Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) 2024 25(6):471-484
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Review

https://doi.org/10.1631/jzus.B2300029

Clinical applications of metagenomics next-generation sequencing in infectious diseases

Ying LIU, Yongjun MA✉

Department of Clinical Laboratory, Affiliated Jinhua Hospital, Zhejiang University School of Medicine, Jinhua 321000, China

Abstract: Infectious diseases are a great threat to human health. Rapid and accurate detection of pathogens is important in the diagnosis and treatment of infectious diseases. Metagenomics next-generation sequencing (mNGS) is an unbiased and comprehensive approach for detecting all RNA and DNA in a sample. With the development of sequencing and bioinformatics technologies, mNGS is moving from research to clinical application, which opens a new avenue for pathogen detection. Numerous studies have revealed good potential for the clinical application of mNGS in infectious diseases, especially in difficult-to-detect, rare, and novel pathogens. However, there are several hurdles in the clinical application of mNGS, such as: (1) lack of universal workflow validation and quality assurance; (2) insensitivity to high-host background and low-biomass samples; and (3) lack of standardized instructions for mass data analysis and report interpretation. Therefore, a complete understanding of this new technology will help promote the clinical application of mNGS to infectious diseases. This review briefly introduces the history of next-generation sequencing, mainstream sequencing platforms, and mNGS workflow, and discusses the clinical applications of mNGS to infectious diseases and its advantages and disadvantages.

Key words: Metagenomics next-generation sequencing (mNGS); Infectious disease; Cerebrospinal fluid (CSF); Oxford Nanopore Technologies (ONT); Microbiome

1 Introduction

Infectious diseases are a major threat to human health. The World Health Organization (WHO) predicts that, by 2050, 13 million deaths will be attributed to infectious diseases (Dye, 2014). When diagnosing infectious diseases, organism detection is crucial. Conventional methods for organism detection include the culture method, immunologic assay, polymerase chain reaction (PCR) assay, and matrix-assisted laser desorption/ionization time-of-flight (MADI-TOF) mass spectrometry (MS). Each of these methods has its advantages and disadvantages (Table 1). It has been reported that about 40% of intestinal infections, nearly 50% of blood infections, and up to 50% of central nervous system infections could not be attributed to a definite pathogenic organism (Han et al., 2019). In these situations, doctors usually use broad-spectrum antibiotics, which not only increase medical expenses but also lead to multiple-resistant bacteria. Hence, a better method for the identification of pathogenic organisms is required.

Next-generation sequencing (NGS) technologies are new methods for sequencing nucleic acids, and metagenomics NGS (mNGS) is a new microbial detection technology that uses NGS methods to detect nucleic acids in samples and identity pathogens including bacteria, viruses, fungi, and parasites. Targeted sequencing and shotgun sequencing are two types of metagenomics. In the targeted sequencing approach, certain conserved regions (16S ribosomal RNA (rRNA), 18S rRNA, and internal transcribed spacer (ITS) regions) are amplified and then sequenced (Salipante et al., 2013; Wagner et al., 2018). Variable regions and conserved regions can be identified for different groups of organisms. However, the species level of organisms is nearly impossible to identify using a targeted sequencing method, and these data cannot be directly used to assess functions. The shotgun metagenomics

✉ Yongjun MA, jhmyj71@sina.cn
Ying LIU, https://orcid.org/0000-0002-5109-5252
Yongjun MA, https://orcid.org/0000-0001-9165-7371

Received Jan. 9, 2023; Revision accepted June 6, 2023;
Crosschecked Apr. 10, 2024; Published online May 17, 2024

method can unbiasedly detect all nucleic acids in samples. It can assign taxonomy, quantify organisms to the species level, and assess gene function (Quince et al., 2017). The advantages and disadvantages of the two approaches are summarized in Table 2. The development and decreased cost of NGS technologies have allowed mNGS to gradually move from the arena of scientific research to clinical application, especially for detecting novel organisms and complicated infections, such as the novel coronavirus and the new bunyavirus (Zhang W et al., 2019; Li et al., 2021). In this review, we briefly introduce the history of NGS, the routine of mNGS, and its clinical applications in infectious diseases, and discuss the technologies. Additionally, mNGS refers only to the shotgun sequencing method.

2 History of NGS

First-generation sequencing (also called Sanger sequencing) began with the invention of the chain termination and chemical degradation methods in 1977 (Sanger and Coulson, 1975; Maxam and Gilbert, 1977). In 1986, ABI Corporation produced the first DNA sequencer, which marked a commercial step towards DNA sequencing (Heather and Chain, 2016). Sanger sequencing is the gold standard method and has the

Table 1 Advantages and disadvantages of different methods for organism detection

Method	Advantages	Disadvantages	Turnaround time
Culture	Low price; Easy operation; High specificity; Gold standard	Manual operation; Low sensitivity; Hard to identify fastidious organisms	Normally 3–5 d; Up to 1–2 weeks for slow-growing organisms (e.g., fungi and tubercle bacillus)
Immunology assay	Easy operation; Relatively low price; Large number of samples in one test	Low sensitivity and specificity; Cross reaction among similar antigens; Relies on antibody quality	Several hours
PCR assay	High sensitivity; Can be quantitative; Can identify several organisms in an assay	High price; Can only identify specific organisms; Cannot distinguish between dead and live organisms	Several hours
MALDI-TOF MS	High specificity; Short time; Low price; Large samples in one test	High-priced equipment; Culture before identification	Several minutes after culture
mNGS	Unbiasedly detects all organisms (bacteria, fungi, viruses, and parasites); Discovers novel organisms; Advantageous for fastidious organisms	High price; Complicated operation and result analysis; Contamination risk; Cannot distinguish between dead and live organisms	24–72 h, average 48 h

PCR: polymerase chain reaction; MALDI-TOF MS: matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; mNGS: metagenomics next-generation sequencing.

Table 2 Advantages and disadvantages of targeted and shotgun sequencing

Approach	Advantages	Disadvantages
Targeted sequencing	Great depth; More precise; Only sequences a special region: 16S rRNA, 18S for fungi, and ITS regions	Unequal amplification; Only looks at one part of one gene; Not enough resolution for species identification; Impossible to directly assess microbial function
Shotgun sequencing	Unbiased sequencing; Greater resolution to genetic content; Can assess function; Can identify novel organisms/genes/genomic features	Insensitive to high-host background and low biomass; Mass and complex dataset; Contamination risk; High cost

rRNA: ribosomal RNA; ITS: internal transcribed spacer.

advantages of reading DNA with a length of 1000 bp with high accuracy (99.99%); however, the low throughput and high cost restricted larger-scale applications. In 2005, the 454 Life Science Corporation invented the method of sequencing by synthesis and produced the first high-throughput sequencing system. Soon afterwards, Illumina introduced the Solexa system in 2006, followed by ABI's Solid system in 2007 (Schatz et al., 2010). These three systems were second-generation sequencing technology, also called NGS. Pacific Biosciences (PacBio) led the development of third-generation sequencing (TGS) technology. In 2011, PacBio released the PacBio RS sequencer, which uses single-molecule real-time (SMRT) technology (van Dijk et al., 2018). SMRT technology is still based on sequencing by synthesis and no PCR amplification is required, as a closed and circular single-stranded DNA (ssDNA) template can be replicated automatically (Xiao and Zhou, 2020). In 2014, Oxford Nanopore Technologies (ONT) launched the first nanopore sequencer—MinION (Heather and Chain, 2016; Wang et al., 2021). This technology sequences single molecules in real-time using a nanoscale protein pore (Wang et al., 2021). In recent years, ONT released a series of platforms such as GridION (2017), PromethION (2018), and PromethION 48 (2019) (Fig. 1).

