

Review Article

Next-generation sequencing in clinical virology: A review of its applications

Samea Khan¹, Mirza Jawad ul Hasnain², Muhammad Tariq Pervez²,
Tanveer Hussain¹ and Khizra Maqsood^{2*}

1. Department of Molecular Biology, Virtual University of Pakistan, Lahore-Pakistan

2. Department of Bioinformatics and Computational Biology, Virtual University of Pakistan, Lahore-Pakistan

*Corresponding author's email: khizra.maqsood@vu.edu.pk

Citation

Samea Khan, Mirza Jawad ul Hasnain, Muhammad Tariq Pervez, Tanveer Hussain and Khizra Maqsood. Next-generation sequencing in clinical virology: A review of its applications. Pure and Applied Biology. Vol. 9, Issue 4, pp2356-2360. <http://dx.doi.org/10.19045/bspab.2020.90250>

Received: 02/03/2020

Revised: 16/06/2020

Accepted: 18/07/2020

Online First: 20/07/2020

Abstract

Next-generation sequencing (NGS) has emerged as the most useful technique for wide-scale detection of infection-causing pathogens. Owing to the increasing demand for cost-effective and quick DNA sequencing protocols, many research groups have reported the use of NGS for lab research and experimentation. NGS has revolutionized the field of DNA sequencing by helping sequence big data and overcomes the limitations associated with Sangers sequencing. In the current manuscript, we have aimed to review the feasibility of using NGS in public and clinical health laboratories for disease diagnostics. We believe that this review will add to the available literature by highlighting the important aspects of NGS in the characterization of various pathogens which cannot be grown using culture media and other conventional approaches.

Keywords: Biotechnology; Clinical application; Next-Generation Sequencing; Virology; Whole viral genome.

Introduction

Viruses are the main cause of several human diseases specifically in developing countries. A lot of research has been done for detection, classification, and characterization of the viruses, advances in the field have helped in the prevention and treatment of many viral infections. In older times, the cell culture technique was used for viral identification. With the advent of an economical and quick technique of Next Generation Sequencing (NGS) revolutionized the filed of viral identification and characterization. Once sequencing has been done, known viruses can be identified using publicly available sequence repositories using sequence alignment strategies [1].

Viruses are highly diverse and over the past decades, there has been a paradigm shift in viral epidemiology. This shift in epidemiology is mostly attributed to anthropological activities that have led to the emergence of new viruses as well as the relocation of known notorious pathogens to new ecological niches [2, 4]. Owing to the extremely diverse nature of the viral pathogens, computational challenges pose a serious hurdle in viral identification. This is because viruses do not have conserved domains such as the 16S ribosomal RNA of the bacteria. This tremendous diversity makes the identification of novel viruses very difficult due to constraints in isolation, amplification, and purification protocols [5].

Advancements in genome sequencing technologies have led to the opening of new avenues for research and discovery in clinical virology. The main advantage of using NGS is that it allows massive parallel sequencing at cost-effective rates. However, due to the availability of many diverse NGS platforms, it is important to choose the right one. Identification of novel pathogens calls for large databases and improved computing power for longer reads and higher coverage. Once reference genomes are accessible for assembly, short-read producing technology can be used to cover metagenomes [6, 7].

Earlier the use of the NGS platforms was not appropriate because of its fast readings for virus exploration or de novo sequence projects. Nevertheless, a frequent improvement in real-time for Illumina platforms rendered the assembly of genomes to de novo at the level of sensitivity equal to a specific Polymerase Chain Reaction (PCR). Pyro sequencing technology (Roche 454) is ideally used in comparison to the other NGS viewpoints that produce the shorter reads, in particular for metagenomes studies [6]. In this review will we provide an overview of next-generation sequencing technologies role in the field of clinical virology especially in viruses' discovery, and reconstruction of the viral genome.

Impact of next-generation sequencing in clinical virology

Next-generation high throughput sequencing technology has made it possible to directly identify, detect, and then discover novel viruses in a biased manner in several animals and plants without the presence of antibodies or prior virus sequences knowledge [8].

