The next generation of ocular pathogen detection

Sharon L. Sabapathypillai
Washington University School of Medicine in St. Louis
Hayley R. James
Washington University School of Medicine in St. Louis
Rachael R. L. Lyerla
Washington University School of Medicine in St. Louis
Lynn Hassman
Washington University School of Medicine in St. Louis

Follow this and additional works at: https://digitalcommons.wustl.edu/oa_4

Part of the Medicine and Health Sciences Commons

Please let us know how this document benefits you.

Recommended Citation

This Open Access Publication is brought to you for free and open access by the Open Access Publications at Digital Commons@Becker. It has been accepted for inclusion in 2020-Current year OA Pubs by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
The Next Generation of Ocular Pathogen Detection
Sharon L. Sabapathypillai, BS, Hayley R. James, MD, Rachael R.L. Lyerla, MD, and Lynn Hassman, MD, PhD

Abstract: Metagenomic next-generation sequencing is a powerful method for pathogen detection that combines advanced genome sequencing technology with cutting-edge bioinformatics to analyze microbial populations. Metagenomic next-generation sequencing has the potential to identify uncommon, unculturable, and even previously unidentified pathogens from a clinical isolate. Of particular interest to ophthalmology, this robust data extraction can occur from very small volume clinical samples. Here we discuss the opportunities and limitations of this technique and their current and future application to ophthalmic diagnostics.

Key Words: endophthalmitis, next-generation sequencing, uveitis

INTRODUCTION
Clinical metagenomics next-generation sequencing (mNGS) is the comprehensive analysis of genetic samples (DNA or RNA) to study microbial populations. Next-generation sequencing (NGS) technology, which is also referred to as deep sequencing, high-throughput sequencing, and massively parallel sequencing, allows the simultaneous sequencing of all the genomic material present in small volume clinical samples. This data-rich investigation relies on NGS technology, which represents a revolution from previous Sanger sequencing methods in that nucleic acid sequencing occurs by rapid, relatively cheap and high-throughput methods, allowing millions of DNA/RNA segments to be simultaneously sequenced. NGS is inherently unbiased and hypothesis-free, theoretically facilitating detection of uncommon, unculturable, and even previously unidentified pathogens from a clinical isolate. Importantly, the term metagenomics encompasses the analysis of entire populations of microorganisms. It draws on the large data provided by NGS, and high-level bioinformatic analysis and interpretation of the data into clinically useful information (Table 1).

Since its inception, mNGS has been used to improve the identification of infectious microbes in many specimens currently evaluated with cultures or targeted polymerase chain reaction (PCR), for diseases such as culture-negative sepsis. This technique has also led to increased awareness of the microbial diversity present in respiratory secretions, cerebrospinal fluids, stool, blood, tissue, gastrointestinal infections, and ocular tissues and is also being studied to translate lung microbiomes to develop a culture-independent treatment algorithm for respiratory infections.

Clinical metagenomics is beginning to play an important role, not only in the identification of the normal flora of the human eye, but also in evaluating pathogens that may be important in ocular diseases. With the increasing advances of clinical metagenomics, there is likely to be a trend toward the use of mass sequencing-based diagnostics in the identification of ocular pathogens and ultimately in the treatment of ocular infections.

TRADITIONAL METHODS OF PATHOGEN DETECTION
Culture is the current gold standard for pathogen detection and requires minimal prior knowledge of candidate microorganisms; however, it is time-consuming and limited to organisms that can be grown on standard culture media. Other traditional methods for pathogen detection in ocular disease include PCR, a technique that is well suited for herpes viruses which replicate to high titers in the eye; however, the sensitivity for organisms like toxoplasmosis can be as low as 28% to 53% for aqueous fluid. PCR reactions are specific to individual organisms and therefore require prior knowledge of likely pathogens and sufficient ocular fluid to perform each reaction. Important drawbacks of PCR include false positives due to multiple amplifications and false negatives if primers mismatch to the organism’s genome. Several techniques have been developed to overcome some of these challenges, including nested or dual target PCR reactions in which a second primer set improves specificity when low pathogen genomes necessitate additional amplification cycles. This technique can improve sensitivity of toxoplasmosis detection to nearly 100%. Additionally, multiplexed PCR reactions conserve sample volumes and allow the analysis of multiple potential pathogens in a single reaction.

