Genomics in Infectious Diseases

Mitesh Shetty

ABSTRACT

For the effective care, rehabilitation, and protection of patients, recognizing and characterizing microorganisms that cause infection are essential. In the diagnostic laboratory, however, not all bacterial species can be cultured successfully. Genomics and whole-genome sequencing (WGS) can greatly enhance human knowledge and understanding of infectious diseases. The ability to assess the microbial community without the need to culture the species has created the ever-growing field of metagenomics and microbiome analysis. Currently, the principal possible applications of WGS in the diagnostic microbiology laboratory for characterizing bacterial pathogen are identification, typing, detection of resistance, and virulence gene detection. In addition, next-generation sequencing (NGS) has helped understand the genome of SARS-CoV-2 early and provided insight into epidemiology, expansion of COVID-19, early and efficient production of the vaccine. The metagenomic sequencing (mNGS) microbial cell-free DNA testing for infection diagnosis is gaining traction. More specifically, for clinicians and specialists in the clinical microbiology community, paradigm shifts in understanding molecular diagnostics are required. A comprehensive clinical review should complement NGS’s diagnostic study, which (a) demonstrates clinical effectiveness, (b) guides the use, and (c) exposes possible fields of misunderstood use. Both conventional culture-based technologies and molecular diagnostics have several strengths and limitations.

Keywords: Clinical microbiology, Genomics, Infectious disease, Metagenomics, Next-generation sequencing, Whole-genome sequencing.

Pediatric Infectious Disease (2021): 10.5005/jp-journals-10081-1296

BACKGROUND

In acutely ill hospitalized patients, the etiology of suspected infections frequently remains undiagnosed, resulting in delayed or insufficient care, extended stays, readmissions, and increased mortality and morbidity. Because of cancer, genetic syndromes, or transplantation, these patients are often immunocompromised, especially if they are in tertiary care medical centers, rendering them extremely vulnerable to infections. Unrecognized infections can quickly lead to hospital outbreaks and endanger the hospital admissions of patients. Historically, infectious diseases were identified due to the use of techniques developed in the late 1800s, such as the Gram staining and culture of the causative agent but the turnaround time-to-results cycle is long (≥48 hours), and it is difficult or impossible for many pathogens to culture. Molecular testing powered by theories, such as PCR, could require multiple individual tests for particular species, but a rare pathogen could still miss. When clinical material is more difficult and involves different organisms, such as fecal samples, the Sanger sequencing method becomes troublesome. In such cases, findings from Sanger sequencing are not accurate and make identifying particular pathogens difficult or impossible. In addition, for these activities, the cost of Sanger sequencing is high and the turnaround period is long. Pulsed-field electrophoresis (PFGE) is currently the “gold standard” method of bacterial subtyping. However, when compared to the sequencing of entire genomes, its ability to distinguish between strains is limited.

In the microbial diagnostic study, a hypothesis-free diagnostic approach that has the potential to detect almost any organism will lead to a drastic paradigm shift. Recently, unbiased metagenomic sequencing (mNGS) has proved to be a very promising microbial identification technology through the detection of possible pathogens in several culturally negative patients. X-linked agammaglobulinemia was first described in 1952 and, as such, is arguably the first PID to have been described. Haemophilus influenzae, the first full genome sequence of a free-living organism, was completed in 1995. Since then, it has sequenced over 23,000 bacteria, 600 fungi, and 4,000 viruses.

TECHNIQUES

Next-generation Sequencing

The potential of next-generation sequencing (NGS) to revolutionize microbiology has been tremendous and will greatly advance knowledge and understanding of infectious diseases. Since whole-genome sequencing (WGS) represents the pinnacle of strain characterization and epidemiological analysis, conventional typing approaches, resistance gene detection and other sequence-based studies (e.g., 16S rDNA PCR) are likely to be replaced shortly. Since its introduction in 2004, the cost of NGS has been decreased by many orders of magnitude. This enables us to look at the microbes’ full DNA fingerprint with increasing speed and decreasing cost. Metagenomics, the study of the collective genome of microorganisms, has demonstrated its ability to diagnose new infectious diseases in areas where standard diagnostics have failed. This has provided valuable new insights into the host-microbe relationship.