Virus discovery

A major purpose to identify novel viruses is to cure the diseases associated with them as seen in the case of the latest coronavirus infectious disease (COVID-19) pandemic. For this purpose, the specimen used is a mix of healthy and viral sequences [1]. The rapid

progress of NGS has helped the technology to evolve independently of the conventionally used sequencing platforms [9, 10]. Moreover, this technology works well with big data and provides accurate and sensitive results in a small amount of time. The technology also uses fascinating features, including viral evolution research and quasi-species identification, antiviral resistance vaccines, etc. Discovery of the novel coronavirus is also a gift of the NGS, apart from many other novel pathogens discovered in the last decade including clusters arenavirus, Ebola virus, Severe Acute Respiratory Syndrome (SARS) Virus, Human Immunodeficiency Virus (HIV), Corona Virus and identification of viral etiology of an outbreak of a disease in honeybees [11, 13].

Reconstruction of whole viral genome

Reconstruction of the whole viral genome is an important application of NGS. This technology helps redesign the entire sequence even when the genome is poorly characterizing or unknown. The process starts either from a clinical sample or culture-enriched viral preparation [13]. The assay used in the reconstruction of the viral genome is entirely different from the shotgun metagenomic sequencing, to random shotgun sequencing and overlapping amplicon sequencing. By using this approach, the full genome of multiple viruses was obtaining that include HIV, pandemic influenza virus, human herpesviruses, and many more [14,15]. The Sequence independent amplification of nucleic acids that are combined with NGS technology and bioinformatics analyses is an encouraging strategy for recognizing pathogens in public and clinical health positions [16].

NGS in clinical virology

NGS technologies expanding the capacity of researchers to analyze viral community from a heterogeneous mixture of habitats and different environments that help in the

discovery of new viruses for example novel arenaviruses that are involved in the transplant-associated disease clusters [1]. The whole viral genome sequencing is a common application of NGS stating either directly from clinical samples or culture-enriched viral preparations. Examples of NGS genome sequencing viruses are influenza virus, HIV, human herpesviruses, and many more [2]. Characterization of viral community is another most common application of NGS.

The technology holds huge promise for clinical testing of infectious disease-causing agents. However, several challenges such as the automation processes, standardizing technical protocols and bioinformatics pipelines, improving reference databases, establishing the skilled quality control measures and testing, and decreasing the cost and turn-around time still need to be overcome [17]. The latest research has highlighted the significance of NGS for molecular characterization, spanning from metagenomics characterization of the unknown pathogens to the epidemiology and the development of viral quasi-species [18]. Whole-genome sequencing (WGS) of clinical bacterial isolates has been demonstrated to notify hospital infection-prevention practices [19]. Viruses are identified as major disease-causing agents; however, pathogenic viruses are not characterized in most patients. Currently, NGS has been implemented to completely detect pathogens of many infectious diseases of unrevealed etiology [20].

Using ideal experimental design NGS based drug resistance testing at a big scale is practicable, supplying finer characterization and the understanding of the evolving viruses and possibly making better the predictive value of the drug resistance testing [21]. The effect of NGS on microbiology has been revolt, with new sequences of microbial genomic being created daily, leading to the

progression of huge databases of genes and genome sequences [22].

Many RNA viruses possess small genomes as compared to other organisms and would become an evident success story for the utilization of NGS technologies [23]. NGS showed the fine polymorphic sub-structure of the virus species, from nucleotide variants with only less than 50% frequency to fractions of 1%. NGS enables the very delicate measurement of complete transcriptomes [24].

Just like every latest technology that is adopted in microbiology, the incorporation of NGS into routine and clinical workflows must be managed very carefully. Microbiome approaches based on NGS, express a major new availability to inspect the potential job of infection-causing pathogens in the pathogenesis of numerous neuro-inflammatory diseases [24].