In addition to PCR, the Goldmann-Witmer coefficient has been used to detect pathogen-specific intraocular antibodies. The Goldmann-Witmer coefficient is helpful in establishing pathogen-specific antibody production within the ocular microenvironment when the ratio of pathogen-specific antibodies to total antibodies is higher in aqueous fluid compared to serum. Researchers have found the usage of intraocular antibody in aqueous humor testing to significantly contribute to the pathogen detection in infectious
5. Bioinformatic analysis involves
   a. data clean-up to remove adapter sequences and low-quality reads
   b. sequence alignment and removal of human genome sequences
   c. alignment of the remaining sequence reads to a pathogen database
   d. taxonomical classification
6. Interpretation of results, understanding that
   b. rare or new pathogens may not be represented in sequence databases
   c. pathogens may be difficult to distinguish from related organisms (ie, mycobacterial species) due to genetic similarity
   d. nonpathogens including normal flora or laboratory contaminants can be inadvertently amplified
7. Ultimately the determination of pathogenicity must factor in the clinical context.

<table>
<thead>
<tr>
<th>TABLE 1. Clinical Application of Next-generation Sequencing</th>
</tr>
</thead>
</table>
| 1. Sample collection  
  a. Important considerations include sample storage and processing if RNA will be assessed |
| 2. Nucleic acid extraction  
  a. Methods specific to DNA or RNA |
| 3. Library preparation  
  a. In the case of RNA specimens, the RNA must first be reverse-transcribed to cDNA  
  b. This step often requires amplification which can introduce bias toward more abundant amplicons |
| 4. Sequencing  
  a. Shotgun (best for truly unbiased pathogen identification)  
  b. Targeted (cheaper, best for amplifying bacterial or fungal species) |
| 5. Bioinformatic analysis involves  
  a. data clean-up to remove adapter sequences and low-quality reads  
  b. sequence alignment and removal of human genome sequences  
  c. alignment of the remaining sequence reads to a pathogen database  
  d. taxonomical classification |
| 6. Interpretation of results, understanding that  
  b. rare or new pathogens may not be represented in sequence databases  
  c. pathogens may be difficult to distinguish from related organisms (ie, mycobacterial species) due to genetic similarity  
  d. nonpathogens including normal flora or laboratory contaminants can be inadvertently amplified |
| 7. Ultimately the determination of pathogenicity must factor in the clinical context. |

**ADVANCED PATHOGEN DETECTION WITH mNGS**

Clinical mNGS is revolutionizing the detection of potential human pathogens due to its ability to detect organisms that can be missed by standard culture or targeted PCR techniques. Additionally, because all genomic material in a sample is sequenced, this technique is inherently unbiased, allowing the identification of organisms that may not have been suspected clinically and discovery of previously unknown organisms. In addition to simple pathogen detection, mNGS offers auxiliary genomic information for evolutionary or epidemiologic tracing, strain identification, and prediction of drug resistance.1

NGS can be done to identify both DNA and RNA sequences. Although RNA sequencing requires special processing, DNA can be deeply sequenced from fresh, frozen, or even paraformaldehyde-fixed tissues.1 Shotgun NGS refers to sequencing performed simultaneously across random DNA fragments and is the most unbiased approach, generally best for identifying most viruses, some fungi, and previously unidentified pathogens.2 Targeted amplicon sequencing can be utilized to amplify microbial rRNA before sequencing.10,11 In this process, conserved microbial elements such as ribosomal RNA (rRNA) elements, 16S rRNA in the case of bacterial species, and 28S rRNA or internal transcribed spacer for fungal species, are first amplified by generic primers, then sequenced to identify the specific bacterial or fungal species.8 Importantly, this technique is less useful for viruses due to their sequence variability, or eukaryote parasites due to the similarity of their rRNA sequences to human. False positive results can occur as a result of presequencing amplification of contaminants. Finally, whole genome sequencing can be performed to quantify organisms, identify strain-specific mutations, specific virulence or antibiotic resistance genes, or even characterize novel pathogens.1 This approach is best used for clinical isolates or samples with low microbial diversity. Multiple sequencing platforms are utilized and are beyond the scope of this review; however, the interested reader is directed to an excellent overview by Gu et al.1

An important consideration in the application of NGS to metagenomic pathogen identification is that in most clinical samples, the vast majority of nucleic acid present will actually be human. Human nucleic acid can be removed before DNA/RNA extraction by lysing the human cells with saponin or other chemicals and degrading any DNA not protected by bacterial or fungal cell walls or viral capsids.12 Alternatively, human ribosomal and mitochondrial RNA can be significantly reduced before sequencing by hybridization to human sequence-specific probes or via Cas9 nuclease digestion.13 Both of these methods result in some concurrent loss of pathogen nucleic acid. Finally, targeted amplicon sequencing can be utilized as noted above to enrich microbial rRNA.14

Quality control is highly critical at every step of this process from sample acquisition and storage to bioinformatics algorithms and these considerations were thoroughly reviewed by Gu et al.1 Practically speaking, careful attention to collection with rapid cryostorage and advanced planning of sample processing and analysis are critical to detect RNA viruses as RNA degrades quickly and will not be detected by DNA sequencing methods without an additional conversion to cDNA. Extensive bioinformatic processing of the data must be performed to filter out host sequences and environmental contaminants and to correctly align remaining potential pathogen sequences with genomic databases.1 Finally, the clinician must interpret the results of mNGS to determine whether the organisms detected are likely to be pathogenic in their patient, which may require consultation with infectious disease specialists.