Next-generation sequencing approaches also known as deep, high performance, or massive parallel sequencing methods, include many sequencing systems that have succeeded the conventional...
Sanger method of the dideoxyribonucleoside chain termination. Different sequencing technologies are available, which differ in their sequencing chemistry, read length, and throughput capability (Table 1).16

Types

Illumina (San Diego, CA, USA) provides a range of platforms of sequence series like iSeq, MiSeq, MiniSeq, NextSeq, HiSeq, and NovaSeq. These platforms use a bridge amplification technique, whereby single DNA molecules are first connected to a flow cell and then amplified locally into a clonal cluster, similar to how a single bacterium grows on a media plate into a colony.19 This is accompanied by sequencing by synthesis is a technique that involves synthesizing the complementary DNA one nucleotide per cycle and an optical readout of fluorescein-labeled nucleotides (A, G, T, or C).

The Ion Torrent platform is provided by Thermo Fish Scientific (Waltham, MA, USA). It clones single DNA molecules on a bead into an emulsion.20 The beads are mounted on a semiconductor chip that comprises a matrix of individual pH sensors. As the DNA clones synthesized a local change in pH shows the nucleotide sequenced. Under the names MinION, GridION, and PromethION, Oxford Nanopore Technologies (Oxford, United Kingdom) offers portable sequencers.21 This technology directs single-stranded DNA through a grid of protein nanopores that collects the DNA sequence through electrical current interruptions. However, there are currently more sequencing errors, lower throughput, and higher per-read costs of the nanopore method than other NGS platforms, which may restrict its usefulness for some applications.3

Pacific BioScience sequencing also provides a sequence for single molecules and longer read length, making them ideal for unresolved genome, transcriptome, and epigenetic research problems. The possible benefits of long sequence readings are still being explored for clinical microbiology, while long readings (>5000 bp) have helped to address some of the drawbacks of short reading data in research environments.16

Sequencing

Usually, microorganism NGS studies adopt two general strategies: targeted amplicon sequencing or WGS (Fig. 1).22,23 The first strategy uses target-specific primers for PCR-mediated amplification to enrich and selectively sequence the genomic regions of interest. This method is also used to analyze well-known genomic regions (e.g., identify known drug-resistant mutants). On the other hand, de novo sequencing is based on the non-target preparation of whole genomes. This is also achieved when the genomes or functional potential of the organism under investigation are unknown or the aim is to identify them.18

Host depletion is used in WGS where the goal is to decrease the relative proportion of human host-context sequences in mNGS data instead of exploiting recognized pathogen targets such as the 16S rRNA gene.2 The value of unbiased metagenomic sequencing is maintained by this method. Most host libraries are normally identical to human sequences with RNA or mitochondrial RNA, and depletion of those human host sequences will indirectly increase non-human microbial readings and consequently enhance the analytical sensitivity.3

The shorter the length of the sequence read, the more contigs and gaps produced during the assembly process. This indicates

| Table 1: Popular sequencing technology16 |
| --- |
| **Traditional sequencing** |
| Sanger sequencing | Still widely used for sequencing short segments of DNA (up to 1000 bp) due to the ease and accuracy of sequencing |
| | Labor, time and cost-intensive for sequencing entire genomes regularly |
| **Next-generation sequencing technologies** |
| Ion semiconductor sequencing (Life Technologies Ion Torrent) | Uses a sequencing-by-synthesis method, detecting changes in pH due to hydrogen ion release with a synthesis of complementary DNA |
| | Higher error rates, particularly homopolymers, than other platforms and poor coverage of extremely AT-rich or GC-rich regions |
| | Ion Torrent Personal Genome Machine (PGM) and newer, higher throughput Ion Proton available |
| **Illumina sequencing** |
| | Uses a sequencing-by-synthesis method, detecting a release of fluorescent labels from incorporated nucleotides to determine the sequence |
| | Current market leader with high sequence throughput, with low error rate and low sequencing cost per base |
| | Limitations of short-read sequences and a longer sequencing run time |
| | Several platforms with moderate (MiSeq), moderate-high (NextSeq), and high (HiSeq) throughput |
| | TruSeq long read technology recently introduced to produce synthetic reads of 10 kb in length (currently only HiSeq 2000/2500) |
| **Emerging technologies** |
| Nanopore sequencing (Oxford Nanopore) | Probably the leader of the pack of benchtop sequencing technologies in development |
| | Detects characteristic disruptions in a current applied across a protein channel or “nanopore” as each nucleotide of a strand of DNA is passed through the nanopore |
| | The method still being refined, but has the capability of generating long-sequence reads |
| | Two portable/affordable benchtop sequencers available – the MinION (disposable USB stick), and the GridION (rack-mountable) |
there is an increased preference for longer read lengths. While microbial genomes are usually smaller and less complex than human genomes, techniques of long-read sequencing are useful for constructing complete, highly precise genomes and sorting out plasmids, repeats, and other complex regions. To achieve a full sequence without gaps, short and long reads are also used to combine the strengths of these different sequence approaches (Fig. 2).24 However, the platforms for short reads are cheaper and can generate large quantities of knowledge. Thus, most laboratories in clinical and public health microbiology use short-read sequencing platforms for routine sequencing.

