Journal of Microbiological Methods 95 (2013) 415–424

Contents lists available at ScienceDirect

Journal of Microbiological Methods

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

Review

Genomics and metagenomics in medical microbiology

Roshan Padmanabhan, Ajay Kumar Mishra, Didier Raoult, Pierre-Edouard Fournier *

Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Aix-Marseille Université, UM63, CNRS 7278, IRD 198, INSERM 1095, Faculté de Médecine, 27 Bd. Jean Moulin, 13005 Marseille, France

ARTICLE INFO

Article history:
Received 29 July 2013
Received in revised form 8 October 2013
Accepted 10 October 2013
Available online 5 November 2013

Keywords:
Next-generation sequencing
Genomics
Metagenomics
Clinical microbiology

ABSTRACT

Over the last two decades, sequencing tools have evolved from laborious time-consuming methodologies to real-time detection and deciphering of genomic DNA. Genome sequencing, especially using next generation sequencing (NGS) has revolutionized the landscape of microbiology and infectious disease. This deluge of sequencing data has not only enabled advances in fundamental biology but also helped improve diagnosis, typing of pathogen, virulence and antibiotic resistance detection, and development of new vaccines and culture media. In addition, NGS also enabled efficient analysis of complex human micro-floras, both commensal, and pathological, through metagenomic methods, thus helping the comprehension and management of human diseases such as obesity. This review summarizes technological advances in genomics and metagenomics relevant to the field of medical microbiology.

© 2013 Elsevier B.V. All rights reserved.

Contents

1. Introduction . . . 415
2. Next generation DNA sequencing platforms . . . 416
   2.1. Roche/454 GS . . . 416
   2.2. Illumina/Solexa . . . 416
   2.3. Life Technologies SOLiD . . . 416
   2.4. Benchtop sequencers . . . 417
   2.5. Single-cell sequencing . . . 417
   2.6. Higher output, new challenges . . . 417
3. Application of genomics in medical microbiology . . . 417
   3.1. Design of polymerase chain reaction (PCR) assays . . . 417
   3.2. Genotyping . . . 418
      3.2.1. Non-sequence based genotyping methods . . . 418
      3.2.2. Sequence-based genotyping methods . . . 418
   3.3. Reverse vaccinology . . . 419
   3.4. Genomics-based design of culture media . . . 419
   3.5. Detection of virulence factors . . . 419
   3.6. Detection of antibiotic resistance . . . 419
   3.7. Taxonomy . . . 419
   3.8. Role of WGS in epidemiology . . . 420
   3.9. Pangenomics . . . 420
4. From genomics to metagenomics — a culture-independent technique to decipher complex floras . . . 420
   4.1. Metagenomic studies of the human microbiome . . . 420
5. Next generation sequencing and its future prospects in medical microbiology . . . 421
References . . . 421

1. Introduction

Infectious diseases are the predominant cause of human mortality worldwide. Humans are susceptible to infections caused by a wide

* Corresponding author. Tel.: +33 491 385 517; fax: +33 491 387 772.
E-mail address: pierre-edouard.fournier@univ-amu.fr (P.-E. Fournier).

0167-7012/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.mimet.2013.10.006

variety of microbial pathogens. Hence, an accurate diagnosis of the causative agent is one of the most crucial steps of disease management. The gold standard for diagnosis in the clinical microbiology laboratory has long been culture. This process includes the isolation, identification and classification of the organism which is followed by the generation of a patient report. This diagnostic strategy takes a minimum of 18 h, followed by further subculture for identification and susceptibility testing. In the 1980s, the introduction of techniques enabling replication of DNA, notably polymerase chain reaction (PCR) (Mullis and Faloona, 1987), has resulted in the development of a multitude of diagnostic tools that helped in improving the diagnosis and characterization of infectious diseases. In 1995, the sequencing of the first bacterial genome, that of *Haemophilus influenzae* (Fleischmann et al., 1995), was a major step forward in microbiology by demonstrating the usefulness of genomics to unveil the full genomic content of a bacterium (Li et al., 2009). Medical applications of microbial genomics, strengthened by the rational, rather than empirical, design of most molecular assays, had a major impact on diagnosis and prevention of infectious diseases, with traditional culture-based diagnostic techniques being partially superseded by molecular detection methods such as PCR, real-time PCR, molecular hybridization-based technologies, and fingerprinting which enabled detection and identification of pathogens directly within specimens without the need for culture (Logares et al., 2012).

In the 2000s, high throughput sequencing (HTS), also named next-generation sequencing (NGS), enabled obtaining thousands to millions of reads per run at decreasing costs but raised new challenges, notably for sequence analysis (Didelot et al., 2012). Among the applications of NGS to medical microbiology, metagenomics i.e., the ability to identify all DNA populations within a given specimen, has been one of the most spectacular, especially when applied to complex human floras (Armougom and Raoult, 2008; Qin et al., 2010; Armougom et al., 2009; Lazarevic et al., 2009). In this article, we review the most relevant equipments and techniques used in genomic and metagenomic strategies, with special reference to the field of medical microbiology and how genomics is propelling diagnostic techniques from an empirical to a rational era.