3 Representative platforms

The Illumina sequencing platform is a type of second-generation sequencing and is widely used globally. By reviewing 108 studies published before 2019 in infectious diseases, more than 50% of the studies used the Illumina platforms (Han et al., 2019). Illumina sequencing platforms perform sequencing by synthesis. By using bridge amplification, fragments are amplified into distinct and clonal clusters. Once cluster generation is completed, the templates are read as sequencing readouts. Illumina technology uses a reversible terminator-based method; fluorescent-labeled deoxynucleoside triphosphate (dNTP) is added and the emission wavelength is identified during each sequencing cycle. Illumina offers a series of platforms, such as the iSeq 100, MiniSeq, and MiSeq systems and the NexSeq 550 instrument. The characteristics and key applications of these platforms are summarized in Table 3.

Nanopore sequencing is a type of TGS that can sequence nucleic acids in real-time without amplification. When single nucleic acid molecules pass through a tiny protein channel (nanopore), they cause changes in the ion current, which can be used to identify the sequence as well as base modifications (Wang et al., 2021). Oxford introduced several platforms: MinION, GridION, and PromethION. MinION is portable so that it can sequence anywhere. Key features of the three platforms are summarized in Table 4. Nanopore sequencing technology has the following advantages: (1) short turnaround time (TAT); (2) PCR-free direct sequencing of original DNA/RNA samples; (3) no preference for GC; and (4) direct acquisition of methylation information. Due to its short TAT, portability, and real-time data analysis, nanopore sequencing has potential applications in healthcare (Sun et al., 2020), for example in the epidemiological surveillance of Ebola (Quick et al., 2016) and Zika (Faria et al., 2017) viruses. However, the disadvantages of high cost and high error rate restrict its clinical applications.

MGI Tech Co., Ltd. (a subsidiary of the Beijing Genomics Institute (BGI) Group) is a representative sequencing corporation in China, which launched a series of NGS equipment including MGISEQ-200, MGISEQ-2000, DNBSEQ-T7, and DNBSEQ-G50. MGI sequence platforms perform the combinatorial probe-anchor synthesis (cPAS) method, and the first

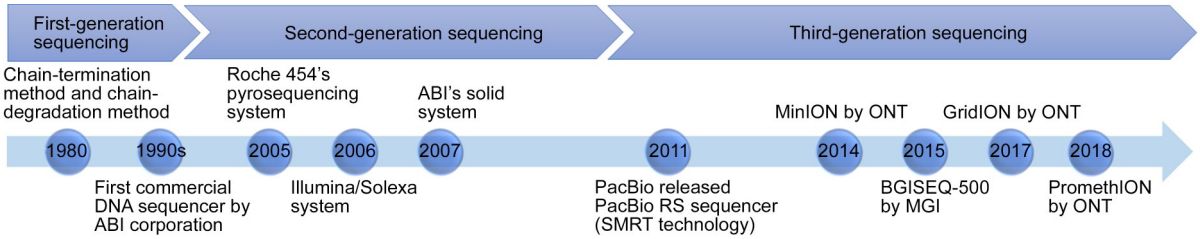

Fig. 1 Timeline of sequencing development and main platforms. SMRT: single-molecule real-time; PacBio: Pacific Biosciences; ONT: Oxford Nanopore Technologies; BGI: Beijing Genomics Institute.

Table 3 Illumina sequencing platforms*

Platform	Run time (h)	Maximum reads per run (million)	Maximum read length (bp)	Key applications
iSeq 100	9.5–19.0	4	2×150	Small whole-genome sequencing (microbe and virus); Targeted gene sequencing
MiniSeq	4–24	25	2×150	Small whole-genome sequencing (microbe and virus); Targeted gene sequencing; Targeted gene expression profiling; 16S metagenomics sequencing
MiSeq	4–55	25	2×300	Small whole-genome sequencing (microbe and virus); Targeted gene sequencing; 16S metagenomics sequencing
NextSeq 550	12–30	400	2×150	Exome sequencing; Targeted gene sequencing; Whole-transcriptome sequencing; Cytogenomic array
NextSeq 1000&2000	11–48	1200	2×150	Small whole-genome sequencing (microbe and virus); Exome and large panel sequencing; Single-cell profiling; Transcriptome sequencing; MicroRNA and small RNA analysis

* Data from the Illumina corporation website (https://www.illumina.com/systems/sequencing-platforms.html).

Table 4 Nanopore sequencing technology platforms*

Platform	Key features
MinION	Powerful: get up to 50 Gb data from a single flow cell. Theoretical max output when system is run for 72 h at 420 bases/s; Portable, sequence anywhere; Real-time: immediate data streaming for rapid, actionable results; Unrestricted read length: generate short to ultra-long (>4 Mb) reads for ultimate experimental flexibility
GridION	Flexible: run up to five independently addressable Flongle or MiniION flow cells; High throughput: as much as 250 Gb of data-streamed in real-time for immediate analysis; Integrated compute: powerful onboard data processing and analysis minimizing IT requirements; Compact: small footprint, suitable for any lab; No capital cost: pay just for consumables
PromethION	On-demand sequencing: run up to 48 independently addressable, high-capacity PromethION flow cells; Ultra-high throughput: generate terabytes of data—streamed in real-time for immediate analysis; Powerful computation: alleviate data analysis bottlenecks; Population-scale sequencing: generate human genomes from as little as $720 per genome for all consumables and instrumentation; No capital cost: pay just for consumables

* Data from the website (https://nanoporetech.com/products).

commercially available sequencer, BGISEQ-500, was announced in 2015 (Huang et al., 2017; Korostin et al., 2020). In December 2021, MGISEQ-200 and MGISEQ-2000 were approved by the National Medical Products Administration for use in DNA and RNA sequencing from human samples, an important step in promoting the development of sequencing (National Medical Products Administration, 2024a, 2024b).

The Association of Biomolecular Resource Facilities (ABRF) led a performance assessment study of DNA sequencing platforms and the results were published in *Nature Biotechnology* in 2021 (Foox et al., 2021). In the ABRF NGS study, a group of sequencing instruments was included: HiSeq/NovaSeq/paired-end 2×250-bp chemistry; Ion S5/Proton; PacBio circular consensus sequencing (CCS); ONT PromethION/MinION, BGISEQ-500/MGISEQ-2000, and GS111. The results of this study provided insights into current common sequencing instruments and are summarized in Table 5.

4 Routine of mNGS

mNGS involves the collection of samples, extraction of nucleic acids, preparation of libraries, sequencing, bioinformatics analyses, and the generation of reports. The procedures for collecting samples, extracting nucleic acids, preparing libraries, and sequencing are called wet experiments, while the procedures of bioinformatics analysis and report generation are called dry experiments (Fig. 2). The collection of clinical specimens is the first step. Various types of samples are suitable for detection by mNGS assay, such as blood, urine, bronchoalveolar lavage fluid (BALF), and tissues. After samples are collected, they should be transported quickly. Long-term storage and repeated freeze-thaw cycles can reduce the ability to detect pathogens, especially for RNA viruses.