Some technologies can sequence and identify a handful of bacterial genomes in a few hours while others have the capacity to sequence and identify 50 to 100 bacterial genomes in a single run between 1 day and 3 days.16

Data Analyses
Data processing is the greatest obstacle related to the implementation of NGS in the clinical microbiology laboratory. However, scientific knowledge of the genomic characteristics and biological history of the microorganism under investigation is needed for further study.7 Well-curated and up-to-date libraries with references are important as microbial pathogens are rapidly emerging and bacteria can share plasmids that also code virulence and antimicrobial resistance traits in various organisms.25 Many open sources and private mNGS software packages are now in existence, including serial super-fast pathogen identification (SURPI),26 Kraken,27 Taxonomer,28 and private pipelines for the detection and characterization of microbial sequences.3,29 Usually, these bioinformatics pipelines (a) pre-process sequencing reads to remove sequenced adapters and low-quality and low-complexity regions;30 (b) align to the human genome to remove human reads; (c) align the processed reads to a curated pathogen database and assign a taxonomic classification to each sequence read; and (d) conduct organizational and statistical analysis with visualization in a graphical user interface.3

Application
Next-generation sequencing’s prospective clinical applications are vast and require clinical microbiology at all levels of the phase of growth.17

Virus Identification
Previous NGS studies have demonstrated evidence-of-concept applications for bacterial identification.18 Next-generation sequencing can be the perfect one-step tool for investigating a wide range of pathogen features.3,31 Connaissance of a pathogen’s virulence profile is essential to predict the severity, the outcome of the infection, and to allow for an early start of the disease risk.
Genomics in Infectious Diseases

Drug Resistance Mutation Testing

Drug resistance poses a challenge in the fight against various viral infections. When present at <15–20% of the viral population, the Sanger method has limited sensitivity for minor variants, while NGS methods can detect approximately 1% of drug-resistant mutations (DRMs). An epidemiological study of AIDS patients has proven that the presence of DRMs in the HIV-1 can predict the success or failure of available anti-HIV treatment. Next-generation sequencing and Sanger sequencing methods were contrasted by several analyzes to collect minority resistant variants, which revealed that Sanger is lacking at least half of the DRMs found by NGS. It has been shown that the existence of such variants indicates an increased risk of failure of therapy. Cytomegalovirus (CMV) drug resistance rates are varied according to patient populations: 5–12.5% in recipients of solid organ transplantation and 2–5% in recipients of hematopoietic stem cell transplantation. CMV drug resistance identification is important in due course because the DRMs will accumulate with continuing drug exposure that could lead to reduced survival of the graft and increased morbidity.

A WGS technique could also be used to detect antimicrobial resistance in microbiology. Some studies tried to verify the precision of WGS with a fair consistency in the prediction of antimicrobial resistance. Although current organism culture susceptibility methods may be faster and more accurate for routine study, WGS methods may be useful for specimens that develop slowly or cannot be cultivated or where phenotypic susceptibility tests are inaccurate. For example, WGS was used to quickly diagnose a case of extensively drug-resistant tuberculosis, reducing the amount of time needed to diagnose the disease and ultimately reducing the likelihood of de novo resistance.

