How to Cite:
Yadav, S., Wagh, V., Pajai, S., Joseph, A. S. M., Jaganathan, R., & Banerjee, D. (2022). Detection of respiratory viruses in clinical samples: Exploring next generation sequencing and clinical metagenomics. *International Journal of Health Sciences*, *6*(S6), 8264–8273. https://doi.org/10.53730/ijhs.v6nS6.12062

Detection of respiratory viruses in clinical samples: Exploring next generation sequencing and clinical metagenomics

Shubham Yadav
Department of Microbiology, Jawaharlal Nehru Medical College, Belagavi (KLE Deemed to be University)
Email: yadavshubham835@gmail.com

Vaishnav Wagh
Department of Multidisciplinary Research Unit, AFMC, Pune
Email: vaishnavwagh04@gmail.com

Sakshi Pajai
Department of Biochemistry, AIIMS Nagpur
Email: Sakshipajai27@gmail.com

Anto Simon Joseph M.
Department of Biotechnology, Sri Krishna Arts and Science College, Coimbatore
Email: simonbscbt@gmail.com

Ravindran Jaganathan
Preclinical Department, Royal College of Medicine Perak, University of Kuala Lumpur
Email: jravimicro@gmail.com

Durba Banerjee
Department of Biotechnology, University of Gautam Buddha
Email: banerjeedurba24@gmail.com

Abstract---Respiratory viral infections accounts for more than 4.25 million fatalities each year, making them the third most common cause of death worldwide. Although the majority of acute respiratory infections are assumed to be caused by viruses, the exact cause is often unknown. Comprehensive mapping of viral genomic sequencing is done, which also shows a significant degree of viral heterogeneity that contributes to the early diagnosis of respiratory illness. The development of next-generation sequencing (NGS), particularly for the detection of unidentified respiratory viruses, has revolutionized the field of novel viral genome detection. This review focusses on different

Manuscript submitted: 9 April 2022, Manuscript revised: 18 June 2022, Accepted for publication: 27 July 2022

models of sequencing techniques available for novel viral genome detection. Although there are still major technical and ethical issues in using these technologies for clinical detection, this technology has great promise for the future as a way to better understand respiratory viruses and make diagnoses that are more precise.

Keywords---comprehensive mapping, next-generation sequencing (NGS), genomic sequencing.

Introduction

Respiratory viral infections (RVIs) accounts for high fatality rates in humans throughout the worldresponsible for causing millions of hospitalizationsper year globally [1]. Overall, this type of airway illness accounts for lower epidemic peaks and higher mortality rate among all age group of children. Furthermore, the outbreak of emerging viruses has increased over past decades including SARS-CoV, human IFV (H5N1, H7N9) [2]. Thus, the outbreak of this severe illness has made researchers to focus mainly on the diagnosis and classification of respiratory virus which are: (URTIs) (e.g., rhinitis, pharyngitis, and laryngitis) and LRTIs (e.g., bronchitis, bronchiolitis, and pneumonia). The co-pathogenesis of this type of coinfections generally involves effective mechanism of host immune response towards invading pathogens. The virulence of the microbe inside the human body affects the normal homeostasis [3]. The effect of the airways infections mainly depends on the virulence of viral strain and host mechanism. The underlying cause for RVIs remains unknown as theviral pathogens have received little attention till date [4]. Although the severity of disease is high, only few preventiveinterventions are currently available. As for now the currently identified viruses does not responsible for causing illness, there are many lower respiratory ailment agents that are yet to be identified using systematic approaches [5].

Standard laboratory methods currently involved in clinical detection includes NAAT, DFA, RADT and RT-PCR techniques.The main drawback of all the available techniques is that, only few targets are available for genome sequencing which results in atypical or novel virus detection[6]. Moreover, the use of sensitive array techniques still does not identify the exact causative agent for the viral infection. This will depend upon the sample time, performance of the individual assays, and novel viral genomes are not detected because of large specificity in assays for detection. New advancement in viral genomic studies paved the way for sequence-independent amplification techniques combined with next-generation sequencing platforms. Sequence-independent NGS approaches are widely used in the field of virology for clinical diagnosis of this pathogenic viral infection. This technique is highly recognized because of their large sequencing of data at a relatively low duration, fast turnover with complete genomic information. For novel pathogen discovery without knowing the complete sequence of the viral pathogen metagenomic sequencing method is highly recommended [7]. The main advantage of using metagenomics sequencing is that novel genomic sequencing is possible without requiring the knowledge on the viral genome; sequencing of many individual genome is possible at a time [8].