High-quality nucleic acid is the key factor for library preparation and sequencing. Certain samples should be processed before nucleic acid extraction.

Table 5 Key results of ABRF next-generation sequencing study

Key performance	Key results
Distribution of genomic coverage	Good performance in all sequencing platforms
Sequencing mismatch rate	ONT has higher mismatch rate than short-read platforms in every context; The mismatch rate of PacBio CCS reads is the same as or even lower than that of short reads in all contexts, except for satellite regions; BGISEQ-500 performs better than HiSeq 2500, 4000, and X10, but MGISEQ-2000 falls short; GS111 has higher mismatch rate than all other short-read platforms, except for satellite regions
Single-nucleotide variant and INDEL detection	Compared with other platforms, BGISEQ-500, MGISEQ-2000, and NovaSeq 2×250 bp have the highest precision and recall rates, but HiSeq 2500 and 4000 have the worst performances
Bacterial genome sequencing	Both within and across platforms, taxonomic composition varies considerably; Each taxon's GC content clearly influences taxonomic composition

ABRF: The Association of Biomolecular Resource Facilities; ONT: Oxford Nanopore Technologies; PacBio: Pacific Biosciences; CCS: circular consensus sequencing; INDEL: insertion and deletion.

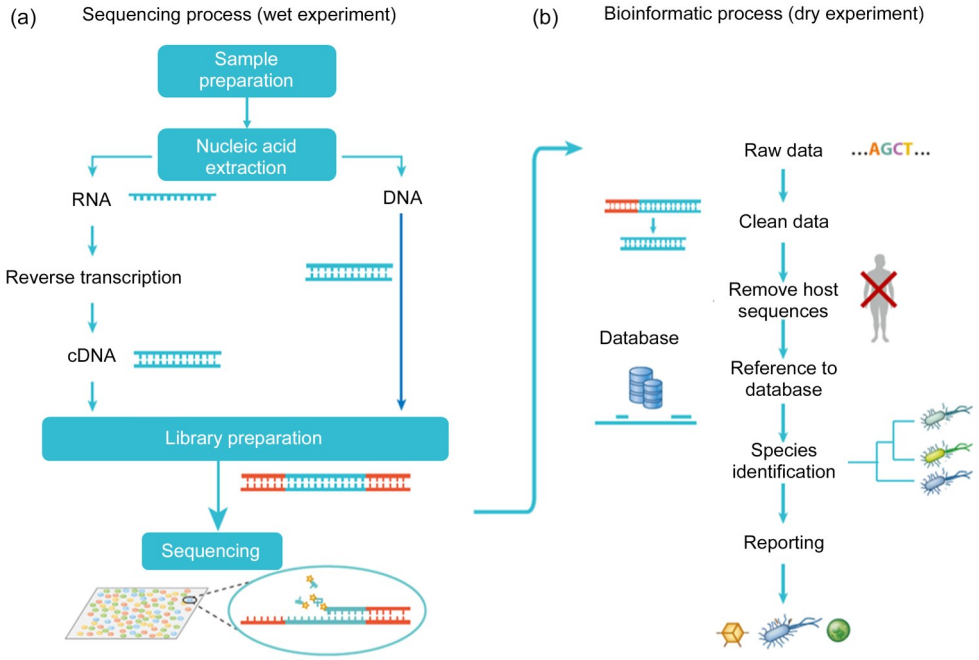

Fig. 2 Workflow of metagenomics next-generation sequencing (mNGS) includes two parts: (a) sequencing process (wet experiment), including sample preparation, nucleic acid extraction, library preparation, and sequencing; (b) bioinformatic process (dry experiment). The average turnaround time is 48 h. cDNA: complementary DNA.

For example, sputum should be liquified with dithiothreitol (DTT) and/or protease K, blood samples should be centrifuged and the plasma collected, and absolute lysis is important for the detection of fungi or mycobacteria. The amount of microbial nucleic acids can be small in clinical samples, as most of the nucleic acids is of human origin. Human nucleic acids can be removed using laboratory and bioinformatics methods. Because the human genome is large, it complicates bioinformatics analysis and can reduce the accuracy of results. If the amount of host DNA is very high (>90%), it is a better choice to remove the human nucleic acid prior to sequencing by using a laboratory method. Filtration, differential lysis, and probe capture methods might be suitable to ensure that minimal host DNA is obtained (Thomas et al., 2012). Suitable extraction methods and kits should be chosen according to the sample type and detection purpose. RNA extraction differs from DNA extraction. RNA should be reversely transcribed into complementary DNA (cDNA), which is used for the library construction. Furthermore, qualified nucleic acid can be used for the library construction. The evaluation criteria for qualified nucleic acid are: (1) the total amount of nucleic acid is ⩽3 ng (⩾30 ng in blood samples); and (2) the percentage of 200–300 bp nucleic acid fragments is ⩽90%. Different sequencing platforms have different methods for the library construction. For example, in the Illumina platforms, the entire procedure for library preparation consists of nucleic acid fragments, end repair, index adaptor ligation, and target fragment amplification. The prepared library is quantified by quantitative real-time PCR (qPCR) or other analytical instruments.

Sterility is very important during all wet experiment procedures. When small amounts of exogenous DNR/RNA are introduced during the operational process, a detectable signal will be produced and a false-positive result may occur. Exogenous nucleic acid can come from laboratory surfaces, consumables, reagents, and normal flora (Salter et al., 2014; Strong et al., 2014). Contaminated specimens not only make bioinformatics analysis complicated, but also create a false-positive report, especially for low-biomass samples such as cerebrospinal fluid (CSF). Studies have shown that an overamplification of sequencing reagent contaminants occurs in CSF samples (CSF RNA inputs typically range from 5 to 50 pg) (Ramachandran and Wilson, 2020). Quality control should be performed for each batch of experiments to control the whole procedure, usually including internal controls, negative and positive external controls, process controls, and contamination controls (Gu et al., 2019). Negative controls could be elution buffer, transport medium, or noninfectious samples, which can detect external, reagent, and cross-sample contamination. Positive controls contain one or more known microbiomes. If the known microbiomes are not detected in the positive controls, the test is a failure.

Bioinformatics analysis is also called dry experimentation. Unfortunately, there are no user-friendly software for analyzing mNGS data. This still requires highly trained professional staff to customize bioinformatics pipelines (Naccache et al., 2014; Flygare et al., 2016). The raw sequencing reads are usually performed in the FASTQ format and a typical bioinformatics pipeline consists of filtering low-quality and low-complexity reads, adaptor trimming, human host and contaminant subtraction, microorganism identification by alignment to a reference database, species identification, and report generation (Li et al., 2021). Low-quality sequences are characterized by Q30 bases of ⩽80%, up to 1% adaptor contamination, and sequence length of <50 bp. Raw data filtering provides an idea of the quality of the data and cleans up and reduces the size of the data, making downstream analysis much easier (del Fabbro et al., 2013). Human host and contaminant removal are somewhat unique to metagenomic analyses. It is necessary to consider the potential contaminants and to subsequently gather the necessary reference sequences.