Clinical Microbiology

Genomic methods are possibly also useful when diagnosing and treating bacterial, mycobacterial, and fungal infections including pathogens, virulence factors, strain forms, and antibiotic markers. Concrete evidence has shown that NGS has a significant potential for microbial identification of primary human specimens by both targeted sequences of rRNA genes and WGS. The existence of pathogens and virulence and/or resistance genes in one series may be investigated. A metagenomic approach may also be used to study the resistome, the study, and understanding of how resistance emerges and evolves, in addition to the detection of pathogens. The intestine is a known antibiotic resistance gene reservoir (ARG) and antibiotic therapy has an impact on the intestinal resistome that can lead to horizontal gene transfer and selection of resistant bacteria. To characterize mixed infections, especially those containing uncultivable or nonviable species, targeted 16S rRNA sequencing by NGS can have immediate clinical application. This strategy was successfully applied directly to the material of the brain abscess, lymph node biopsy, cystic fibrosis spuata, and mastoid abscess material.
Outbreak Management

More apparent and immediate use of WGS is the typing of bacterial pathogens for epidemiological monitoring, infection control, and outbreak analysis. Typing is unique to an organism and must be continually validated. Whole-genome sequencing, on the other hand, has the potential to trump conventional typing techniques by either in silico typing or superior differential ability. For example, multi-locus sequence typing (MLST) which is typically done by sequencing a collection of housekeeping genes, can be replicated by mapping WGS reads for the reference sequences of the genome, or by using the Simple Local Alignment Search Method (BLAST) for identifying alleles of the housekeeping genes. Whole-genome sequencing-based typing was used to identify a newly emerging CTX-M-15 producing Klebsiella pneumoniae with sequence type (ST) 1427. Moreover, genomic phylogenetic analysis was used to trace the transmission of CTX-M-15-producing ST15 K. pneumoniae between patients treated in a single center and the subsequent interinstitutional spread by patient referral. Whole-genome sequencing also enables controls to be introduced to prevent the spread of resistant bacterial clones. The outbreak of inter-institutional spread of colistin-resistant carbapenemase-producing K. pneumoniae (KPC) in the Netherlands was managed by the relocation of all positive residents to a different location outside the institution where patients were cared for by a dedicated team. It represents the value of NGS as a highly discriminatory method for separating clones with unique properties from the information obtained for patient care, infection prevention, and developmental studies. Genomics in Infectious Diseases

Molecular Case Finding

In complex and detailed outbreaks, NGS databases can retrospectively be checked for incidents. This can contribute to the identification of cases not identified in conventional epidemiological studies. In a study, a Dutch patient without a recent history of international travel was isolated from a long-term care center by a New Delhi Metallo-ß-lactamase-5 (NDM-5)-producing K. pneumoniae ST16 strain. The results of molecular cases showed that the Dutch strain was clonally linked to the four patients isolated in Denmark in 2014, but the cases in Danish and Dutch hospitals had no apparent epidemiological relations.

Taxonomy

The phylogenetic aspect was added to the taxonomy by Darwin. By incorporating 16S rRNA gene sequencing, the taxonomy of bacterial organisms has changed drastically. Usage of WGS for taxonomic purposes can be used for more genes to differentiate between organisms than the conventional approach for hybridizing DNA–DNA or 16S rRNA sequencing.

Zoonotic transmission

The first zoonotic transmission tests were focused on low-discriminatory approaches like serotyping. Next-generation sequencing gives these subjects a new viewpoint. Increased discrimination can show variations in previously undiscovered bacterial strains in animals and humans. In combination with epidemiological data, the source may track possible zoonotic infections. Next-generation sequencing also enables a detailed study of the manipulation of specific microbiota with implications for the transmission of interspecies through the use of antibiotics.

The analysis of the bacterial genomes improves awareness about microbial evolution.