Recent review in clinical detection using bioinformatics methods only discusses about the application, method of sequencing techniques but they failed to discuss about the application of these techniques in novel viral pathogen detection. The main objective of the review are (i) to explain the role of NGS in identifying mutated DNA patterns of virus infected patients (ii) describe the contribution of these sequencing techniques in the detecting viral infections using several case studies (iii) bioinformatics challenges in using these techniques for clinical detection

Next Generation Sequencing

Table 1
Summary of currently available NGS techniques with their advantages and disadvantages [4]

Method	Amplification Method	Sequencing technique	Advantages	Disadvantages
Roche 454	Emulsion PCR	Pyrosequencing	Longer sequencing read length, highly flexible and parallel sequencing can be done	Only read lengths of about ~100-200bp obtained so far
SOLiD	Emulsion PCR	Ligation	Shows higher accuracy for sequences which possess coverage of around 30X	Requires higher instrumental cost; short reads are obtained
Illumina	Bridge amplification *in situ*	Reversible terminators	Error free reads, enables higher base calling of genome	Multiple days required to get result
Helicos	No amplification	Reversible terminators	Provides exact quantification of sequence loaded	Longevity of the results obtained remains unsolved
PacBio	Linear amplification	Fluorescently labelled dNTPs	Higher potential for obtaining very long reads, epigenetic characterization of genomes were obtained	Error rate is too higher
Ion Torrent	Emulsion PCR	Detection of released H^+	Faster sequence run with minimal error rate	More hands on time required

In current practice most of the NGS technology follows slightly modified protocol for achieving clonal sequencing which are listed below (Table 1)[9].

Among the most studies performed for detection of respiratory virus, three sequencing methods are widely used: PCR amplification, metagenomic and target enrichment sequencing. Clinical studies on viral detection using NGS highlights the usage in detection of viral strains with atypical subtypes [10]. NGS techniques can be viewed as an adjunct to prevention and control (IPC) strategies in order to uncover putative transmission among people and monitor possible spread which provides useful information on virus typing and amino acid alterations [11][12].

For instance, *Roy et al.* discovered that the complete genomic analysis successfully identified the influenza outbreak that had been missed by the traditional IPC method (based on HA and NA sequencing) [13]. From the clinical studied performed, NGS results proves that it can combined along with diagnostic test results to enhance the clinical diagnosis of patients (observed in majority of SARS-CoV-2 infections) [4].

Advantages

They are more sensitive and specific than earlier methods like virus culture and direct immunofluorescence, and they are quick and reasonably priced. Assays are grouped to identify numerous pathogens in a single test and can be rapidly designed to detect novel or emerging infections [14].

Barriers to routine use and financial costs of performing NGS

Although research shows that NGS has several advantages, writers nonetheless identified a number of obstacles that clinical laboratories should be aware of before considering using NGS for routine diagnosis: (i) Among the standard available methods, the cost of performing NGS is still rather high[15] (ii) The entire NGS work process flow requires larger turnaround times and requirement of labor-intensive workers especially for library preparation step [16](iii) Because of the volume and complexity of sequencing data, advanced downstream bioinformatic studies must be carried out by knowledgeable laboratory personnel[17] (iv) Low viral genetic material may occur from host or microbial cell contamination, particularly in nasal specimens[18][19].

NGS Pyrosequencing (Roche Diagnostics)

Process flow of 454 sequencing includes:

- Step 1: dsDNA in solution is first sheared and appropriate size were chosen
- Step 2: The linkers are attached to the ends of DNA sequences with a spacer between them
- Step 3: Attachment of invitro clones to the beads followed by immobilizing it on picotiter plate which contains single bead. In case of polony method, the beads were immobilized on a matrix attached with a monolayer
- Step 4: One by one, bases are poured across the picotiter plate, and the discharge of light serves to indicate integration. Ligation of anchor primersthat are annealed in one of four positions. Only complementary oligoswill ligate when a population of fluorescently tagged degenerate monomers is added to the monolayer throughout each cycle[20].

