Mini-Review

iMedPub Journals
http://www.imedpub.com

Archives of Clinical Microbiology
ISSN 1989-8436

2021
Vol. 12 No. S5:176

Diagnostic Metagenomics as Diagnostic Tools in Infectious Diseases

Kendall Kling[1], Teresa Zembower[1,2] and Chao Qi[1*]

[1] Department of Pathology, Northwestern University Feinberg School of Medicine, Clinical Microbiology Laboratory, Northwestern Memorial Hospital, Chicago IL, USA

[2] Department of Medicine, Division of Infectious Diseases, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

***Corresponding author:** Chao Qi

Department of Pathology, Northwestern University Feinberg School of Medicine, Clinical Microbiology Laboratory, Northwestern Memorial Hospital, Chicago IL, USA

fqi451@northwestern.edu

Abstract

Treatment of infections with culture-negative results represents a challenge to clinicians. Culture-independent broad-range nucleic acid sequencing methodologies capable of identifying pathogens at the genomic level have emerged as powerful diagnostic tools. While traditional Sanger sequencing only allows a single target detection, Metagenomic Next-Generation Sequencing (mNGS) enables the detection of all types of organisms, including viruses, bacteria, fungi, and parasites, in a single-assay. Its potential for broad-range pathogen detection with high sensitivity offers a diagnostic opportunity for patients with a wide-range of differential diagnoses. Currently there is no FDA-approved metagenomic diagnostic test. Clinical laboratories face many obstacles in implementing mNGS assays. Streaming sequencing process, developing and validating a comprehensive database, and standardizing the bioinformatics pipelines for data analysis are among the challenges to be addressed. When the technology becomes more readily available in clinical laboratories, its value for improving infectious diseases diagnosis will be realized.

Keywords: Infections; Microbiology; Diagnosis; Bacteria; Diseases

Received: September 17, 2021; **Accepted:** October 01, 2021; **Published:** October 08, 2021

Citation: Kling K, Zembower T, Qi C (2021) Diagnostic Metagenomics as Diagnostic Tools in Infectious Diseases. Arch Clin Microbiol Vol.12 No. S5: 176.

Introduction

The clinical microbiology laboratory remains a critical tool for the physician in diagnosis of infectious diseases. Classic diagnostic methods for pathogen detection involve culturing samples from patients, waiting for growth, and performing *in vitro* susceptibility testing. While this form of diagnostics certainly has stood the test of time, it has some major flaws. Firstly, many patients have or are receiving antimicrobial therapy at the time of sample collection, which can greatly impact the organism's ability to grow in culture and can cause false negative results. In addition, many fastidious organisms that require certain specific growth conditions cannot or will not grow with standard culture media, also leading to false negative results. Anaerobic bacteria are also infamously difficult to grow in the laboratory and require very strict transport media after sample collection, likely leading to under diagnosis. It is evident that new technologies are needed to fill the gaps that culture-based diagnostics cannot provide. DNA sequencing techniques, both traditional chain terminations Sanger sequencing and newer next generation sequencing techniques, are alternative diagnostic modalities for culture negative infections that are gaining a foothold in modern medical practice.

Literature Review

Diagnostic application of next generation sequencing for infectious disease

Next Generation Sequencing (NGS) emerged in the mid 1990's as an alternative to traditional Sanger sequencing as a massively parallel, high throughput, and automated process that could analyze clinical samples [1]. Initially, NGS was utilized in human microbiome studies and has a major niche in cancer research and therapeutics. Now its application extends to the diagnosis of clinical infectious diseases [2]. It allows unbiased detection of all organisms present in a sample, including viruses, bacteria, fungi, and parasites. Its potential for pathogen detection in culture negative samples has been extensively explored.

Pathogen detection with NGS involves the following steps: Nucleic acid extraction from clinical specimens, library preparation, next-generation sequencing, bioinformatic analysis of sequence data with any combination of trimming, assembly, annotation, typing, and phylogeny [3]. There are at least 20 commercially available NGS platforms. They differ in sequencing techniques to achieve massive parallel sequencing. The sequences obtained are

characterized as "reads" which can be either short (75-600 base pairs) or long (100-10,000 base pairs), and having (or the presence of) more reads increases the sensitivity of detecting lower levels of pathogen nucleic acids in a background of high-level host DNA [3]. The turnaround time varies significantly depending on the sequence technology and analytic programs used, and can range from 6 hours to 7 days [3].

