1038/nrg2484>
- Wilson, M.R., Naccache, S.N., Samayoa, E., Biagtan, M., Bashir, H., Yu, G., Salamat, S.M., Somasekar, S., Federman, S., Miller, S., Sokolic, R., Garabedian, E., Candotti, F., Buckley, R.H., Reed, K.D., Meyer, T.L., Seroogy, C.M., Galloway, R., Henderson, S.L., Gern, J.E., DeRisi, J.L., Chiu, C.Y.: Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N. Engl. J. Med.* **370**, 2408–2417 (2014). <https://doi.org/10.1056/NEJMoa1401268>

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.