NGS Ilumina Sequencing

Process flow involves:

- Step 1: single-stranded sequences are attached to a glass plate precoated with oligos complementary to the adapters
- Step 2: Bridge amplification occurs were single stranded DNA gets binds to the end of primer
- Step 3: Successive PCR amplification results in the production ofclusters of amplified molecules (serve as clones for subsequent sequencing)
- Illumina sequencing uses fluorescently labelled reversible terminators, so as to temporarily terminate the reaction by incorporation of single base sequence at the end. A high-resolution digital image is used to capture which nucleotide gets incorporated at the end of DNA cluster[21].

Metagenomics

Shotgun sequencing is an alternate way to study the genomic sequence of uncultured microbial consortium.Metgenomics plays a major role in determining the genomic sequence which are responsible for causing SARS-CoV-2 viral disease[22]. Our understanding on viral infections and coinfections in patients limited by uneven broad testing for respiratory viruses among people under investigation for SARS-CoV-2[23]. The case study states that for the first two months in the state of Georgia,theyassessed the ability of RNA metagenomic next-generation sequencing (mNGS) to identify SARS-CoV-2. The main problems is how to interpret metagenomic analysis results using NGS in terms of what is therapeutically appropriate for treatment. Absolute agreement between the metagenomic sequencing data and the RT- Real time PCR data was done 47.7%[24]. DNA is extracted from group of cells,instead of targeting a specific genomic sequence, all DNA sequence that are sheared into tiny fragments are sequenced. As a result, numerous genomes present in the sample (i.e., reads) that correspond to different chromosomal regions were obtained[25][26]. The reads obtained from genomic loci that are taxonomically useful (like the 16S), while others will come from coding sequences that reveal information about the biological processes encoded in the genome [27].

Work flow

Steps involved in metagenomicsequencing includes:

- Extraction of DNA sequences
- Preparation of library
- DNA sequencing and assembly
- Genomic annotation
- Statistical reportthat stores the generated results under a database-structured computational repository enabling advanced data management, processing, mining, and meta-mining capabilities[28][29].

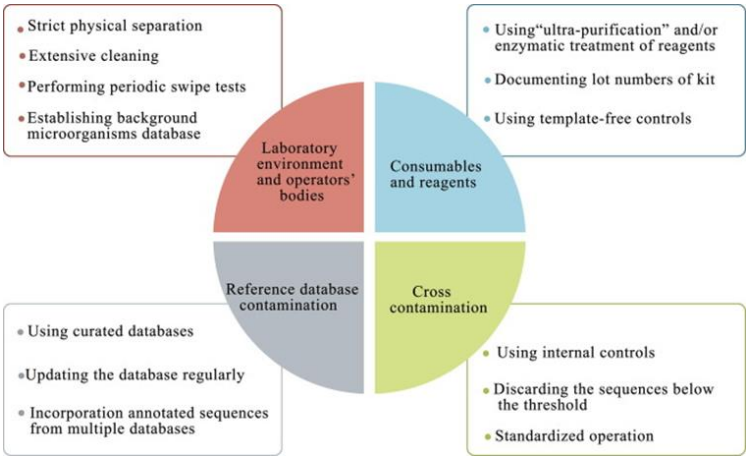

Figure 1. Possible source of contamination for mNGS [9]

Bioinformatic challenges involved in high throughput screening

The difficulty faced while directly recovering viral genomes from clinical specimens using NGS techniques involves noise from host cells and minimal viral RNA quantities[15]. A brand-new amplification method called NetoVir is used for high-throughput sample preparation from host infections[30][9]. Notably, the cause for viral infections remains unsolved because of their complexity in determing genome sequence[31][32]. The main purpose in creation of these bioinformatics tools utilization of specialized computational tools is necessary for bioinformatics analysis of omics and other biological information [33][30].The main purpose behind the development of these tools is to obtain genetic sequence in a simpler way, which are then gradually applied to build more intricate applications[31][34]. Major application includes usage of software for building alignment of genomic sequences, doing statistical analysis and validating data obtained from them[35][36]. Integration of NGS into regular analytical checks for authenticity/contamination at various stages of genomic detection would be the ideal strategy to utilize it for the benefit of customer:

- Application of these technologies in a control laboratory will assist guarantee the integrity and quality of genomic sequence that have been tested. As a result of introducing this technique in viral detection the cost involved in analyzing viral detection is too low.
- The affirmation that the data derived from the software complies with the safety and quality standards that could be extremely beneficial in value addition and establishing it with exception from opponents[36].