Report generation is the final step of mNGS. One of the significant challenges of mNGS in clinical applications is how to interpret the results and generate a reasonable and accurate clinical report (Chiang and Dekker, 2020). mNGS technology detects microorganisms in an unbiased manner. Hence, it is important to distinguish strictly pathogenic microorganisms from opportunistic pathogenic microorganisms. Due to the difficulty of DNA extraction and the low possibility of contamination, strictly pathogenic microorganisms are considered to be present when even one read is mapped to the reference sequence, such as *Mycobacterium tuberculosis* (MTB), *Cryptococcus neoformans*, and *Legionella*. Opportunistic microorganisms may be pathogenic if present in sterile fluid

(blood, CSF, and puncture fluid), while they may be symbiotic if present in sputum, urine, or bronchoalveolar fluid. Therefore, more attention should be paid to clinical symptoms and conventional test results (X-ray, culture, immunology assay, and PCR), especially in immunodeficient or immunocompromised patients, and a suitable and correct report should be generated.

5 Applications of mNGS in infectious diseases

In 2014, *The New England Journal of Medicine* published the first clinical application of mNGS. Clinically actionable information was obtained from mNGS in a 14-year-old boy with meningoencephalitis who was diagnosed with neuroleptospirosis. The boy received targeted antibiotic treatment and eventually recovered (Wilson et al., 2014). This initial report opened up a new field for the clinical application of mNGS in the diagnosis of infectious diseases. Recently, the clinical value of mNGS in infectious diseases has been widely researched and studies have shown that mNGS is more sensitive than traditional methods (Miao et al., 2018; Duan et al., 2021).

5.1 Nervous system infections

Various pathogenic microbes can invade the central nerval system (CNS), presenting as encephalitis, meningitis, or abscess. CNS infections are emergency diseases with serious symptoms, rapid progression, and high mortality. However, up to 50% of CNS infections are undiagnosed by conventional tests (culture, immunologic assay, and PCR assay) (Granerod et al., 2010). Diagnostic methods that are timely and accurate are therefore urgently needed. CSF and brain tissue can be used for microorganism detection via mNGS. Miller et al. (2019) collected 95 samples of CSF and validated the test performance of mNGS. The results showed that the sensitivity and specificity of mNGS was 73% and 99%, respectively. In a large, prospective multicenter study of 213 patients with infectious and noninfectious CNS diseases, the mNGS-positive detective rates for definite CNS infection and tuberculous meningitis were 57.0% and 27.3%, respectively, and the sensitivities for cryptococcal meningitis and cerebral aspergillosis were 76.9% and 80.0%, respectively (Xing et al., 2020). In another multicenter study, mNGS of CSF from 204 patients with meningitis or encephalitis was found to improve the diagnosis of neurologic infections and provided some useful information in some cases (Wilson et al., 2019). However, in this study, researchers also indicated that mNGS did not detect all infections, while mNGS had good concordance with other CSF methods. In general, mNGS has a higher detection rate than conventional methods, especially for simultaneous multiple pathogens (Piantadosi et al., 2021), but it cannot replace the use of traditional methods such as culture (Xing et al., 2020; Zhang et al., 2020). The use of mNGS with other tests can increase accuracy, especially for multiple infections and complicated and rare diseases.

5.2 Respiratory system infection

Pathogenic microorganisms can invade the respiratory system, especially the lungs, presenting mainly as pneumonia. Typical symptoms are fever, chest pain, cough, and difficulty in breathing. Microorganisms are difficult to identify, although medical history, physical examination, and chest X-ray can help doctors diagnose the cause of the infection. Pulmonary infections are urgent and have a high mortality, with more complicated situations in old and immunocompromised patients (Azoulay et al., 2020; Peng et al., 2021). In China, nearly half of the patients with pulmonary infections cannot be diagnosed with a definite pathogenic bacterium (Zhu et al., 2018). An mNGS assay can provide rapid and precise detection for pulmonary infections. It is customary to collect specimens of BALF, sputum, throat swab, and tissue for mNGS. One study evaluated the value of using mNGS for 240 patients with peripheral pulmonary infection. The positive rate of mNGS was 89% compared to the 25.73% rate of traditional methods (Huang et al., 2020). In another study, the detection rate of mNGS was 65% in patients with lower respiratory tract infection (LRTI), where that of the culture method was only 20% (Chen et al., 2020). In a single-center, prospective cohort study, rapid metagenomics achieved 96.6% sensitivity and 88.0% specificity when testing 171 BALF and 121 sputum samples from patients with LRTI (Mu et al., 2021). Data reflected that the detection rate of mNGS was higher than that of conventional cultures.

For detecting respiratory viruses, multiple studies have suggested that the sensitivity and specificity

of mNGS were lower than those of the PCR method (Prachayangprecha et al., 2014; Thorburn et al., 2015; Qian et al., 2020). Qian et al. (2020) found 50% sensitivity and 100% specificity of mNGS with FilmArray Respiratory Panel (FA-RP) as referent and a lower mNGS sensitivity compared to PCR in the detection of certain viruses, such as influenza A and rhinovirus; however, mNGS had the advantage in detecting rare and novel viruses.

For *Mycobacterium* detection, the diagnostic accuracy of mNGS for BALF samples and lung biopsy tissue specimens did not differ significantly (Zhu et al., 2021), but there were controversies concerning the sensitivity. In the study by Zhu et al. (2021), the results showed that the sensitivity of mNGS for MTB was superior to that of conventional culture methods and Xpert. Zhou et al. (2019) reported that the sensitivity of mNGS to MTB was similar to that of Xpert and higher than that of conventional culture methods. Another study showed a lower sensitivity to MTB detection with the Illumina and Nanopore platforms (Guo et al., 2021). The reasons for the inconsistent reports could be that absolute lysis before nucleic acid extraction is essential due to the thick *Mycobacterium* cytoderm, the low read ratio of MTB, and existent bias in different studies. Liu X et al. (2021) suggested that combining mNGS with conventional detection methods (Xpert, culture, and acid-fast bacillus smear using microscopy) could increase the detection rate of MTB.

For pulmonary fungal infections, there were also controversies relating to the sensitivity. In the study by Yang et al. (2021), mNGS had a significantly higher sensitivity than conventional tests in diagnosing pulmonary fungal infections (80.0% vs. 44.4%). However, in the study by Fang et al. (2020), no significant differences were noted between the mNGS and conventional test methods. Both studies were single-center and small-sample-size analyses, so further research should be conducted. mNGS can detect fungi that are difficult to test for using traditional methods; however, mNGS has no advantage for cryptococcal or *Candida* species detection (Qian et al., 2020; Zheng et al., 2021).

5.3 Bloodstream infections

Bloodstream infections triggering sepsis can be caused by a wide variety of different microorganisms. Sepsis is a clinical syndrome that can rapidly lead to tissue damage, organ failure, and even death. Sepsis is challenging to diagnose and treat. In the sepsis guidelines, a blood culture is the only microbiology test recommended to support the diagnosis of sepsis (Rhodes et al., 2017). However, the sensitivity of blood culture is only 5%–50% and the TAT is long (incubation time of 5–7 d) (Lamy et al., 2016). Hence, clinicians need a rapid and precise detection method to determine the pathogenic microorganism in patients with bloodstream infections for correct diagnosis and treatment.

mNGS assay can detect cell-free DNA/RNA (cfDNA/RNA) extracted from plasma to determine the pathogenic microorganism causing sepsis. A study showed that the sensitivity and specificity of mNGS for the diagnosis of bloodstream infections were 87.1% and 80.2%, respectively, compared with the composite reference standard (Jing et al., 2021). Another study showed 93.7% agreement between mNGS and blood culture results in a cohort of 350 patients with a sepsis alert, and that more than 85% of the mNGS results were delivered the day after the sample was received, with 53.7% of the reports identifying one or more microorganisms (Blauwkamp et al., 2019). The results from this study showed that mNGS assay is not only accurate and sensitive compared to blood culture but also time-effective in the diagnosis of sepsis. In critically ill patients who could not be diagnosed by blood culture, mNGS showed higher pathogen detection rates (Geng et al., 2021). This may be because DNA from pathogens could remain longer in plasma. The advantages of detecting cfDNA are that it is a non-invasive approach and has higher sensitivity due to less human cellular background. Hong et al. (2018) reported that detecting cfDNA in the peripheral blood of patients with invasive fungal infections could provide a rapid and actionable treatment when a biopsy is not possible.