Cell-free DNA

Microbial cell-free DNA (mcfDNA) is becoming an attractive infectious disease diagnostic modality, enabling wide-range pathogen identification, noninvasive screening, and rapid diagnosis. This is an advantages diagnostic approach for clinical infection, though this modern technique does have drawbacks and is not routinely applied in most laboratories. For immunocompromised pediatric patients, invasive fungal disease (IFD) presents a major risk of morbidity and mortality. Because of their prolonged neutropenic disorders, children and adolescents receiving chemotherapy, undergoing hematopoietic stem cell transplantation, or experiencing bone marrow failure are at elevated risk for IFD. To increase survival and decrease morbidity, prompt diagnosis of IFD with the administration of aggressive and appropriate antifungal therapy is essential. Invasive, sluggish to deliver a result, and sometimes lack sensitivity and identifying species levels are traditional diagnostic technics for fungal infection. Microbial cell-free DNA methods have been employed to classify non-human sequences and associate them with established bacterial, viral, and fungal pathogens genomic databases, for sequencing circulating cfDNA. In immunocompromised pediatric patients at risk of IFD, it revealed several infectious pathogens.

In a cohort of 10 immunocompromised patients with febrile neutropenia and pneumonia, 70% of cases, mcfDNA correlated with standard microbiological testing indicating that this technology may be useful in this clinical setting, especially in patients in whom tissue diagnosis is not feasible for bronchoscopy or biopsy. The findings of mcfDNA, however, have to be viewed with caution. And in combination with other radiological and clinical results in the laboratory.

Urinary cell-free DNA can be derived from its genome-wide methylation mark profile that can map the tissues of cell-free DNA origin and measure the abundance of a wide variety of viral and bacterial pathogens and the degree of host tissue injury from the contact between host and microorganism. This assay was used to test for urinary tract infection in patients with kidney transplantation and to find excellent evidence for the usefulness of this assay to differentiate between infections and infectious diseases and to determine the seriousness of the disease.

Meningitis

Extreme health conditions linked to elevated rates of morbidity and mortality worldwide are bacterial encephalitis and meningitis. In >50% of patients with acute encephalitis, however, unique pathogens are not known. In the wide variety of probable viral, autoimmune, neoplastic, paraneoplastic, Parameningeal, and toxic causes, subacute and recurrent meningitis are diagnostically difficult. A final diagnosis can require weeks or months of testing or remain unsolved and require inadequate or even harmful empirical treatment approaches. To identify all neuroinvasive pathogens, routine microbiological testing is often inadequate. Encephalitis can be caused by >100 different pathogens. Conventional approaches can be difficult to test all neurological pathogenic agents. Infectious agents could be identified independently by mNGS in a targeted way. It can classify all pathogens in a single run and can be a promising method for investigating clinical-causative pathogens with atypical characteristics in central nervous infection. In patients with...
diagnostically difficult subacute or chronic meningitis, subarachnoid neurocysticercosis, and toxoplasmic encephalitis, mNGS of cerebrospinal fluid (CSF) or brain tissue screens for almost all possible CNS infections and can identify novel or unexpected pathogens.48,49 Metagenomic sequencing CSF collected from meningitis or encephalitis patients enhanced the detection of neurological infections and, in some cases, offered actionable details.50

Other Clinical Utility
Patients with unexplained ocular infections, including chronic intraocular rubella infection, were diagnosed with mNGS. In pneumonia, mNGS and 16S NGS both identified pathogens using quantitative approaches.3 However, with the presence of commensal oral flora species, we need to be vigilant. Although numerous studies have studied the stool microbiome, only a handful have tried using mNGS techniques to diagnose related clinical diseases, such as diarrhea.3

COVID-19
Early understanding of the SARS-CoV-2 genome offered unprecedented insight into viral spread dynamics and affected response strategies. Global scientists knew the possible origin of the virus in <60 days of reporting, how similar it is to the viruses which are better known, and what treatments are available.51 This is the first time, due to NGS, that a full genome of a novel infectious agent has been made available to the public in such a limited time since the first case was identified to the WHO.51

As the virus transfers from host to host, it evolves across individuals and can acquire mutations that make it more virulent, enabling it to potentially affect more individuals. This evolutionary arms race across the globe with a multitude of strains makes it essential for researchers to quickly establish potential treatment and possible preventative measures to combat and halt the virus from continually evolving and affecting more people. Tracking the virus allows us to identify its source, develop future treatments based on how it evolves, and put appropriate potential preventative measures in place to limit its distribution.52 sequencing-based molecular modeling shows the variation in a vital surface protein SARS-CoV-2, which also has consequences for potentially vaccine policies.51 The results of NGS raise the possibility of a higher rate of efficacy of potential coronavirus vaccines and could enable the global scientific community to better understand COVID-19 epidemiology and spread. Viruses that mutate quickly over short periods make it difficult to develop effective vaccines that protect individuals from infection.

