

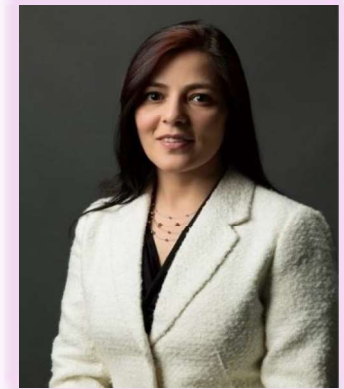


Sequencing

By Megha Agrawal*, PhD
Co Editor-in-Chief

Biotechnology Kiosk, 2, 7 (2020)

DOI: <https://doi.org/10.37756/bk.20.2.7.2>



Next Generation Sequencing Technology in Clinical Diagnostics

Abstract

The promise of next generation sequencing offers hope for the current DNA sequencing technology to evolve from the applications in basic research to transition to the clinical diagnostics. This advancement in the sequencing technology is happening in part due to the introduction of high throughput and benchtop instruments that offer fully automated cost-effective sequencing along with faster assay times. This development is believed to remove the bottleneck of the complex and cumbersome library preparation that include isolation of nucleic acids and the resulting amplified and barcoded DNA with sequencing adapters. Here, we present a brief overview of the principles of next generation sequencing and automation of library preparation along with the diagnostic applications of next generation sequencing in human immunogenetics. Finally, an outlook is presented.

Keywords: microfluidics, multiplexing, genome, NGS.

***E-Mail: megha@biotechkiosk.com**

To cite this article: Agrawal M; Next Generation Sequencing Technology in Clinical Diagnostics, Biotechnology Kiosk, Vol 2, Issue 7, PP: 10-17 (2020); DOI: <https://doi.org/10.37756/bk.20.2.7.2>

Introduction: The Process of Next Generation Sequencing

Over the last few decades, there has been tremendous growth in DNA sequencing technology. This growth has witnessed the evolution of the technology in phases that include the discovery of the double helix DNA shape and the complete sequencing of a human genome along with a continuously widening range of applications in research, agriculture and public health [1-4]. Currently, next generation sequencing (NGS) is employed to sequence complete genomes. However, the usual costs of sequencing a human genome are somewhat high that lie around the \$1000-mark [4]. Researchers have shown that to reduce the costs, a viable way is to increase the throughput of machines. This has resulted in a decrease in sequencing costs per base [4]. A number of studies have shown major successes with regard to reliability along with data that are considered prerequisites for NGS to be applied to clinical diagnostics [4, 5].

Figure 1 schematically shows the various process elements involving the NGS. These include sample pre-processing, library preparation, sequencing itself and bioinformatics [4]. It is to be noted that all modern sequencing technologies require dedicated sample preparation to yield the sequencing library loaded onto the instrument regardless of the underlying principles of the respective sequencing method [4]. The NGS process involves sequencing libraries that consist of DNA fragments of a defined length distribution with oligomer adapters at the 5' and 3' end for barcoding along with the actual sequencing process. Subsequently, the generated data is

analyzed using bioinformatics after sequencing (Figure 1) [4].

Routine laboratory practice in NSG is very important that involves reliable and standardized implementation and also quality control measures for all stages of the NSG process [4]. There are some challenges that are encountered at each of the workflow steps shown in Figure 1. Especially, during library preparation, major challenges include complexity of protocols, contamination and cost [4]. Researchers have shown that the extraction of a sufficient amount of DNA from the sample input without extracting disturbing inhibitors can result in changing the complexity depending on the sample material. Further, carry-over contamination can occur during sequencing runs that can lead to significant errors [6]. In addition, a significant challenge involves the bioinformatics component in the NGS. Studies have shown that the handling of the extremely large amounts of data generated by high-throughput NGS often requires considerable information technology (IT) resources. However, for such comprehensive analysis, there are now enough advanced IT solutions available [4]. These challenges must be overcome to ensure high quality sequencing results [4].

Studies have shown an inherent problem with sample contamination as libraries are usually prepared in parallel [6]. It has been shown that the costs incurred during library preparation derive from laboratory equipment, trained personnel's hands-on time and reagent cost [4].