NGS supplies a practical approach to the complete-genome sequencing of HCV that detects low-frequency variants and permits examination of virus genetic diversity within the sample. Such devices can examine and sequence in a rapidly parallel way a complex mix of nucleic acid plants, rendering them excellent instruments for viral metagenomics [24].

In clinical studies, NGS has been an efficient tool for the identification and discovery of new infectious agents. Nevertheless, as the NGS method is unbiased, very small genome-to-genome ratios of clinical samples are highly influenced by the viruses. Thus, it is necessary to increase the response to the discovery of new diseases that contain viruses in the clinical samples to the improvement of pathogenic genetic material or the depletion of the host genetic material.

Conclusion

In this current article, the importance of next-generation sequencing has been elaborated within the field of clinical virology that how

beneficially it works in this department. It helps in the detection of harmful pathogens and open ways for its advanced treatments. NGS possesses great value for virology. The supreme benefit of NGS is that it possesses the ability to characterize numerous different pathogens that are not cultivable utilizing the conventional approaches. So, next-generation sequencing plays a versatile role in clinical virology. Despite the multiple application of NGS in virus discovery, the high sensitivity of this technique also makes it prone to unintentional contamination. To avoid contamination one approach might be the simultaneous analysis of the primers and controls used for deep sequencing and enrichment. Equipment high cost, a requirement of high computational power to analyze large sequences dataset and the needs of highly skilled personnel with proper expertise are the other limiting factors of NGS. However, NGS is a relatively recent technology, and these issues must be addressed before it becomes fully exploitable at the clinical laboratory level.

Authors' contributions

Conceived and designed the experiments: MT Pervez & T Hussain, Performed the experiments: S Khan, MJU Hasnain & K Maqsood.

References

- Zhang J, Chiodini R, Badr A & Zhang G (2011). The impact of next-generation sequencing on genomics. *J Genet Genomics* 38: 95–109.
- Szpara ML, Parsons L & Enquist LW (2010). Sequence Variability in Clinical and Laboratory Isolates of Herpes Simplex Virus 1 Reveals New Mutations. *J Virol* 84: 5303–5313.
- Cantalupo PG & Pipas JM (2019). Detecting viral sequences in NGS data. *Curr Opin Virol* 39: 41–48.
- Gondard M, Temmam S, Devillers E, Pinarello V, Bigot T, Chrétien D, Aprelon R, Vayssier-Taussat M, Albina E & Eloit M (2020). RNA Viruses of *Amblyomma variegatum* and *Rhipicephalus microplus* and Cattle Susceptibility in the French Antilles. *Viruses* 12.
- O'Callaghan-Gordo C, & Antó J (2020). COVID-19: The disease of the anthropocene. *Environ Res* 187: 109683.
- Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, de Bourcy C, Egger R, Dimitrov B, Juan YF & Sheu J (2019). Metagenomic next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PloS One* 14: e0218318.
- Aguilera-Diaz A, Vazquez I, Ariceta B, Mañú A, Blasco-Iturri Z, Palomino-Echeverría S, Larrayoz MJ, García-Sanz R, Prieto-Cond MI, Chillón M & del C (2020). Assessment of the clinical utility of four NGS panels in myeloid malignancies. Suggestions for NGS panel choice or design. *PloS One* 15: e0227986.
- Cremer J, Morley U, Pas S, Wolthers K, Vennema H, Duizer E & Benschop K (2019). Highly sensitive parechovirus CODEHOP PCR amplification of the complete VP1 gene for typing directly from clinical specimens and correct typing based on phylogenetic clustering. *J Med Microbiol* 68: 1194–1203.
- Gaspareto KV, Ribeiro RM, de Mello Malta F, Gomes-Gouvêa MS, Muto NH, Mendes-Correa MC, Rozanski A, Carrilho FJ, Sabino EC & Pinho JRR (2016). HCV inter-subtype 1a/1b recombinant detected by complete-genome next-generation sequencing. *Arch Virol* 161: 2161–2168.
- Chang ST, Sova P, Peng X, Weiss J, Law GL, Palermo RE & Katze MG (2011). Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and regulation of noncoding RNA expression in a CD4+ T cell line. *mBio* 2.
- Zhang Y, Qi Y, Lo ECM, McGrath C, Mei ML & Dai R (2020). Using next-generation sequencing to detect oral microbiome change following periodontal interventions: A systematic review. *Oral Dis.* n/a.