Different sequencing strategies have been developed, but in general they can be classified as either targeted amplicon sequencing or shotgun metagenomics [2]. In the targeted amplicon sequencing method, the 16S rRNA gene is amplified prior to undergoing NGS. Species and strain identification can be challenging with this method due to the well conserved nature of the 16S rRNA gene sequence among closely related species [2]. For example, closely related Streptococci, Staphylococci, Enterococci, and some *Enterobacteriaceae* cannot be identified at the species level with this method [4]. Sequencing of other genes has been postulated to improve ability to distinguish species within genera (i.e., *rpoB* gene), but these methodologies have not yet taken hold given lack of established reference databases [2]. For fungal infections, some methodologies sequence the 18S rRNA gene; however, the preferred amplicon is the ribosomal Internal Transcribed Spacer (ITS) gene due to its higher variability and more extensive database than 18S [2]. In the shotgun method the DNA is randomly fragmented, sequenced, and then reassembled based on overlap sequences [2]. This technique can provide higher resolution for species identification, detect antimicrobial resistance markers, and separate strains of the same species based on the genome sequences [2]. Bioinformatics algorithms are required to analyze and interpret the results [2].

Retrospective studies of the diagnostic value of metagenomic sequencing have largely focused on proof-of-concept and diagnostic specificity and sensitivity in comparison with culture. In 2016, Cummings et al., analyzed 15 sputum samples from cystic fibrosis patients and subjected them to traditional culture and Illumina Inc. (San Diego, CA) 16S rRNA NGS and found that NGS detected every pathogen that grew in culture in addition to 3 bacteria that culture missed (*Streptococcus agalactiae*, *Burkholderia cepacia*, and *Achromobacter xylosoxidans*) and concluded that NGS was better for species level identification [1]. In 2017, Ruppe et al., performed metagenomic NGS on 24 bone and joint fluid specimens and found 100% agreement with culture in monomicrobial samples. At least 182 additional pathogens were identified, but NGS had more difficulty in species level identification for polymicrobial specimens [5]. Several studies were published in 2019, including one by Charalampous et al., who evaluated 40 respiratory specimens and compared Nanopore metagenomics sequencing to traditional culture. They found NGS had 91% sensitivity and 100% specificity, and 5 bacteria were identified in NGS but not culture [6]. Also, Zinter et al. retrospectively collected 41 respiratory specimens from immunocompromised children and found that in the 24 culture negative specimens, metagenomics NGS was able to identify a pathogen in 11 (46%) [7]. They also found that in the 17 clinical positive (culture or non-NGS PCR) cases, NGS detected the known pathogens and differed only in Coronavirus species identification in 3 cases [7]. NGS also detected 3 additional viral pathogens in the clinical positive cases [7]. Additionally, Wilson et al., applied metagenomic NGS to 58 Central Nervous System (CNS) infections in critically ill hospitalized patients and found that NGS identified 13 causes that routine testing did not. However, NGS failed to detect a pathogen in 26 cases that were diagnosed *via* conventional methods including serology or other tissue specimens [8].

A limited number of reference laboratories offer NGS diagnostics to infectious diseases physicians and have been studied. The most utilized service is the Karius Test™ (Karius, Inc., Redwood, CA) which is performed on cell-free plasma and performs metagenomic NGS *via* Illumina NextSeq, but notably does not perform RNA sequencing so will not detect RNA viruses [9]. IDbyDNA Inc. (Salt Lake City, UT) performs DNA and RNA sequencing of respiratory specimens and also uses Illumina technology [9]. Finally, the University of California San Francisco offers a DNA/RNA mNGS *via* Illumina platform [9]. In 2021, Hogan et al., analyzed conducted a retrospective cohort analysis of all patients for whom Karius testing was performed for all indications at 5 US institutions over 1.5 years [9]. They analyzed 82 Karius tests from 39 adults and 43 children, 53 of who were immunocompromised, and found that Karius had a positive impact in 7.3%, negative impact in 3.7%, no impact in 86.6%, and indeterminate impact in 2.4% [9]. In 2019, the UCSF Illumina CerebroSpinal Fluid (CSF) mNGS platform was evaluated by Miller et al., who studied 95 CSF specimens and compared NGS results to conventional testing [10]. They and found a sensitivity of 73% with specificity of 99% compared with original test results [10]. Of the 21 mNGS results considered false-negatives, 4 were RNA viruses (1 enterovirus and 3 West Nile virus (WNV) diagnosed by CSF serology), 4 were DNA viruses (2 VZV, 1 HSV-2, 1 EBV, diagnosed by PCR), 9 were bacteria (diagnosed by culture), and 4 were fungi (diagnosed by culture and/or antigen testing) [10]. The mNGS-negative but clinically positive specimens tended to be diseases diagnosed by serologic testing, low colony count bacteria, and high host genomic DNA background that appeared to interfere with results [10]. The mNGS detected 18 pathogens not found on conventional tests including HIV, other viruses, and one Bacillus species [10]. In 2021, Azar et al. studied the IDbyDNA platform on 30 immunocompromised patients' lower respiratory tract specimens compared to conventional testing and found that NGS increased the total sensitivity 23% and had the greatest impact on bacterial pathogens [11].