The SARS-CoV-2 test is a quick, accurate, and cost-effective tool to detect the SARS-CoV-2 virus. It can detect 237 viral-specific SARS-CoV-2 targets and 5 human gene expression controls to assess the presence of the virus in a variety of sample types (isolates, throat, and nasal swabs) and will cover all potential serotypes to help better understand epidemiology outbreaks.52 The first NGS test approved for use under the U.S. FDA Emergency Use Authorization (EUA) is the Illumina COVIDSeq Test.51 Whole-genome sequencing of SARS-CoV-2 virus can be compared using the phylogenetic analysis to identify how the virus is mutating and spreading and whether different strains are emerging (Fig. 3).

Limitation
It is important to consider the drawbacks of WGS. Most studies are currently focused on single nucleotide variants or SNPs known from comparisons with a reference genome sequence. The quality of the sequence and the genome assembly as well as the quality and selection of the reference genome are thus dependent on the analyzes.16 A large proportion of phylogenetic data is omitted by existing comparative studies. Whole-genome sequencing data may provide comprehensive genomic information, this does not translate to knowledge about gene

Fig. 3: This phylogeny shows the evolutionary relationships of the coronavirus (Courtesy: Nexstrain) Genomic epidemiology of novel coronavirus—Global subsampling. Ref: https://nextstrain.org/ncov/global?c=region
expression and transcription alone.16 The lack of validity and usefulness comparisons in clinical trials is a major drawback of WGS in clinical laboratories.16

**Quality**

As is true in all cases of testing used in public health and diagnostic laboratories, there is a need for rigorous quality assurance and standardization.16 With the increasing adoption of NGS for clinical microbiology, well-defined and widely sequenced microbial reference species will be required as controls and competence materials.18 Both positive and negative external controls are to be performed at each series and handled with the same lot of reagents and procedural workflow as separate samples.2 Samples should be treated not only in a sterile way but also to eliminate exogenous nucleic acid contamination.2 Regardless of the mechanism, an evaluation of potential cross-contamination should always be included in the analysis of metagenomic data, especially when a sequencing run involves a sample with high pathogen levels.3 It may also be beneficial to validate the findings of mNGS with Sanger sequencing about the existence or absence of species.3 The FDA has developed a database named FDARGOSS in conjunction with other federal agencies (FD A database for regulatory-grade microbial sequences; BioProject 231221). In addition, tests must be standardized, streamlined, checked, and ultimately automated if a solution is to progress from concept evidence to NGS-based diagnostics in US clinical laboratories and hospitals.17 To support the verification process for genomic sequencing-based testing methods, the National Institute for Standards and Technology (NIST) is developing whole-genome microbial reference standards.17

**Conclusion**

Metagenomic sequencing is a cutting-edge technology that has completely altered the way clinical diagnostics are conducted.3 Next-generation sequencing allows uniform typing procedures for all pathogens (“one test fits all”). But NGS approaches are intended to complement, rather than replace, traditional diagnostic testing at this point.18 As NGS technology continues to grow and evolve, the analytical performance of each test, the clinical validity for different pathogens, and, most importantly, the clinical niche for NGS will need to be identified by the research and medical communities.17 As with any strong new technology, molecular diagnosis faces difficulties in data analysis and reporting but can diagnose infections more easily and completely and provide important clinical management knowledge. Despite the difficulty and challenges that NGS assays present in clinical microbiology, they provide a remarkable opportunity to advance the field of clinical microbiology.17

