Computational and Structural Biotechnology Journal 16 (2018) 108–120

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Mini Review

Highlighting Clinical Metagenomics for Enhanced Diagnostic Decision-making: A Step Towards Wider Implementation

Jessica D. Forbes [a,b,c], Natalie C. Knox [c], Christy-Lynn Peterson [c], Aleisha R. Reimer [c,*]

[a] Department of Internal Medicine, University of Manitoba, Winnipeg, Manitoba, Canada
[b] University of Manitoba IBD Clinical and Research Centre, Winnipeg, Manitoba, Canada
[c] National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB, Canada

ARTICLE INFO

Article history:
Received 7 November 2017
Received in revised form 19 February 2018
Accepted 25 February 2018
Available online 27 February 2018

Keywords:
Clinical metagenomics
Diagnosis
Clinical microbiology laboratory
Pathogen detection
Culture-independent diagnostic test
Next-generation sequencing

ABSTRACT

Clinical metagenomics (CMg) is the discipline that refers to the sequencing of all nucleic acid material present within a clinical specimen with the intent to recover clinically relevant microbial information. From a diagnostic perspective, next-generation sequencing (NGS) offers the ability to rapidly identify putative pathogens and predict their antimicrobial resistance profiles to optimize targeted treatment regimens. Since the introduction of metagenomics nearly a decade ago, numerous reports have described successful applications in an increasing variety of biological specimens, such as respiratory secretions, cerebrospinal fluid, stool, blood and tissue. Considerable advancements in sequencing and computational technologies in recent years have made CMg a promising tool in clinical microbiology laboratories. Moreover, costs per sample and turnaround time from specimen receipt to clinical management continue to decrease, making the prospect of CMg more feasible. Many difficulties, however, are associated with CMg and warrant further improvements such as the informatics infrastructure and analytical pipelines. Thus, the current review focuses on comprehensively assessing applications of CMg for diagnostic and subtyping purposes.

Contents

* Corresponding author at: National Microbiology Laboratory, Public Health Agency of Canada, 1015 Arlington Street, Winnipeg, MB R3E 3R2, Canada.
E-mail address: Aleisha.Reimer@canada.ca (A.R. Reimer).

https://doi.org/10.1016/j.csbj.2018.02.006

1. Introduction

Infectious diseases are a leading cause of morbidity and mortality worldwide. Recent estimates suggest that approximately 19% of global deaths are attributed to infectious diseases [1]. According to the World Health Organization, lower respiratory tract infections are at present, the most common communicable disease causing 3.2 million deaths in 2015; enteric disease and tuberculosis caused 1.4 million deaths each and HIV/AIDS was responsible for 1.1 million deaths [2]. The identification and characterization of pathogenic microorganisms including bacteria, viruses, parasites or fungi that cause infections are critical for the clinical management of patients and the prevention of transmission. In addition, novel, emergent and re-emergent pathogens such as MERS, Ebola, Zika, and the spread of multidrug-resistant pathogens further emphasize the importance of effective diagnostics.

Many syndromes are complicated by the capability of a wide array of pathogens to cause clinically indistinguishable diseases. As a result, accurate diagnosis often requires a battery of traditional microbiological methods such as culture, nucleic acid amplification tests (e.g. polymerase chain reaction; PCR) and serologic assays. Rapid developments have recently been made in the modernization of clinical microbiology laboratories with the employment of multiplex syndromic panels (e.g. BDMax, FilmArray and others), matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) and whole genome sequencing (WGS). These methods have played an increasingly important role in clinical microbiology laboratories due to their ability to reduce turnaround time (TAT) from specimen collection to clinically actionable result and through the detection of non-cultivable or fastidious pathogens. However, due to limitations of current diagnostic methodologies (reviewed in [3]) such as requiring a priori knowledge of the pathogen, missed diagnoses occur in 20–60% of cases dependent on the particular syndrome [4–7]. As a consequence, broad-spectrum antibiotics are generally empirically administered, obviating the use of targeted therapies and ultimately resulting in increased mortality along with excess healthcare-associated costs.