Bacterial DNA has been detected in the blood of healthy people and its taxonomic composition differs from that of septic patients (Gosiewski et al., 2017). It was suggested that blood is not a sterile environment and that bacteria continuously move through the blood without causing sepsis. Hence, clinical physicians should consider combining cfDNA analysis with clinical information to distinguish between pathogenic and opportunistic organisms.

5.4 Ocular infections

In the case of ocular infections, the traditional method is to take a corneal scrape and use it for culture or PCR assay. Sadly, more than 50% of all presumed ocular infections do not have an identified pathogen (Gallon et al., 2019). mNGS opened the way to explore pathogens in ocular infections, especially rare organisms. For example, a Chinese woman in New York was diagnosed with Malayan filariasis infection using mNGS assay with the patient's eye discharge sample and subcutaneous tissue sample (Gao et al., 2016). Due to the difficulty in obtaining large amounts of samples, the research has mainly involved small sample retrospective studies (Parekh et al., 2020; Lalitha et al., 2021; Low et al., 2022). It is therefore still unclear whether mNGS could be a useful diagnostic tool for ocular infections. In addition, little is known about the ocular microbiome or its function. Further research should be conducted on ocular infections.

6 Advantages of mNGS in clinical applications

The greatest advantage of mNGS is that it is an unbiased method for the detection of all nucleic acids in samples without culture. It can not only identify pathogenic microorganisms but also discover novel pathogenic microorganisms, especially in complicated diseases and those with multiple infectious agents. With the development of NGS technology, mNGS now has the advantage of timeliness compared to the culture method, which usually takes two to three days to obtain results and even more than one week in cases of fastidious bacteria, such as MTB (Mu et al., 2021). The average TAT of routine mNGS is 48 h (Han et al., 2019; Gu et al., 2021). A TAT of 6 h was reported for detecting pathogenic microorganisms using mNGS technology with the Nanopore platform (Gu et al., 2021; Mu et al., 2021). Furthermore, mNGS is less affected by the use of antibiotics, and more information can be obtained with one test, such as drug resistance and evolution characteristics (Zhang XX et al., 2019).

7 Limitations of mNGS in clinical applications

Due to a number of factors, the clinical application of mNGS has lagged behind research despite its potential and recent successes. As the first limitation, mNGS is less sensitive when host background is high. For example, 97%–99% of the sequences might be human in an infected CSF sample from a patient with encephalitis (Wilson et al., 2018). Various host depletion methods have been reported (Marotz et al., 2018; Nelson et al., 2019); for example, the reprocessing of clinical samples with the nonionic surfactant saponin resulted in a reduction of background human DNA and improved the sensitivity of NGS to detecting pathogens (Hasan et al., 2016). Differential lysis could reduce human nucleic acid; however, it also depletes viral and certain bacterial nucleic acids (Nelson et al., 2019). Thus, before applying these methods, they should be validated to ensure that the method will not deplete the pathogens. Until now, there has been a lack of systematic research on the influence of the host background on the results of mNGS. Recent research showed that spiked phages used as internal controls were reliable indicators of sensitivity loss, and the exclusion of samples with high-host background could increase the sensitivity from 73% to 89% (Miller et al., 2019). Increasing the sequencing depths is another approach to reducing the false-negative rate, but sequencing costs also increase and cost-effectiveness should be considered in clinical applications. A read depth of 20 million is generally sufficient and cost-efficient for pathogen detection (Liu DL et al., 2021). Other diagnostic tests that are less sensitive to the host background should be considered, such as 16S rRNA gene PCR assay, which can also test positivity (Flurin et al., 2022).

The second limitation is that clinical mNGS lacks universal reference standards, test validation, and quality assurance, since it is a new technology and differs from conventional methods. Clinical conventional methods are mature and relatively easy to operate by following standardized protocols in a licensed clinical laboratory. In clinical laboratories, any changes need to be validated to ensure accurate and reproducible results. Many efforts have been made in the areas of validation of laboratory protocols, data analysis algorithms, reference databases, and the establishment of vigorous quality control metrics. Miller et al. (2019) developed and validated a clinical mNGS assay for the diagnosis of infectious causes of meningitis and encephalitis from CSF in a licensed microbiology laboratory. Their results demonstrated the analytic performance

of a laboratory-validated mNGS assay for pathogen detection, to be used clinically for the diagnosis of neurological infections from CSF samples. It is believed that more validated clinical mNGS assays will be used in the diagnosis of infectious diseases in licensed microbiology laboratories.

Third, the greatest challenge of mNGS is how to make a definite report using metagenomics data. Different from conventional methods, mNGS is a technology that detects unbiased nucleic acid in samples and then mass data are obtained through one sequencing, thus making result analysis more complicated. The first step is to determine whether an organism is detected. A normalized bioinformatics parameter, for example stringent mapped read number (SMRN), is used to describe the differences between organisms that are clinically relevant and those that are not. According to varying situations, each laboratory should establish its own technical threshold. The next step is to determine whether an organism should be reported and it is important to determine the role of the organisms (infection, colonization, and contamination). In practice, it is almost impossible to establish universal criteria to determine whether an organism causes infections. mNGS results cannot be used for clinical diagnosis. Whether the results of mNGS are positive or negative, professional physicians should make the correct clinical decision by combining other microbiological tests (culture, serology, and PCR assay), clinical manifestations, epidemiological history, and other related factors. Hence, to avoid misinterpretation of mNGS results, report teams should include clinical microbiologists, infectious disease specialists, and treating physicians.

The fourth limitation concerns cost. Although costs decrease with the development of NGS technology, the overall per-sample cost is relatively high. For example, the accepted amount of sequencing data for each sample of human feces is 6–9 GB and the corresponding price for library construction and sequencing ranges from $100 to $300 (Liu YX et al., 2021). The estimated cost for reagents per sample (excluding labor) was $27.20–$61.40 and $269.70 for Illumina and Nanopore sequencing, respectively (Gu et al., 2021). If labor cost is considered, the cost per sample increases. Although mNGS is more expensive compared to traditional methods, Bajaj et al. (2020) reported that mNGS was cost-effective for predicting and potentially preventing hospitalizations in cirrhosis. It is anticipated that, with the advancement of the technology, mNGS will be able to reduce the cost of managing infectious patients.

The advantages and disadvantages of mNGS in clinical applications are summarized in Table 6.