Development of implementation of metagenomic assays in clinical laboratories

Lacking an FDA-approved commercial diagnostic metagenomic test is the major obstacle for bringing the technology into clinical microbiology laboratories. The key steps to develop a diagnostic metagenomic test involves optimizing the nucleic acid extraction for a variety of sample types, controlling sequencing quality, creating sequence libraries from negative controls to remove the contaminations introduced during sample collection

and library construction, standardizing bioinformatic analysis process, and establishing threshold for reporting clinically significant organisms. As no gold standards currently exist, the practices in clinical laboratories vary significantly. Recently, Miller et al., developed and validated a shot-gun based metagenomic sequencing assay for pathogen detection in CSF [10]. A laboratory developed bioinformatics pipeline was used in the assay [10]. Using 95 clinical samples tested with conventional clinical testing methods, the assay was reported to have 73% sensitivity and 99% specificity, and 81% positive percent agreement and 99% negative percent agreement after judication [10]. Our group validated a 16S rRNA metagenomic assay using the Ion 16S metagenomics kit and Ion Reporter metagenomics workflow for bacterial detection in clinical samples from sterile sources [12]. We established a threshold for reporting detected organisms based on serial testing of positive controls and characterized its diagnostic performance by testing 98 samples that had results produced by manual Sanger sequencing [12]. The assay had an overall sensitivity of 88% at the species level and 76% specificity, whereas for genus level NGS had 100% sensitivity. It produced positive results in 13 cases where traditional 16S was negative [12]. We showed that the assay is reproducible and produced results superior to single target Sanger sequencing.

Development of a metagenomic assay for viral detection was reported by Sander van Boheemen group [13]. Compared to the single target PCR, the developed protocol had a sensitivity of 83% and specificity of 94% respectively, and more samples were positive by the NGS assay [13].

The clinical niche for diagnostic metagenomics in providing diagnosis for patients with negative results by conventional tests and broad differentials are well accepted. The commercially available products like the Ion 16S metagenomics kit and Ion Reporter metagenomics workflow provide reagents for wet-lab steps as well as the bioinformatic data analysis package and curated reference database for the dry-lab process, making it possible to streamline the workflow and reduce many variables influencing the sensitivity. Efforts to standardize the laboratory process are underway. The College of American Pathologists' (CAP) published eighteen requirements in an accreditation checklist for the NGS analytic and bioinformatics processes as part of the molecular pathology checklist [14]. Clinical and Laboratory Standards Institute (CLSI) has integrated specific requirements for the application of NGS in Microbiology into the "Nucleic acid sequencing methods in diagnostic laboratory medicine" guidelines [15]. A regulatory framework guiding the development and implementation of diagnostic metagenomic tests in clinical laboratories is available. Budowle published validation guidelines for NGS and microbial forensic applications that outlined the steps and specific criteria for assay validation [16]. In a report by Schlaberg et al., strategies for how to design metagenomic test validation to fulfill regulatory requirements were discussed [17].

As Gargis et al., indicated in a review article on quality assurance for implementing NGS tests in clinical microbiology and public health laboratories, additional standards and guidance are needed to move diagnostic metagenomics into clinical laboratories [18]. For assay development, developments of comprehensive reference databases that are validated and accessible to all clinical laboratories are critical to generate accurate identification. Bioinformatics pipelines that can be used to analyze sequence data from the commonly used sequencing platforms should be standardized and made available to the clinical laboratories. For assay validation, executable guidelines to characterize assay performance specifics and commercially available quality control panels are needed. For quality management, development of CAP proficient test challenges for NGS-based assays for infectious diseases is an indispensable laboratory component to ensure the reliability of NGS-based test results.

