Journal of Clinical Virology 58 (2013) 346–350

Contents lists available at ScienceDirect

Journal of Clinical Virology

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

Review

Next-generation sequencing technologies in diagnostic virology

Luisa Barzon*, Enrico Lavezzo, Giulia Costanzi, Elisa Franchin, Stefano Toppo, Giorgio Palù

Department of Molecular Medicine, University of Padova, Via A. Gabelli 63, I-35121 Padova, Italy

ARTICLE INFO

Article history:
Received 24 November 2012
Received in revised form 1 March 2013
Accepted 2 March 2013

Keywords:
Genome sequencing
Bioinformatics
Deep/high throughput sequencing
Virus discovery
Viral metagenomices/viromics
Viral epidemiology

ABSTRACT

The data deluge produced by next-generation sequencing (NGS) technologies is an appealing feature for clinical virologists that are involved in the diagnosis of emerging viral infections, molecular epidemiology of viral pathogens, drug-resistance testing, and also like to do some basic and clinical research. Indeed, NGS platforms are being implemented in many clinical and research laboratories, as the costs of these platforms are progressively decreasing. We provide here some suggestions for virologists who are planning to implement a NGS platform in their clinical laboratory and an overview on the potential applications of these technologies in diagnostic virology.

Contents

1. Introduction

Novel DNA sequencing techniques, referred to as "next-generation" sequencing (NGS), provide high speed and throughput that can produce an enormous volume of sequences in a single run at relatively low cost. The most important advantage of these platforms is the ability to determine the sequence data from single DNA fragments of a library, avoiding the need for cloning in vectors prior to sequence acquisition. These techniques have provided

Abbreviations: NGS, next generation sequencing; HIV, human immunodeficiency virus; HBV, hepatitis B virus; HCV, hepatitis C virus; EBV, Epstein-Barr virus; HCMV, human cytomegalovirus; HSV, herpes simplex virus.

* Corresponding author. Tel.: +39 0498218946; fax: +39 0498272355.
E-mail address: luisa.barzon@unipd.it (L. Barzon).

1386-6532/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jcv.2013.03.003

a great contribution to the research in many fields of life sciences[1] and are being increasingly introduced in clinical laboratories, with many diagnostic applications in human genetics, oncology, as well as in microbiology and virology.[2,3] We provide here some suggestions for virologists who are planning to implement a NGS platform in their clinical laboratory and an overview on the potential applications of these technologies in diagnostic virology.

2. Implementation of NGS technology in the clinical virology laboratory

The data deluge produced by NGS technologies is an appealing feature for clinical virologists that are involved in the diagnosis of emerging viral infections, molecular epidemiology of viral pathogens, drug-resistance testing, and also like to do some basic and clinical research. Indeed, NGS platforms are being implemented in many clinical and research laboratories, as the costs of these platforms are progressively decreasing.

Different NGS platforms are available in the market, which are best suited for some applications but not for others. Thus, before acquiring a NGS platform, a careful analysis of the diagnostic and research needs of the laboratory should be done and expertise in NGS laboratory protocol set-up and NGS data assembly and analysis should be achieved. To this aim, we strongly recommend to establish a "wet lab team" and a "bioinformatics team" that collaborate to provide answers to the laboratory needs, before getting the NGS platform. Requirements of the "wet lab team" are skills in molecular biology techniques and biotechnologies, while requirements for the "bioinformatics team" are expertise in data management, implementation of sequence alignment algorithms, design of custom working pipelines and statistical analysis.

In the settlement of instruments and protocol workflow for NGS, care should be taken to separate pre- and post-PCR phases and to avoid the risk of nucleic acid contamination (e.g., during DNA nebulisation) also with the other conventional molecular diagnostic methods in the laboratory. Most important, in order to be used in routine testing, diagnostic tests performed by NGS technologies require analytical and clinical validation according to current guidelines and recommendations that are used for molecular assays and must be under the supervision of quality assurance and quality control programmes.