Future prospectives

Testing laboratories should soon integrate their techniques and analytical pipelines as NGS becomes a well-established tool[6][10]. The sector will appear to be challenged toward more affordable, high-throughput, and user-friendly solutions for analysis as a result of technological advancements and rising competition [16].

Conclusion

Over the past ten years, the development of NGS technology has completely changed how we understand viral genomes [31].But its a long way to go until NGS technology is utilized to identification of genetic pattern, despite the fact that it has been widely employed to diagnosis genetic illness particularly in the field of virology. Notably, NGS has become less expensive over time, which will entice more laboratories to use it in their studies and put more work into optimizing NGS protocols[15]. Additionally, softwares like Geneious and Qiagen CLC Workbench, provides an intuitive graphical user interface for analyzing complicated databases, which could be helpful for people without familiar in bioinformatics background[37][38]. Future solutions to the existing NGS platform difficulties may result from the quick advancement of NGS technologies and bioinformatics pipelines[37]. Therefore, in the near future, we may anticipate that NGS technique will be predominantly used in wet lab for routine work at a cheap cost and providing more effective therapy, and monitoring new viruses [39].

References

1. Babiker et al., “Metagenomic sequencing to detect respiratory viruses in persons under investigation for COVID-19,” J. Clin. Microbiol., vol. 59, no. 1, 2021, doi: 10.1128/JCM.02142-20.
2. D. Radford, D. Chapman, L. Dixon, J. Chantrey, A. C. Darby, and N. Hall, “Application of next-generation sequencing technologies in virology,” pp. 1853–1868, 2012, doi: 10.1099/vir.0.043182-0.
3. Nieto, F. Pozo, M. Vidal-García, M. Omeñaca, I. Casas, and A. Falcón, “Corrigendum: Identification of rare PB2-D701N mutation from a patient with severe influenza: Contribution of the PB2-D701N mutation to the pathogenicity of human influenza [Front. Microbiol. 8, 575 (2017)] doi: 10.3389/fmicb.2017.00575,” Front. Microbiol., vol. 8, no. JUN, pp. 1–2, 2017, doi: 10.3389/fmicb.2017.01080.
4. Ed. John LMokili,ForestRohwer, “Metagenomics and future perspectives in virus discovery,” Curr. Opin. Virol., vol. 2, no. 1, pp. 63–77, 2012, doi: https://doi.org/10.1016/j.coviro.2011.12.004.
5. Liu et al., “An Optimized Metagenomic Approach for Virome Detection of Clinical Pharyngeal Samples With Respiratory Infection,” Front. Microbiol., vol. 11, no. July, 2020, doi: 10.3389/fmicb.2020.01552.
6. Benefits of NGS”, [Online]. Available: https://sapac.illumina.com/science/technology/next-generation-sequencing/beginners.html
7. C. Ginocchio and A. J. McAdam, “Current best practices for respiratory virus testing,” J. Clin. Microbiol., vol. 49, no. 9 SUPPL., pp. 44–48, 2011, doi: 10.1128/JCM.00698-11.
8. Qi et al., “Detection of respiratory pathogens in clinical samples using metagenomic shotgun sequencing,” J. Med. Microbiol., vol. 68, no. 7, pp. 996–1002, 2019, doi: 10.1099/jmm.0.000968.
9. C. Gaston et al., “Evaluation of Metagenomic and Targeted Next-Generation Sequencing Workflows for Detection of Respiratory Pathogens from Bronchoalveolar Lavage Fluid Specimens,” J. Clin. Microbiol., 2022, doi: 10.1128/jcm.00526-22.