8 Future directions

Although mNGS has disadvantages, it is on the road to maturity and it is gradually being recognized. With the localization of the technology, it could be applied to a wide range of clinical applications: (1) obtaining information on virulence and drug resistance; (2) tracking pathogen outbreaks and epidemiological investigation; and (3) identifying pathogen colonization prior to disease onset. We believe that, in the future, with the development of sequencing technology, the cost and TAT of mNGS will continue to decrease, the experimental process of mNGS will be easy to operate, and all procedures will be designed to be operated with a push of a button; using cloud-based platforms, the whole procedure will be automated and bioinformatics experts will not be required (Chiu and Miller, 2019).

Table 6 Advantages and disadvantages of mNGS

Advantages	Disadvantages
Unbiased tests, including bacteria, fungi, viruses, and parasites;	High price;
Can identify rare and novel pathogens;	Can be insensitive to high-host background and low biomass;
Short turnaround time: average 48 h;	Complicated operations and lack of universal reference standards, test validation, and quality assurance;
Minimally influenced by antibiotics;	Difficult to discriminate pathogens from colonizers or contaminants
More clinical applications in microorganism identification, antibiotic/antiviral resistance prediction, microbiome analyses, and transcriptomics	

mNGS: metagenomic next-generation sequencing.

9 Conclusions

mNGS is a revolutionary technology that is different from conventional tests. A single test can identity all pathogens, including bacteria, fungi, viruses, and parasites. mNGS is especially helpful for difficult-to-test, rare, and novel pathogens. A full understanding of the advantages and disadvantages of mNGS aids its clinical application. Currently, there is a general consensus that: (1) mNGS is unlikely to replace traditional tests in the short term but it is a good auxiliary and important method for detecting microorganisms; (2) the results of mNGS cannot be used alone for diagnosis; (3) a large-scale prospective study with a large sample size is needed to advance research; and (4) validation of the mNGS workflow and a friendly bioinformatics pipeline also need further research in clinical applications.

Acknowledgments

This work was supported by the Medicine and Health, Science and Technology Plan Project of Zhejiang (Nos. 2020KY1009 and 2021KY387) and the Jinhua Science and Technology Planning Project Social Development Key Project (No. 2021-3-072), China.

Author contributions

Ying LIU wrote the manuscript. Yongjun MA had the idea, revised the manuscript, and applied the project. Both authors have read and approved the final manuscript, and therefore, agreed to be responsible for the research integrity and for all aspects of the work.

Compliance with ethics guidelines

Ying LIU and Yongjun MA declare that they have no conflict of interest.

This review does not contain any studies with human or animal subjects performed by either of the authors.

References

Azoulay E, Russell L, van de Louw A, et al., 2020. Diagnosis of severe respiratory infections in immunocompromised patients. *Intensive Care Med*, 46(2):298-314. https://doi.org/10.1007/s00134-019-05906-5

Bajaj JS, Acharya C, Sikaroodi M, et al., 2020. Cost-effectiveness of integrating gut microbiota analysis into hospitalisation prediction in cirrhosis. *GastroHep*, 2(2):79-86. https://doi.org/10.1002/ygh2.390

Blauwkamp TA, Thair S, Rosen MJ, et al., 2019. Analytical and clinical validation of a microbial cell-free DNA sequencing test for infectious disease. *Nat Microbiol*, 4(4):663-674. https://doi.org/10.1038/s41564-018-0349-6

Chen HB, Yin YY, Gao H, et al., 2020. Clinical utility of in-house metagenomic next-generation sequencing for the diagnosis of lower respiratory tract infections and analysis of the host immune response. *Clin Infect Dis*, 71(Suppl 4):S416-S426. https://doi.org/10.1093/cid/ciaa1516

Chiang AD, Dekker JP, 2020. From the pipeline to the bedside: advances and challenges in clinical metagenomics. *J Infect Dis*, 221(Suppl 3):S331-S340. https://doi.org/10.1093/infdis/jiz151

Chiu CY, Miller SA, 2019. Clinical metagenomics. *Nat Rev Genet*, 20(6):341-355. https://doi.org/10.1038/s41576-019-0113-7

del Fabbro C, Scalabrin S, Morgante M, et al., 2013. An extensive evaluation of read trimming effects on Illumina NGS data analysis. *PLoS ONE*, 8(12):e85024. https://doi.org/10.1371/journal.pone.0085024

Duan HX, Li X, Mei AH, et al., 2021. The diagnostic value of metagenomic next-generation sequencing in infectious diseases. *BMC Infect Dis*, 21:62. https://doi.org/10.1186/s12879-020-05746-5

Dye C, 2014. After 2015: infectious diseases in a new era of health and development. *Philos Trans R Soc B Biol Sci*, 369(1645):20130426. https://doi.org/10.1098/rstb.2013.0426

Fang XW, Mei Q, Fan XQ, et al., 2020. Diagnostic value of metagenomic next-generation sequencing for the detection of pathogens in bronchoalveolar lavage fluid in ventilator-associated pneumonia patients. *Front Microbiol*, 11:599756. https://doi.org/10.3389/fmicb.2020.599756

Faria NR, Quick J, Claro IM, et al., 2017. Establishment and cryptic transmission of Zika virus in Brazil and the Americas. *Nature*, 546(7658):406-410. https://doi.org/10.1038/nature22401

Flurin L, Wolf MJ, Mutchler MM, et al., 2022. Targeted metagenomic sequencing-based approach applied to 2146 tissue and body fluid samples in routine clinical practice. *Clin Infect Dis*, 75(10):1800-1808. https://doi.org/10.1093/cid/ciac247

Flygare S, Simmon K, Miller C, et al., 2016. Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. *Genome Biol*, 17:111. https://doi.org/10.1186/s13059-016-0969-1

Foox J, Tighe SW, Nicolet CM, et al., 2021. Performance assessment of DNA sequencing platforms in the ABRF next-generation sequencing study. *Nat Biotechnol*, 39(9):1129-1140. https://doi.org/10.1038/s41587-021-01049-5

Gallon P, Parekh M, Ferrari S, et al., 2019. Metagenomics in ophthalmology: hypothesis or real prospective? *Biotechnol Rep (Amst)*, 23:e00355. https://doi.org/10.1016/j.btre.2019.e00355

Gao D, Yu QF, Wang GQ, et al., 2016. Diagnosis of a malayan filariasis case using a shotgun diagnostic metagenomics assay. *Parasit Vectors*, 9:86.

https://doi.org/10.1186/s13071-016-1363-2

Geng SK, Mei Q, Zhu CY, et al., 2021. Metagenomic next-generation sequencing technology for detection of pathogens in blood of critically ill patients. *Int J Infect Dis*, 103:81-87. https://doi.org/10.1016/j.ijid.2020.11.166

Gosiewski T, Ludwig-Galezowska AH, Huminska K, et al., 2017. Comprehensive detection and identification of bacterial DNA in the blood of patients with sepsis and healthy volunteers using next-generation sequencing method—the observation of DNAemia. *Eur J Clin Microbiol Infect Dis*, 36(2):329-336. https://doi.org/10.1007/s10096-016-2805-7

Granerod J, Ambrose HE, Davies NWS, et al., 2010. Causes of encephalitis and differences in their clinical presentations in England: a multicentre, population-based prospective study. *Lancet Infect Dis*, 10(12):835-844. https://doi.org/10.1016/s1473-3099(10)70222-x