12. Palacios G, Druce J, Du L, Tran T, Birch C, Briese T, Conlan S, Quan PL, Hui J & Marshall J (2008). A new arenavirus in a cluster of fatal transplant-associated diseases. *N Engl J Med* 358: 991–998.
13. Towner JS, Sealy TK, Khristova ML, Albariño CG, Conlan S, Reeder SA, Quan PL, Lipkin WI, Downing R & Tappero JW (2008). Newly Discovered Ebola Virus Associated with Hemorrhagic Fever Outbreak in Uganda. *Plos Pathog* 4: e1000212.
14. Bibby K, Viau E, & Peccia J (2011). Viral metagenome analysis to guide human pathogen monitoring in environmental samples. *Lett. Appl Microbiol* 52: 386–392.
15. Willerth SM, Pedro HAM, Pachter L, Humeau LM, Arkin AP, & Schaffer DV (2010). Development of a Low Bias Method for Characterizing Viral Populations Using Next Generation Sequencing Technology. *Plos One* 5: e13564.
16. Kwok H, Tong AHY, Lin CH, Lok S, Farrell PJ, Kwong DLW & Chiang AKS (2012). Genomic Sequencing and Comparative Analysis of Epstein-Barr Virus Genome Isolated from Primary Nasopharyngeal Carcinoma Biopsy. *Plos One* 7: e36939.
17. Prachayangprecha S, Schapendonk CME, Koopmans MP, Osterhaus ADME, Schürch AC, Pas SD, van der Eijk AA, Poovorawan Y, Haagmans BL & Smits SL (2014). Exploring the potential of next-generation sequencing in detection of respiratory viruses. *J Clin Microbiol* 52: 3722–3730.
18. Lefterova MI, Suarez CJ, Banaei N & Pinsky BA (2015). Next-Generation Sequencing for Infectious Disease Diagnosis and Management: A Report of the Association for Molecular Pathology. *J Mol Diagn JMD* 17: 623–634.
19. Van Borm S, Belák S, Freimanis G, Fusaro A, Granberg F, Höper D, King DP, Monne I, Orton R & Rosseel T (2015). Next-generation sequencing in veterinary medicine: how can the massive amount of information arising from high-throughput technologies improve diagnosis, control, and management of infectious diseases? *Methods Mol Biol Clifton NJ* 1247: 415–436.
20. Greninger AL, Zerr DM, Qin X, Adler AL, Sampoleo R, Kuypers JM, Englund JA & Jerome KR (2017). Rapid Metagenomic Next-Generation Sequencing during an Investigation of Hospital-Acquired Human Parainfluenza Virus 3 Infections. *J Clin Microbiol* 55: 177–182.
21. Goldberg B, Sichtig H, Geyer C, Ledebor N & Weinstock GM (2015). Making the Leap from Research Laboratory to Clinic: Challenges and Opportunities for Next-Generation Sequencing in Infectious Disease Diagnostics. *mBio* 6.
22. Van Laethem K, Theys K & Vandamme AM (2015). HIV-1 genotypic drug resistance testing: digging deep, reaching wide? *Curr Opin Virol* 14: 16–23.
23. Next generation sequencing of viral RNA genomes
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3708773/>.
24. Suzuki T, Kawada, JI, Okuno Y, Hayano S, Horiba K, Torii Y, Takahashi Y, Umetsu S, Sogo T & Inui A (2017). Comprehensive detection of viruses in pediatric patients with acute liver failure using next-generation sequencing. *J Clin Virol* 96: 67–72.