10. D. Wagner et al., “VPipe: an Automated Bioinformatics Platform for Assembly and Management of Viral Next-Generation Sequencing Data,” Microbiol. Spectr., vol. 10, no. 2, pp. 1–10, 2022, doi: 10.1128/spectrum.02564-21.
11. Li et al., “Direct next-generation sequencing of virus-human mixed samples without pretreatment is favorable to recover virus genome,” Biol. Direct, vol. 11, no. 1, pp. 1–10, 2016, doi: 10.1186/s13062-016-0105-x.
12. C. Carbo et al., “Performance of Five Metagenomic Classifiers for Virus Pathogen Detection Using Respiratory Samples from a Clinical Cohort,” Pathogens, vol. 11, no. 3, pp. 1–14, 2022, doi: 10.3390/pathogens11030340.
13. De Vita, “Clinical cases,” Psicoter. e Sci. Um., vol. 8960, no. 3, pp. 516–518, 2021, doi: 10.3280/PU2021-003009.
14. H. Graf et al., “Unbiased detection of respiratory viruses by use of RNA sequencing-based metagenomics: A systematic comparison to a commercial PCR panel,” J. Clin. Microbiol., vol. 54, no. 4, pp. 1000–1007, 2016, doi: 10.1128/JCM.03060-15.
15. Thorburn, S. Bennett, S. Modha, D. Murdoch, R. Gunson, and P. R. Murcia, “The use of next generation sequencing in the diagnosis and typing of respiratory infections,” J. Clin. Virol., vol. 69, pp. 96–100, 2015, doi: 10.1016/j.jcv.2015.06.082.
16. Thorburn, S. Bennett, S. Modha, D. Murdoch, R. Gunson, and P. R. Murcia, “Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19 . The COVID-19 resource centre is hosted on Elsevier Connect , the company ’ s public news and information ,” no. January, 2020.
17. Fig2-Outline-of-the-454-and-polony-sequencing-process-Both-systems-first-fragment-the.”
18. Capobianchi, M. R., Giombini, E., & Rozera, “Next-generation sequencing technology in clinical virology,” Clin. Microbiol. Infect., vol. 19(1), pp. 21–25, 2013, doi: https://doi.org/10.1111/1469-0691.12056.
19. Gandamayu, I. B. M., Antari, N. W. S., & Strisanti, I. A. S. (2022). The level of community compliance in implementing health protocols to prevent the spread of COVID-19. International Journal of Health & Medical Sciences, 5(2), 177-182. https://doi.org/10.21744/ijhms.v5n2.1897
20. Shao et al., “Next-generation sequencing as an advanced supplementary tool for the diagnosis of pathogens in lower respiratory tract infections: An observational trial in Xi’an, China,” Biomed. Reports, vol. 16, no. 2, 2022, doi: 10.3892/br.2021.1497.
21. J.-Y. Chien, C.-J. Yu, and P.-R. Hsueh, “Utility of Metagenomic Next-Generation Sequencing for Etiological Diagnosis of Patients with Sepsis in Intensive Care Units.,” Microbiol. Spectr., p. e0074622, 2022, doi: 10.1128/spectrum.00746-22.
22. Imai et al., “Whole genome sequencing of influenza A and B viruses with the MinION sequencer in the clinical setting: A pilot study,” Front. Microbiol., vol. 9, no. NOV, pp. 1–8, 2018, doi: 10.3389/fmicb.2018.02748.
23. Lewandowski et al., “Metagenomic nanopore sequencing of influenza virus direct from clinical respiratory samples,” J. Clin. Microbiol., vol. 58, no. 1, 2020, doi: 10.1128/JCM.00963-19.
24. Sabour, M. Sabour, and S. Ghorbian, “Clinical Applications of Next-Generation Sequencing in Cancer Diagnosis,” Pathol. Oncol. Res., vol. 23, no. 2, pp. 225–234, 2017, doi: 10.1007/s12253-016-0124-z.