**UTILITY AND POTENTIAL IN UVEITIS**

Diagnosis of intraocular infections can be complex, as clinical presentations of specific pathogens can have significant overlap. Traditionally, diagnosis was dependent upon diagnostic testing like PCR and bacterial cultures. Unfortunately, culture-based assays have a relatively low sensitivity, ranging from 40% to 70%15,16 and are time-consuming. Molecular diagnostic testing like PCR is limited by its prespecified targets and can only recognize a fraction of ocular pathogens. These techniques are additionally limited by the relatively small sample volume that can be safely obtained from the anterior chamber or the vitreous. Given these limitations, it is not surprising that the causative agent is unidentified in >50% of cases of presumed infectious uveitis.16,17 Metagenomics has the potential to radically increase the diagnostic yield due to its unbiased nature and high throughput. Additionally, it has the capability to detect bacteria, fungi, DNA
and RNA viruses, and parasites from a <50 μL of intraocular fluid.18,19

This technique has been used to identify causative organisms in previously unidentified intraocular infections. Gonzales et al used mNGS to identify rubella virus RNA from 6 patients with recurrent or chronic hypertensive anterior or intermediate uveitis which had previously tested negative for typical viral PCR [herpes simplex virus 1 and 2 (HSV-1 and HSV-2), varicella-zoster virus, and cytomegalovirus (CMV)].20 thus expanding the etiologic spectrum for Fuch iridocyclitis to include rubella. Doan et al also utilized the technique to identify organisms including CMV, human herpes virus 6, HSV-2, human T lymphotropic virus 1, Klebsiella pneumoniae, and Candida dubilensis in 8 of 36 archived vitreous samples that had previously tested negative for HSV, varicella-zoster virus, CMV, and Toxoplasma gondii.19 Importantly, in this study, mNGS also correctly identified control ocular specimens with HSV-1, C. neoformans, and T. gondii, and did not detect anything beyond laboratory contaminants in control noninfectious and noninflammatory samples, establishing a framework for sensitivity and specificity of this novel technique for use in ocular samples.

Remarkably, Doan et al demonstrated that ocular rubella sequences isolated from 1 patient with bilateral chronic anterior and intermediate uveitis aligned with a strain known to circulate in the geographic region and time at which the patient had an unexplained febrile illness as a child,18 illuminating the potential for mNGS to facilitate epidemiologic investigations.

In a different approach, deep-sequencing performed after 16s amplification of ocular specimens from 21 patients with presumed infectious endophthalmitis and 7 control patients showed good concordance with bacterial and fungal culture, and additionally revealed Streptococcus and Pseudomonas in culture-negative samples.21 In a follow-up study, Lee et al used whole genome sequencing to quantify pathogen genomes and discovered that unlike endophthalmitis caused by other organisms, the burden of S. epidermidis organisms did not correlate with severity of endophthalmitis.22

Lee et al also revealed for the first time, a presumably nonpathogenic human anellovirus, torque teno virus (TTV), in 57.1% of the culture-positive specimens and 100% of the culture-negative specimens, but none of the control samples.21 TTV is found at high titers in patients with multiple diseases, including multiple sclerosis, systemic lupus erythematosus, fever of unknown origin, rheumatic diseases, and hepatitis, and has been associated with increased rates of posttransplant infection and transplant rejection,23 suggesting that TTV may be associated with altered immune responses and/or may confer addition pathogenicity to other disease states. In support of this, Lee et al also discovered that the presence of TTV conferred a 5-fold risk for primary vitreoretinal lymphoma.27 The first patient showed the presence of both Epstein-Barr virus and human herpes virus 8, both of which cause primary effusion lymphoma.27 The second patient was negative for infectious agents; however, analysis of the host sequence revealed the S243N mutation in MYD88, which has been associated with the activated B-cell subtype of diffuse large B cell lymphoma.27 Because this mutation is less common than the L265P mutation traditionally associated with primary vitreoretinal lymphoma,27 it may have been missed by typical PCR analysis of the aqueous specimen, suggesting that mNGS is also more sensitive for detecting noninfectious ocular disease.