**References**

1. Bleeker-Rovers CP, Vos FJ, de Kleijn EMHA, et al. A prospective multicentre study on fever of unknown origin: the yield of a structured diagnostic protocol. Medicine 2007;86(1):26–38. DOI: 10.1097/MD.0b013e31802fe858.
2. Ewig S, Torres A, Angeles Marcos M, et al. Factors associated with unknown aetiology in patients with community-acquired pneumonia. Eur Respir J 2002;20(5):1254–1262. DOI: 10.1183/09031936.02.0192001.
3. Gu W, Miller S, Chiu CY. Clinical metagenomic next-generation sequencing for pathogen detection. Annu Rev Pathol 2019;14(1):319–338. DOI: 10.1146/annurev-pathmechdis-012418-012751.
4. Ruud HD, Erik B, Monika AC. Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol 2017;243:16–24. DOI: 10.1016/j.jbiotec.2016.12.022.
5. Pallen MJ. Diagnostic metagenomics: potential applications to bacterial, viral and parasitic infections. Parasitology 2014;141(14):1856–1862. DOI: 10.1017/S0031182014000134.
6. Han D, LIR, Shi J, et al. Liquid biopsy for infectious diseases: a focus on microbial cell-free DNA sequencing. Theranostics 2020;10(12):5501–5513. DOI: 10.7150/thno.45554.
7. Bergholz TM, Moreno Switt Al, Wiedemann M. Omics approaches in food safety: fulfilling the promise? Trends Microbiol 2014;22(5):275–281. DOI: 10.1016/j.tim.2014.01.006.
8. Andreá I, Moreno S, Viviana T. Infectious diseases in the genomic era. Rev chil infect 2015;32(5):571–576. DOI: 10.4067/S0716-101820150000600103.
9. Simmer PJ, Miller S, Carroll KC. The promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. Clin Infect Dis 2018;66(5):778–788. DOI: 10.1093/cid/ciy881.
10. Greninger AL, Messacar K, Dunnebacke T, et al. Clinical metagenomic identification of Balanuthia Mandrillaris encephalitis and assembly of the draft genome: the continuing case for reference genome sequencing. Genome Med 2016;8(1):1. DOI: 10.1186/s13073-015-0257-9.
11. Wilson MR, Naccache SN, Samaya E, et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. New Eng J Med 2014;370(25):2408–2417. DOI: 10.1056/NEJMoa1401268.
12. Miao Q, Ma Y, Wang Q, et al. Microbiological diagnostic performance of metagenomic next-generation sequencing when applied to clinical practice. Clin Infect Dis 2018;67(Suppl_2):S231–S240. DOI: 10.1093/cid/ciy693.
13. Bruton OC. Agammaglobulinemia. Pediatrics 1952;9:722–728.
14. Bruton OC, Apt L, Gitlin D, et al. Absence of serum gamma globulins. AMA Am J Dis Child 1952;84:632–636.
15. Pannaraj PS, Pediatric Infectious Disease Society. Available at: https://www.pids.org/news/411-how-next-generation-sequencing-is-changing-pediatric-infectious-diseases.html.
16. Kwong JC, McCallum N, Sintchenko V, et al. Whole genome sequencing in clinical and public health microbiology. Pathology 2015;47(3):199–210. DOI: 10.1097/PAT.0000000000000235.
17. Goldberg B, Sichtig H, Geyer C, et al. Making the leap from research laboratory to clinic: challenges and opportunities for next generation sequencing in infectious disease diagnostics. mBio 2015;6(6):e01888-15. DOI: 10.1128/mBio.01888-15.
18. Leteforova MI, Suarez CJ, Banaei N, et al. Next-generation sequencing for infectious disease diagnosis and management: a report of the association for molecular pathology. J Mol Diagn 2015;17(6):623–634. DOI: 10.1016/j.jmoldx.2015.07.004.
19. Bentley DR, Balasubramanian S, Svardlow HP, et al. Accurate whole genome sequencing using reversible terminator chemistry. Nature 2008;456(7218):53–59. DOI: 10.1038/nature07517.
20. Rothberg JM, Hinz W, Rearick TM, et al. Integrated semiconductor device enabling non-optical genome sequencing. Nature 2011;475(7356):348–352. DOI: 10.1038/nature10242.
21. Grenier AL, Naccache SN, Federman S, et al. Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. Genome Med 2015;7(1):99. DOI: 10.1186/s13073-015-0220-9.
22. Radford AD, Chapman D, Dixon L, et al. Application of next-generation sequencing technologies in virology. J Gen Virol 2012;93(PT 9):1853–1868. DOI: 10.1099/vir.0.043182-0.
23. Weinstock GM. Genomic approaches to studying the human microbiota. Nature 2012;489(7415):250–256. DOI: 10.1038/nature11553.