2. Next generation DNA sequencing platforms

Developed by Frederick Sanger in 1977, chain termination-based DNA sequencing, now known as Sanger sequencing, was the primary technology in the first generation of commercial sequencing applications. Applied Biosystems introduced the first commercial DNA sequencer utilizing a slab gel electrophoresis, ABI Prism 310, which was later replaced by ABI Prism 3700 with 96-well capillaries, which made the sequencing faster and more accurate. These capillary sequencing machines and their associated software for assembly became the main tools for human genome project completion in 2001. Even though Sanger-based sequencers have long been considered to be gold standards, transition to next generation sequencing (NGS) methodologies had a major impact on several areas of medicine. A major scientific breakthrough in the field of genomics and microbiology came from the sequencing of *H. influenzae* in 1995 which took 13 months to complete (Török and Peacock, 2012). Since then, over 7000 complete bacterial genome sequences have been published [http://www.genomesonline.org]. Three major NGS platforms have emerged post human genome project, including the 454/Roche, Illumina/Solexa and SOLiD platforms. A brief and comprehensive history, and technical comparison and their software requirements of these and newer sequencing technology can be found in Liu et al. (2012) and Pareek et al. (2011).

The workflow of HTS can be summarized as follows: the first step is the production of a library of fragments of the DNA population present within a target specimen. These library fragments are then clonally amplified to get a sufficient signal for detection by the instrument's optical or electronic system. Finally, HTS sequencers perform sequencing reactions using various methods such as pyrosequencing, reversible dye terminator sequencing-by-synthesis (SBS), sequencing by chained ligation, sequencing by unchained ligation or ion-sensitive SBS. Each methodology has its own strengths and weaknesses. Depending on the library construction approach and the type of sequencing platform, HTS reactions can be classified as shotgun sequencing, in which each DNA fragment is sequenced individually; paired-end sequencing, where both ends of linear fragments are tagged and sequenced; or mate pair sequencing during which both ends of previously circularized fragments are sequenced. The latter two methods facilitate assembly of sequenced reads. Here we very briefly describe the major NGS platforms used in medical microbiology and some of the emerging newer generation single molecule sequencing methods.

2.1. Roche/454 GS

Roche/454 sequencers use the pyrosequencing technology in which pyrophosphates are generated while the DNA polymerase adds nucleotides to the template DNA. Roche/454 platforms generate approximately one million reads with a minimum output of 35 Mb and a read size of up to 1000 bps However, this technology suffers from a few drawbacks including the impossibility to decipher homopolymeric regions of more than 5 nucleotides. The Roche/454 platforms are widely used for genomic sequencing and metagenomic studies. In a recent study of the role of colonic mucosa-associated *Escherichia coli* in Crohn's disease, inflammatory bowel disease (IBD) and colorectal cancer, bacterial genes relevant to pathogenesis were identified using the GS-FLX Titanium (Prorok-Hamon et al., 2013). GS-FLX sequencing was also successfully employed in a metagenomic study aiming at identifying the numerous pathogens involved in brain or liver abscess, and pleuritis (Sibley et al., 2012). Another study using the benchtop GS Junior and GS-FLX Titanium evaluated the dysbiosis of the salivary microbiota associated with inflammatory responses in IBD patients. The results showed that *Bacteroidetes* were significantly increased with a concurrent decrease in *Proteobacteria* of IBD patients (Said et al., 2013).

2.2. Illumina/Solexa

Illumina/Solexa sequencers employ the sequencing-by-synthesis method using reversible dye termination nucleotides. Illumina's HiSeq 2500 sequencer has an output of 600 Gb that compensates its short read length (2×100 bps). The Illumina Genome Analyzer on the other hand produces read lengths of 35 bp with a >99% raw base accuracy and an overall throughput of approximately 5 Gb over a three day run (Anderson and Schrijver, 2010). Seth-Smith and colleagues illustrated a viable method for whole bacterial genome sequencing directly from clinical samples without the need for culture using multiplexed Illumina Genome Analyzer II and HiSeq sequencers. These authors sequenced *Chlamydia trachomatis*, an obligate intracellular pathogen causing sexually transmitted infections and the blinding disease trachoma in humans, from discarded clinical swabs. Their methodology included the immunomagnetic separation (IMS) for targeted bacterial enrichment with multiple displacement amplification (MDA) followed by whole-genome amplification. They showed that this approach (IMS–MDA) is useful in microbes which are difficult to culture or fastidious bacteria of clinical concern (Seth-Smith et al., 2013).

2.3. Life Technologies SOLiD

SOLiD (sequencing by oligonucleotide ligation and detection), one of the first HTS sequencing platforms to be developed, relies on sequencing by chain ligation. This platform generates an output of 180 Gb per run but suffers from a short read length (35–75 bp) (Stranneheim and Lundeberg, 2012; Ståhl and Lundeberg, 2012) that limited its use to a few applications such as whole genome re-sequencing, targeted re-sequencing, genotyping (Engelthaler et al., 2011), RNAseq analysis (Ozsolak and Milos, 2011; Van Vliet, 2010) and ChIP-Seq (Schmidt

Download English Version:

https://daneshyari.com/en/article/2090116

Download Persian Version:

https://daneshyari.com/article/2090116

Daneshyari.com