**UTILITY AND POTENTIAL IN CONJUNCTIVITIS, CORNEAL TRANSPLANTS, OR OTHER DISEASES OF THE OCULAR SURFACE**

Metagenomics also has the potential to change the way we analyze and treat diseases of the ocular surface, including diseases of the cornea and conjunctiva. The organisms populating the ocular surface have previously been identified using culture-based techniques; however, as in the case of culture-negative endophthalmitis, this methodology has the potential to miss many microorganisms that do not successfully grow in the lab. With the application of expanded molecular sequencing, such as mNGS, researchers suggest that the normal microbiota of the healthy ocular surface is more diverse than previously thought, although still far less diverse than skin, oral, and intestinal microbiomes. In addition, the ocular microbiome hosts relatively few bacterial taxa that are relatively stable across individuals, and this may point to a role for the microbiome in immune modulation.

Metagenomics RNA sequencing has been used to help identify potential causative pathogens in infectious conjunctivitis and keratitis. Lalitha et al (2020) extracted RNA from inferior fornix conjunctival swabs of the 14 eyes sampled, 12 (86%) were positive for human pathogens, with adenovirus being the most common etiology followed by a fungal organism, Vittaforma Corneae. Seitzman et al also utilized metagenomic sequencing of aqueous fluid to identify a case of Capnocytophaga keratitis that could not be cultured. Thus, mNGS may also be used for the identification of causative organisms that could help target
appropriate antibiotic use and decrease the inappropriate use of antibiotics in the case of viral, fungal, or noninfectious conjunctivitis.

Metagenomics may also be applicable for screening for microorganisms present in the storage media of human donor cornea tissue. The standard commercially available cornal storage media contain penicillin and streptomycin for antibacterial coverage and amphotericin B for antifungal coverage. Parekh et al (2019) extracted genetic material from multiple bacterial and fungal organisms such as Pseudomonas spp, Comamonas spp, and Malassezia sp which were present in both the room temperature and hypothermic storage media. Importantly, additional methods to confirm the presence of live organisms were not performed, highlighting the fact that mNGS cannot confirm the presence of live microorganisms, nor their eradication by antimicrobials in the organ culture medium.

**CURRENT AND FUTURE APPLICABILITY OF NGS**

Patients who are immune-compromised in tertiary care medical centers are vulnerable to infections by both common and uncommon viruses, bacteria, fungi, and parasites. Recovery of such pathogens via routine cultures can be limited by the ability of pathogens to grow in the laboratory, and use of antimicrobial drugs. Most current PCR-based methods are individual tests for specific organisms and will miss uncommon or previously untested organisms. NGS represents an unbiased and hypothesis-free diagnostic approach that can detect nearly any pathogen in live microorganisms, nor their eradication by antimicrobials in the organ culture medium.

mNGS is currently not utilized by clinical laboratories due to the extensive bioinformatics analysis required; however, analysis algorithms like the one utilized by the Van Gelder laboratory for application to ocular samples are currently being developed. The Proctor Foundation at the University California San Francisco leads efforts aimed at analysis of clinical samples using mNGS (https://proctor.ucsf.edu/Proctor-Lab).

Although mNGS holds significant promise in improving the diagnostic yield of intraocular fluid specimens in uveitis and intraocular infection, it is not without its own limits. Current methods are costly, time-intensive, and require specialized bioinformatics skills. Because metagenomics collects all genomic material present in the specimen, DNA sequences present may not represent causative pathogens, or even actual living organisms. Additionally, environmental and laboratory contaminants may be inadvertently sequenced and even amplified during the sequencing process. Although bioinformatic analysis attempts to account for and exclude these, the clinician must consider the possibility that results could represent such false positives when interpreting results. Directed PCR should be performed when the pretest probability for a specific organism is high, whereas mNGS results could represent such false positives when interpreting results. Directed PCR should be performed when the pretest probability for a specific organism is high, whereas mNGS results could represent such false positives when interpreting results.

Finally successful adoption of any of these techniques into clinical practice requires an awareness that these techniques have the potential to amplify irrelevant genomic material, resulting in data that can complicate or confuse clinical decision making.

Future developments that will make this technology more clinically applicable include streamlined and automated bioinformatics data processing, improved data storage, and reduced sequencing costs. Additionally, the ability to identify live organisms with certainty will require optimized collection and sample storage methods to preserve RNA integrity and expanded use of RNA sequencing methods.

**REFERENCES**