Gu W, Miller S, Chiu CY, 2019. Clinical metagenomic next-generation sequencing for pathogen detection. *Annu Rev Pathol*, 14:319-338. https://doi.org/10.1146/annurev-pathmechdis-012418-012751

Gu W, Deng XD, Lee M, et al., 2021. Rapid pathogen detection by metagenomic next-generation sequencing of infected body fluids. *Nat Med*, 27:115-124. https://doi.org/10.1038/s41591-020-1105-z

Guo YF, Li HN, Chen HB, et al., 2021. Metagenomic next-generation sequencing to identify pathogens and cancer in lung biopsy tissue. *eBioMedicine*, 73:103639. https://doi.org/10.1016/j.ebiom.2021.103639

Hasan MR, Rawat A, Tang P, et al., 2016. Depletion of human DNA in spiked clinical specimens for improvement of sensitivity of pathogen detection by next-generation sequencing. *J Clin Microbiol*, 54(4):919-927. https://doi.org/10.1128/JCM.03050-15

Han DS, Li ZY, Li R, et al., 2019. mNGS in clinical microbiology laboratories: on the road to maturity. *Crit Rev Microbiol*, 45(5-6):668-685. https://doi.org/10.1080/1040841X.2019.1681933

Heather JM, Chain B, 2016. The sequence of sequencers: the history of sequencing DNA. *Genomics*, 107(1):1-8. https://doi.org/10.1016/j.ygeno.2015.11.003

Hong DK, Blauwkamp TA, Kertesz M, et al., 2018. Liquid biopsy for infectious diseases: sequencing of cell-free plasma to detect pathogen DNA in patients with invasive fungal disease. *Diagn Microbiol Infect Dis*, 92(3):210-213. https://doi.org/10.1016/j.diagmicrobio.2018.06.009

Huang J, Liang XM, Xuan YK, et al., 2017. A reference human genome dataset of the BGISEQ-500 sequencer. *Gigascience*, 6(5):gix024. https://doi.org/10.1093/gigascience/gix024

Huang J, Jiang EL, Yang DL, et al., 2020. Metagenomic next-generation sequencing versus traditional pathogen detection in the diagnosis of peripheral pulmonary infectious lesions. *Infect Drug Resist*, 13:567-576. https://doi.org/10.2147/IDR.S235182

Jing CD, Chen HB, Liang Y, et al., 2021. Clinical evaluation of an improved metagenomic next-generation sequencing test for the diagnosis of bloodstream infections. *Clin Chem*, 67(8):1133-1143. https://doi.org/10.1093/clinchem/hvab061

Korostin D, Kulemin N, Naumov V, et al., 2020. Comparative analysis of novel MGISEQ-2000 sequencing platform vs Illumina HiSeq 2500 for whole-genome sequencing. *PLoS ONE*, 15(3):e0230301. https://doi.org/10.1371/journal.pone.0230301

Lalitha P, Prajna NV, Sikha M, et al., 2021. Evaluation of metagenomic deep sequencing as a diagnostic test for infectious keratitis. *Ophthalmology*, 128(3):473-475. https://doi.org/10.1016/j.ophtha.2020.07.030

Lamy B, Dargère S, Arendrup MC, et al., 2016. How to optimize the use of blood cultures for the diagnosis of bloodstream infections? A state-of-the art. *Front Microbiol*, 7:697. https://doi.org/10.3389/fmicb.2016.00697

Li N, Cai QQ, Miao Q, et al., 2021. High-throughput metagenomics for identification of pathogens in the clinical settings. *Small Methods*, 5(1):2000792. https://doi.org/10.1002/smtd.202000792

Liu DL, Zhou HW, Xu T, et al., 2021. Multicenter assessment of shotgun metagenomics for pathogen detection. *eBioMedicine*, 74:103649. https://doi.org/10.1016/j.ebiom.2021.103649

Liu X, Chen YL, Ouyang H, et al., 2021. Tuberculosis diagnosis by metagenomic next-generation sequencing on bronchoalveolar lavage fluid: a cross-sectional analysis. *Int J Infect Dis*, 104:50-57. https://doi.org/10.1016/j.ijid.2020.12.063

Liu YX, Qin Y, Chen T, et al., 2021. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell*, 12(5):315-330. https://doi.org/10.1007/s13238-020-00724-8

Low L, Nakamichi K, Akileswaran L, et al., 2022. Deep metagenomic sequencing for endophthalmitis pathogen detection using a Nanopore platform. *Am J Ophthalmol*, 242:243-251. https://doi.org/10.1016/j.ajo.2022.05.022

Marotz CA, Sanders JG, Zuniga C, et al., 2018. Improving saliva shotgun metagenomics by chemical host DNA depletion. *Microbiome*, 6:42. https://doi.org/10.1186/s40168-018-0426-3

Maxam AM, Gilbert W, 1977. A new method for sequencing DNA. *Proc Natl Acad Sci USA*, 74(2):560-564. https://doi.org/10.1073/pnas.74.2.560

Miao Q, Ma YY, Wang QQ, et al., 2018. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. *Clin Infect Dis*, 67(suppl_2):S231-S240. https://doi.org/10.1093/cid/ciy693

Miller S, Naccache SN, Samayoa E, et al., 2019. Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. *Genome Res*, 29(5):831-842. https://doi.org/10.1101/gr.238170.118

Mu SR, Hu L, Zhang Y, et al., 2021. Prospective evaluation of a rapid clinical metagenomics test for bacterial pneumonia. *Front Cell Infect Microbiol*, 11:684965.

https://doi.org/10.3389/fcimb.2021.684965

Naccache SN, Federman S, Veeraraghavan N, et al., 2014. A cloud-compatible bioinformatics pipeline for ultrarapid pathogen identification from next-generation sequencing of clinical samples. *Genome Res*, 24(7):1180-1192. https://doi.org/10.1101/gr.171934.113

National Medical Products Administration, 2024. Citing Electronic Sources of Information. No. 20183220257. https://www.nmpa.gov.cn/datasearch/search-info.html?nmpa=aWQ9YmIwNjM4Zjk5NzBiM2ZjYWUxMmU4NjE1NGE4NjBiZDYmaXRlbUlkPWZmODA4MDgxODNjYWQ3NTAwMTgzY2I2NmZlNjkwMjg1

National Medical Products Administration, 2024. Citing Electronic Sources of Information. No. 20183220258. https://www.nmpa.gov.cn/datasearch/search-info.html?nmpa=aWQ9ZGZjNWJjYWJhMzI4YTY3MDdmOWM0MWUzODFmY2Q4NTkmaXRlbUlkPWZmODA4MDgxODNjYWQ3NTAwMTgzY2I2NmZlNjkwMjg1

Nelson MT, Pope CE, Marsh RL, et al., 2019. Human and extracellular DNA depletion for metagenomic analysis of complex clinical infection samples yields optimized viable microbiome profiles. *Cell Rep*, 26(8):2227-2240.e5. https://doi.org/10.1016/j.celrep.2019.01.091

Parekh M, Romano V, Franch A, et al., 2020. Shotgun sequencing to determine corneal infection. *Am J Ophthalmol Case Rep*, 19:100737. https://doi.org/10.1016/j.ajoc.2020.100737

Peng JM, Du B, Qin HY, et al., 2021. Metagenomic next-generation sequencing for the diagnosis of suspected pneumonia in immunocompromised patients. *J Infect*, 82(4):22-27. https://doi.org/10.1016/j.jinf.2021.01.029