Conclusion

Infectious diseases continue to be a leading cause of morbidity and mortality in the world and appropriate treatment requires accurate and timely testing of clinical specimens. Conventional culture-based techniques are subject to many potentially false-negative results due to fastidious nature of pathogens, improper culture conditions, and impaired viability. Culture-independent broad-range nucleic sequencing methodologies, identifying the DNA or RNA of novel or unexpected pathogens in a single-assay, have emerged as alternative diagnostic modalities. Although the traditional Sanger sequencing continues to be utilized for single target sequencing for bacterial and fungal identification in culture negative samples, we anticipate it will be replaced by more recently developed mNGS technologies for more comprehensive organism coverage, improved sensitivity, high throughput, and potential for automation. Wide use of diagnostic metagenomics is hampered due to the technical complexity and difficulty in standardization. When the technology becomes more readily available in clinical laboratories, its value for improving infectious diseases diagnosis will be realized.

References

1. Cummings LA, Kurosawa K, Hoogestraat DR, SenGupta DJ, Candra F, et al. (2016) Clinical next generation sequencing outperforms standard microbiological culture for characterizing polymicrobial samples. Clin Chem 62: 1465-1473.
2. Boers SA, Jansen R, Hays JP (2019) Understanding and overcoming the pitfalls and biases of Next Generation Sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. Eur J Clin Microbiol Infect Dis 38: 1059-1070.
3. Patricia JS, Steven M, Carroll KC (2018) Understanding the promises and hurdles of metagenomic next-generation sequencing as a diagnostic tool for infectious diseases. Clin Infect Dis 66: 778-788.
4. Sibley CD, Peirano G, Church DL (2012) Molecular methods for pathogen and microbial community detection and characterization: Current and potential application in diagnostic microbiology. Infect Genet Evol 12 505-521.

5. Ruppe E, Lazaredic V, Girard M, Mouton W, Ferry T, et al. (2017) Clinical metagenomics of bone and joint infections: A proof of concept study. Nat 7: 1-12.

6. Charalampous T, Kay GL, Richardson H, Aydin A, Baldan R, et al. (2019) Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. Nat Biotechnol 37: 783-792.

7. Zinter MS, Dvorak CC, Mayday MY, Iwanaga K, Ly NP, et al. (2019) Pulmonary metagenomic sequencing suggests missed infections in immunocompromised children. Clin Infect Dis 68: 1847-1855.

8. Wilson M, Sample H, Zorn K, Arevalo S, Yu G, et al. (2019) Clinical metagenomics sequencing for diagnosis of meningitis and encephalitis. N Eng J Med 380: 2327-2340.

9. Hogan CA, Yang S, Garner OB, Green DA, Gomez CA, et al. (2021) Clinical impact of metagenomic Next-Generation Sequencing of plasma cell-free dna for the diagnosis of infectious diseases: A multicenter retrospective cohort study. Clin Infect Dis 72: 239-245.

10. Miller S, Naccache SN, Samayoa E, Messacar K, Arevalo S, et al. (2019) Laboratory validation of a clinical metagenomic sequencing assay for pathogen detection in cerebrospinal fluid. Genome Res 29: 831-842.

11. Azar MM, Schlaberg R, Malinis MF, Bermejo S, Schwarz T, et al. (2021) Added diagnostic utility of clinical metagenomics for the diagnosis of pneumonia in immunocompromised adults. Chest 159: 1356-1371.

12. Malcynski M, Zembower T, Zhu A, Qi C (2021) Diagnostic performance of Ion 16S metagenomics kit and ion reporter metagenomics workflow for bacterial pathogen detection in culture negative clinical specimens from sterile sources. Disease Diagn Microbiol Infect 101: 115451.

13. Sander VB, Rijn VAL, Pappas N, Carbo EC, Vorderman RH, et al. (2020) Retrospective validation of a metagenomic sequencing protocol for combined detection of RNA and DNA viruses using respiratory samples From pediatric patients. J Mol Diagn 22: 196-207.

14. Aziz N, Zhao Q, Bry L, Driscoll DK, Funke B, et al. (2014) College of American pathologists' laboratory standards for next-generation sequencing clinical tests. Arch Pathol Lab Med 139: 481-493.

15. CLSI (2014) Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine: Approved guideline- 2nd edition. MM09-A2.

16. Budowle B, Connell ND, Bielecka-Oder A, Colwell RR, Corbett CR, et al. (2014) Validation of high throughput sequencing and microbial forensics applications. Invest Genet 5: 9.

17. Schlaberg R, Chiu CY, Miller S, Procop GW, Weinstock G (2017) Validation of metagenomic next-generation sequencing tests for universal pathogen detection. Arch Pathol Lab Med 141: 776-786.

18. Gargis AS, Kalman L, Lubin IM (2016) Assuring the quality of next-generation sequencing in clinical microbiology and public health laboratories. J Clin Microbiol 54: 2857-2865.