3. Choice of the NGS platform

Different NGS methods are commercially available and novel and improved platforms are continuously being developed and released. These NGS methods have different underlying biochemistries[4] and differ in sequencing protocol, throughput, and sequence length.[2] Thus, the SOLiD system (Life Technologies), characterised by extremely high throughput but very short reads, may be more suitable for applications such as large whole genome resequencing or RNA-sequencing projects; while other platforms, like 454 (Roche Diagnostics), Ion Torrent (Life Technologies), and Illumina sequencing systems (Illumina) provide data suitable for de novo assembly, even though the relatively limited throughput of 454 and Ion Torrent PGM restricts their application to small bacterial- and viral-size genomes. In contrast, the relative long length of 454 FLX (and its smaller version GS Junior) reads allows deep sequencing of larger amplicons, with applications in microbial and viral metagenomics, analysis of viral quasispecies, and viral haplotypes reconstruction.[5]

Besides these platforms, which are now consolidated and spread in laboratories all around the world, new instruments have been recently launched (PacBio RS from Pacific Biosciences)[6] or will be released in the near future (GridION and MinION from Oxford Nanopore Technologies),[7,8] the so called "third generation" technologies. These systems will probably take sequencing applications to a next level of performance, since they claim to be able to sequence single DNA molecules and reach read lengths above 10,000 bp, but a settling time will be needed before routine application in diagnostics.

The choice of the NGS platform for diagnostic virology should take into consideration the following aspects:

- *Diagnostic applications*: The types of diagnostic applications and the number of different tests to set up are relevant in the choice of a NGS platform. For a single application (e.g., deep sequencing of amplicons), small and relatively inexpensive instruments, like GS Junior and Ion Torrent PGM, are available. If the type and number of tests are not defined a priori and a flexible platform is required, Illumina sequencing systems provide several advantages, especially in terms of throughput, that could respond to most diagnostic needs, though reads are shorter than those from 454 FLX.
- *Costs*: The costs of tests by using NGS methods could be much higher or lower than those of conventional molecular methods and this should be taken into account. E.g., resequencing a full viral genome by conventional cycle sequencing is generally less expensive than by using NGS methods. Likewise, detection of a known single nucleotide mutation by real-time PCR is less expensive than by using deep sequencing of amplicons.
- *Speed*: The turnaround time of NGS protocols and data analysis is a critical issue if the NGS platform is to be used in "real-time" diagnostics, aiming to provide useful information for disease prevention or therapeutic interventions. Thus, for most diagnostic applications, the laboratory staff and/or the vendor should develop simplified laboratory protocols for library preparation and bioinformatics tools for easy data analysis and interpretation.
- *Throughput*: Different applications require different throughputs (e.g., shotgun metagenomics applications for pathogen detection require a much higher throughput than targeted resequencing applications).
- *Accuracy*: While a very high accuracy is needed for mutation detection (e.g., in drug resistance testing), this issue is less relevant for pathogen detection by shotgun approaches.
- *Read length*: Long reads are needed for deep sequencing of amplicons and de novo sequencing of genomes, and are also useful for pathogen discovery.
- *Upgrading*: NGS technologies are rapidly improving and evolving, thus upgradable instruments should be preferred.
- *Automation*: Since NGS protocol workflow is generally very long and complex, automation of library preparation protocols is strongly recommended.

4. Pitfalls of NGS in diagnostic virology

NGS technologies offer great chances for diagnostic virology, but there are several pitfalls in comparison with applications in diagnostic microbiology and other diagnostic applications.

- *Low sensitivity*: The sensitivity of NGS methods for the detection of viral sequences is relatively low due to the small size of the viral genome in comparison with host genome, that invariably contaminates clinical specimens or viral cultures. This is particularly relevant when viral load is very low. Thus, for viral genome detection, conventional PCR and cycle sequencing may be more convenient than NGS sequencing.
- *Whole viral genome sequencing*: The presence of contaminating host genome sequences is also a problem for whole viral genome sequencing. To enrich viral particles in the sample, contaminating