Piantadosi A, Mukerji SS, Ye S, et al., 2021. Enhanced virus detection and metagenomic sequencing in patients with meningitis and encephalitis. *mBio*, 12(4):e0114321. https://doi.org/10.1128/mBio.01143-21

Prachayangprecha S, Schapendonk CME, Koopmans MP, et al., 2014. Exploring the potential of next-generation sequencing in detection of respiratory viruses. *J Clin Microbiol*, 52(10):3722-3730. https://doi.org/10.1128/jcm.01641-14

Qian YY, Wang HY, Zhou Y, et al., 2020. Improving pulmonary infection diagnosis with metagenomic next generation sequencing. *Front Cell Infect Microbiol*, 10:567615. https://doi.org/10.3389/fcimb.2020.567615

Quick J, Loman NJ, Duraffour S, et al., 2016. Real-time, portable genome sequencing for Ebola surveillance. *Nature*, 530(7589):228-232. https://doi.org/10.1038/nature16996

Quince C, Walker AW, Simpson JT, et al., 2017. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol*, 35(9):833-844. https://doi.org/10.1038/nbt.3935

Ramachandran PS, Wilson MR, 2020. Metagenomics for neurological infections – expanding our imagination. *Nat Rev Neurol*, 16(10):547-556. https://doi.org/10.1038/s41582-020-0374-y

Rhodes A, Evans LE, Alhazzani W, et al., 2017. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. *Intensive Care Med*, 43(3):304-377. https://doi.org/10.1007/s00134-017-4683-6

Salipante SJ, Sengupta DJ, Rosenthal C, et al., 2013. Rapid 16S rRNA next-generation sequencing of polymicrobial clinical samples for diagnosis of complex bacterial infections. *PLoS ONE*, 8(5):e65226. https://doi.org/10.1371/journal.pone.0065226

Salter SJ, Cox MJ, Turek EM, et al., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*, 12:87. https://doi.org/10.1186/s12915-014-0087-z

Sanger F, Coulson AR, 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol*, 94(3):441-448. https://doi.org/10.1016/0022-2836(75)90213-2

Schatz MC, Delcher AL, Salzberg SL, 2010. Assembly of large genomes using second-generation sequencing. *Genome Res*, 20(9):1165-1173. https://doi.org/10.1101/gr.101360.109

Strong MJ, Xu GR, Morici L, et al., 2014. Microbial contamination in next generation sequencing: implications for sequence-based analysis of clinical samples. *PLoS Pathog*, 10(11):e1004437. https://doi.org/10.1371/journal.ppat.1004437

Sun X, Song L, Yang WJ, et al., 2020. Nanopore sequencing and its clinical applications. *Methods Mol Biol*, 2204:13-32. https://doi.org/10.1007/978-1-0716-0904-0_2

Thomas T, Gilbert J, Meyer F, 2012. Metagenomics—a guide from sampling to data analysis. *Microb Inform Exp*, 2:3. https://doi.org/10.1186/2042-5783-2-3

Thorburn F, Bennett S, Modha S, et al., 2015. The use of next generation sequencing in the diagnosis and typing of respiratory infections. *J Clin Virol*, 69:96-100. https://doi.org/10.1016/j.jcv.2015.06.082

van Dijk EL, Jaszczyszyn Y, Naquin D, et al., 2018. The third revolution in sequencing technology. *Trends Genet*, 34(9):666-681. https://doi.org/10.1016/j.tig.2018.05.008

Wagner K, Springer B, Pires VP, et al., 2018. Molecular detection of fungal pathogens in clinical specimens by 18S rDNA high-throughput screening in comparison to ITS PCR and culture. *Sci Rep*, 8:6964. https://doi.org/10.1038/s41598-018-25129-w

Wang YH, Zhao Y, Bollas A, et al., 2021. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol*, 39(11):1348-1365. https://doi.org/10.1038/s41587-021-01108-x

Wilson MR, Naccache SN, Samayoa E, et al., 2014. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. *N Engl J Med*, 370(25):2408-2417. https://doi.org/10.1056/NEJMoa1401268

Wilson MR, O'Donovan BD, Gelfand JM, et al., 2018. Chronic meningitis investigated via metagenomic next-generation sequencing. *JAMA Neurol*, 75(8):947-955. https://doi.org/10.1001/jamaneurol.2018.0463

Wilson MR, Sample HA, Zorn KC, et al., 2019. Clinical

metagenomic sequencing for diagnosis of meningitis and encephalitis. *N Engl J Med*, 380(24):2327-2340. https://doi.org/10.1056/NEJMoa1803396

Xiao TT, Zhou WH, 2020. The third generation sequencing: the advanced approach to genetic diseases. *Transl Pediatr*, 9(2):163-173. https://doi.org/10.21037/tp.2020.03.06

Xing XW, Zhang JT, Ma YB, et al., 2020. Metagenomic next-generation sequencing for diagnosis of infectious encephalitis and meningitis: a large, prospective case series of 213 patients. *Front Cell Infect Microbiol*, 10:88. https://doi.org/10.3389/fcimb.2020.00088

Yang L, Song JX, Wang YB, et al., 2021. Metagenomic next-generation sequencing for pulmonary fungal infection diagnosis: lung biopsy versus bronchoalveolar lavage fluid. *Infect Drug Resist*, 14:4333-4359. https://doi.org/10.2147/IDR.S333818

Zhang W, Wu TF, Guo MM, et al., 2019. Characterization of a new bunyavirus and its derived small RNAs in the brown citrus aphid, *Aphis citricidus*. *Virus Genes*, 55(4): 557-561. https://doi.org/10.1007/s11262-019-01667-x

Zhang XX, Guo LY, Liu LL, et al., 2019. The diagnostic value of metagenomic next-generation sequencing for identifying *Streptococcus pneumoniae* in paediatric bacterial meningitis. *BMC Infect Dis*, 19:495. https://doi.org/10.1186/s12879-019-4132-y

Zhang Y, Cui P, Zhang HC, et al., 2020. Clinical application and evaluation of metagenomic next-generation sequencing in suspected adult central nervous system infection. *J Transl Med*, 18:199. https://doi.org/10.1186/s12967-020-02360-6

Zheng Y, Qiu XJ, Wang T, et al., 2021. The diagnostic value of metagenomic next-generation sequencing in lower respiratory tract infection. *Front Cell Infect Microbiol*, 11: 694756. https://doi.org/10.3389/fcimb.2021.694756

Zhou X, Wu HL, Ruan QL, et al., 2019. Clinical evaluation of diagnosis efficacy of active *Mycobacterium tuberculosis* complex infection via metagenomic next-generation sequencing of direct clinical samples. *Front Cell Infect Microbiol*, 9:351. https://doi.org/10.3389/fcimb.2019.00351

Zhu N, Zhou DB, Li SQ, 2021. Diagnostic accuracy of metagenomic next-generation sequencing in sputum-scarce or smear-negative cases with suspected pulmonary tuberculosis. *BioMed Res Int*, 2021:9970817. https://doi.org/10.1155/2021/9970817

Zhu YG, Tang XD, Lu YT, et al., 2018. Contemporary situation of community-acquired pneumonia in China: a systematic review. *J Transl Int Med*, 6(1):26-31. https://doi.org/10.2478/jtim-2018-0006