24. Koren S, Schatz MC, Walenz BP, et al. Hybrid error correction and de novo assembly of single-molecule sequencing reads. Nat Biotechnol 2012;30(7):693–700. DOI: 10.1038/nbt.2280.

25. Gwinn M, MacCannell D, Armstrong GL. Next-generation sequencing of infectious pathogens. JAMA 2019;321(9):893–894. DOI: 10.1001/jama.2018.21669.

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

27. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 2014;15(3):R46. DOI: 10.1186/gb-2014-15-3-r46.

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

29. Pan W, Ngo TTM, Camunas-Soler J, et al. Simultaneously monitoring immune response and microbial infections during pregnancy through plasma cRNA sequencing. Clin Chem 2017;63(11):1695–1704. DOI: 10.1373/clinchem.2017.723888.

30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for high-throughput sequence data. Bioinformatics 2014;30(15):2114–2120. DOI: 10.1093/bioinformatics/btu170.

31. Aanensen DM, Feil EJ, Holden MTG, et al. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive staphylococcus aureus in Europe. mbio 2016;7(3):e00444-16. DOI: 10.1128/mBio.00444-16.

32. Gibson RM, Schmotzer CL, Quinones-Mateu ME. Next-generation sequencing to help monitor patients infected with HIV: ready for clinical use? Curr Infect Dis Rep 2014;16(4):401. DOI: 10.1007/s11908-014-0401-5.

33. Lataillade M, Chiarella J, Yang R, et al. Prevalence and clinical significance of HIV drug resistance mutations by ultra-deep sequencing identifies zoonotic transmission of MRSA isolates with the novel mecA homologue mecC. EMBO Mol Med 2013;5(4):509–515. DOI: 10.1002/emmm.201202413.

34. Armstrong AE, Rossoff J, Hollemon D, et al. Cell-free DNA next-generation sequencing successfully detects infectious pathogens in pediatric oncology and hematopoietic stem cell transplant patients at risk for invasive fungal disease. Pediatr Blood Cancer 2019;66(7):e27734. DOI: 10.1002/pbc.27734.

35. Maschmeyer G. Invasive fungal disease: better survival through early diagnosis and therapeutic intervention. Expert Rev Anti Infect Ther 2011;9(3):279–281. DOI: 10.1586/eri.11.11.

36. Camargo JF, Ahmed AA, Lindner MS, et al. Next-generation sequencing of microbial cell-free DNA for rapid non-invasive diagnosis of infectious diseases in immunocompromised hosts. F1000 Res 2019;8:1194. DOI: 10.12688/f1000research.19766.3.

37. Cheng AP, Burnham P, Lee JR, et al. A cell-free DNA metagenomic sequencing assay that integrates the host injury response to infection. PNAS 2019;116(37):18738–18744. DOI: 10.1073/pnas.1906320116.

38. Xing X-W, Zhang J-T, Ma Y-B, et al. Metagenomic next-generation sequencing for diagnosis of infectious encephalitis and meningitis: a large, prospective case series of 213 patients. Front Cell Infect Microbiol 2020;10:88. DOI: 10.3389/fcimb.2020.00088.

39. Hu Z, Weng X, Xu C, et al. Metagenomic next-generation sequencing as a diagnostic tool for toxoplasmic encephalitis. Ann Clin Microbiol Antimicrob 2018;17(1):45. DOI: 10.1186/s12941-018-0298-1.

40. Wilson MR, O’Donovan BD, Gelfand JM, et al. Chronic meningitis investigated via metagenomic next-generation sequencing. JAMA Neurol 2018;75(8):947–955. DOI: 10.1001/jamaneurol.2018.0463.

41. Wilson MR, Sample HA, Zorn KC, et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. N Engl J Med 2019;380(24):2327–2340. DOI: 10.1056/NEJMoa1803396.

42. Obrehter Michael, Complex Disease Genomics, Microbial Genomics, 2020. Available at: https://sapac.illumina.com/company/news-center/feature-articles/illumina-perspective-on-the-novel-coronavirus--covid-19--outbreak.html.

43. https://thermofisher.mediaroom.com/2020-03-25-Italian-Researchers-Identify-New-SARS-CoV-2-Gene-Variants-That-Provide-Clues-to-Coronavirus-Epidemiology.