25. Marz et al., “Challenges in RNA virus bioinformatics,” Bioinformatics, vol. 30, no. 13, pp. 1793–1799, 2014, doi: 10.1093/bioinformatics/btu105.
26. Conceição-Neto et al., “Modular approach to customise sample preparation procedures for viral metagenomics: A reproducible protocol for virome analysis,” Sci. Rep., vol. 5, no. August, pp. 1–14, 2015, doi: 10.1038/srep16532.
27. Pappas et al., “Virus Bioinformatics,” Encycl. Virol., no. January, pp. 124–132, 2021, doi: 10.1016/b978-0-12-814515-9.00034-5.
28. N. S. Sahajpal et al., “High-throughput next-generation sequencing respiratory viral panel: A diagnostic and epidemiologic tool for sars-cov-2 and other viruses,” Viruses, vol. 13, no. 10, 2021, doi: 10.3390/v13102063.
29. Principle and work flow of Illumina sequencing.” https://www.cd-genomics.com/blog/principle-and-workflow-of-illumina-next-generation-sequencing/
30. S. Lu, S. Lin, H. Zhang, L. Liang, and S. Shen, “Methods of respiratory virus detection: Advances towards point-of-care for early intervention,” Micromachines, vol. 12, no. 6, 2021, doi: 10.3390/mi12060697.
31. S. Prachayangprecha et al., “Exploring the potential of next-generation sequencing in detection of respiratory Viruses,” J. Clin. Microbiol., vol. 52, no. 10, pp. 3722–3730, 2014, doi: 10.1128/JCM.01641-14.
32. S. Roy, J. Hartley, H. Dunn, R. Williams, C. A. Williams, and J. Breuer, “Whole-genome Sequencing Provides Data for Stratifying Infection Prevention and Control Management of Nosocomial Influenza A,” Clin. Infect. Dis., vol. 69, no. 10, pp. 1649–1656, 2019, doi: 10.1093/cid/ciz020.
33. S. shan Su et al., “Diagnostic performance of the metagenomic next-generation sequencing in lung biopsy tissues in patients suspected of having a local pulmonary infection,” BMC Pulm. Med., vol. 22, no. 1, pp. 1–11, 2022, doi: 10.1186/s12890-022-01912-4.
34. T. J. Sharpton, “An introduction to the analysis of shotgun metagenomic data,” Front. Plant Sci., vol. 5, no. JUN, pp. 1–14, 2014, doi: 10.3389/fpls.2014.00209.
35. T. Kustin et al., “A method to identify respiratory virus infections in clinical samples using next-generation sequencing,” Sci. Rep., vol. 9, no. 1, pp. 1–8, 2019, doi: 10.1038/s41598-018-37483-w.
36. T. Thomas, J. Gilbert, and F. Meyer, “Metagenomics - a guide from sampling to data analysis,” Microb. Inform. Exp., vol. 2, no. 1, p. 3, 2012, doi: 10.1186/2042-5783-2-3.
37. V. Hourdel et al., “Rapid Genomic Characterization of SARS-CoV-2 by Direct Amplicon-Based Sequencing Through Comparison of MinION and Illumina iSeq100TM System,” Front. Microbiol., vol. 11, no. September, pp. 1–9, 2020, doi: 10.3389/fmicb.2020.571328.
38. Widana, I.K., Sumetri, N.W., Sutapa, I.K., Suryasa, W. (2021). Anthropometric measures for better cardiovascular and musculoskeletal health. Computer Applications in Engineering Education, 29(3), 550–561. https://doi.org/10.1002/cae.22202
39. X. Wang, S. S. Braid, and W. D. Rawlinson, “Detection of respiratory viruses directly from clinical samples using next-generation sequencing: A literature review of recent advances and potential for routine clinical use,” no. April, pp. 1–12, 2022, doi: 10.1002/rmv.2375.

40. Y. Wang et al., “Metagenomic analysis of viral genetic diversity in respiratory samples from children with severe acute respiratory infection in China,” Clin. Microbiol. Infect., vol. 22, no. 5, pp. 458.e1-458.e9, 2016, doi: 10.1016/j.cmi.2016.01.006.
41. Y. Y. Qian et al., “Improving Pulmonary Infection Diagnosis with Metagenomic Next Generation Sequencing,” Front. Cell. Infect. Microbiol., vol. 10, no. January, pp. 1–9, 2021, doi: 10.3389/fcimb.2020.567615.