Recent and continuous improvements of next-generation sequencing (NGS) technology have effectively transformed biomedical research. The application of next-generation sequencing (NGS) approaches such as WGS in clinical microbiology laboratories is wide-ranging including for purposes of outbreak management, pathogen surveillance and subtyping and zoonotic transmission determination. Few laboratories are at present applying a culture-independent high-throughput sequencing approach for diagnostic purposes [8]. An NGS-based approach can offer a relatively unbiased pathogen detection through the use of bioinformatics methods and comprehensive reference databases (Fig. 1). From a clinical standpoint, the implementation of clinical metagenomics (CMg) appears to be promising in numerous disciplines including infectious diseases. Thus, CMg has the ability to function as a single assay that can be employed for diagnostic purposes, subtyping, antimicrobial resistance (AMR) detection and virulence profiling. Herein, we discuss the rapidly emerging field of CMg and provide a comprehensive review of NGS culture-independent diagnostic applications thereby describing the potential suitability of this diagnostic assay to be routinely implemented in frontline laboratories.

2. What Is Metagenomics?

Metagenomics has previously been used to evaluate the microbial community within a sample or environment, for example, interrogating the gut microbiome and its association with chronic diseases such as inflammatory bowel disease [9], obesity [10] and type 2 diabetes [11]. Analogous to the increased popularity of NGS and bioinformatics, metagenomics is progressively being applied as a novel infectious disease diagnostic assay.

There are two approaches that can be used to examine the microbiome of a given specimen environment, shotgun metagenomics and targeted-amplicon sequencing. Key differences between each approach are described elsewhere [3]. Briefly, shotgun metagenomics attempts to sequence the entire genetic content present in a sample whereas targeted-amplicon represents a more biased approach to a particular group of microorganisms.

Of the two methods, shotgun metagenomics is less taxonomically biased and capable of higher taxonomic resolution as it aims to amplify the whole genomes of every organism present in a specimen. As such, it allows for extended characterization of the microbial population, including subtypes, AMR and pathogenic gene carriage. For this reason, this method tends to lend itself to CMg diagnosis, as can be seen in the proportion of current CMg studies using shotgun metagenomics methodologies (Fig. 2B, Suppl. Table S1). However, there are many issues inherent to shotgun metagenomics; for example, overwhelming quantities of host DNA are often sequenced in comparison to the small fraction of microbial DNA, which is dependent on the biological specimen type. Thus, it can be difficult to obtain high sequence coverage for microbes of interest in specimens where host cells are abundant [12]. In addition, dependent on the sequencing depth required, shotgun metagenomics is significantly more costly than target-amplicon sequencing.

As previously stated, the use of targeted-amplicon sequencing, for pathogen detection is biased due to its inability to query microorganisms across multiple kingdoms (e.g. virus, eukaryotes and prokaryotes) in a biological specimen. In addition, this approach does not provide any additional characterization beyond phylogenetic information. The use of a ubiquitous and taxonomically informative universal genetic marker is used to capture the phylogenetic information of the targeted microorganisms in a given environment. The 16S ribosomal RNA (rRNA) gene, for example, represents the most common marker gene for bacteria and archaea [13]. The gene is ubiquitous in varying copy numbers [14] and provides sufficient sequence variability for taxonomic resolution [15]. Confident taxonomic assignment below the genus level is often difficult due to lack of resolution. Further, accurate classification below the family level of taxonomy has recently been questioned [16]. Several other marker genes with similar limitations to the 16S rRNA gene exist. These marker genes target different taxonomic groups such as the 18S rRNA gene for eukaryotes [17], internal transcribed spacer (ITS) region for fungi [18] and rpoB [19], cpn60 [20], 5S rRNA and 23S rRNA for bacteria and archaea [21,22]. While this method is less suited for an unbiased pathogen detection approach in comparison to shotgun metagenomics, some studies (discussed below) have applied it in a clinical diagnostic setting. Accordingly, we have provided a brief overview of its methods and analytical processes.