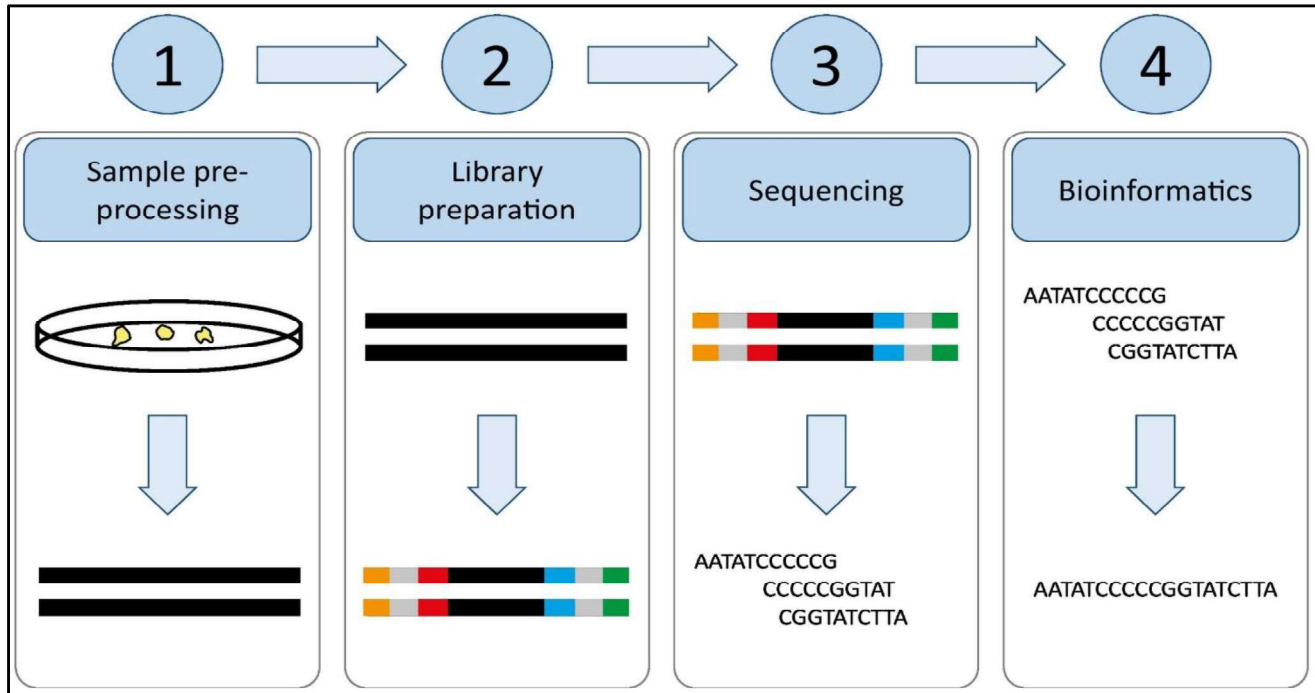


Figure 1: Schematic depiction of workflow for NGS that shows extraction of the nucleic acid from the sample (e.g. blood, sputum, bone marrow, bacterial cultures, etc.). This is followed by library preparation that yields a population of DNA fragments of defined lengths with defined oligomer sequences at both ends to be compatible with the applied sequencing technique. Subsequently, the actual sequencing run occurs on the respective system. Resulting sequence data is then processed by a bioinformatics analysis pipeline [Source: Biotechnology Advances, 2020].

The Automation: Advanced Library Preparation and Performing Complex Protocols

Previous studies have suggested that an automation based NGS process can provide the viable solution to perform complex protocols with high reproducibility, leading to reduced error rates [7]. In general, there are four main scenarios for which automation of library preparation can be considered. These include sequencing centers, diagnostics laboratories, point-of-care and research (Figure 2) [4].

The biotechnology research and industry has already employed sequencing centers and diagnostic laboratories. For future technology development, the point-of-care components are predicted opportunities of NGS [4]. Further, it is envisioned that microfluidic solutions could be an attractive option in sequencing centers along with pipetting workstations that could be used in research (Figure 2). It has been shown that pipetting workstations can be used for high throughput scenarios. Whereas, microfluidic solutions can be employed for low throughput applications [4].

It has been shown by reducing human interaction with the reagents and samples contamination risk can be significantly lowered. This also leads to decreased costs per sample due to the fact that hands-on time is reduced or fewer reagents are used due to miniaturization [4]. However, there are some disadvantages of automation techniques due to the involvement of high number of resources that are needed to initially provide

and implement the instruments in the laboratory workflow and the costs associated with the maintenance. Despite these issues, researchers have shown that profitable efficiency can be achieved by using the system intensively. Another route of profitability is going for low investments and also if high prices are paid for high quality analysis [4].

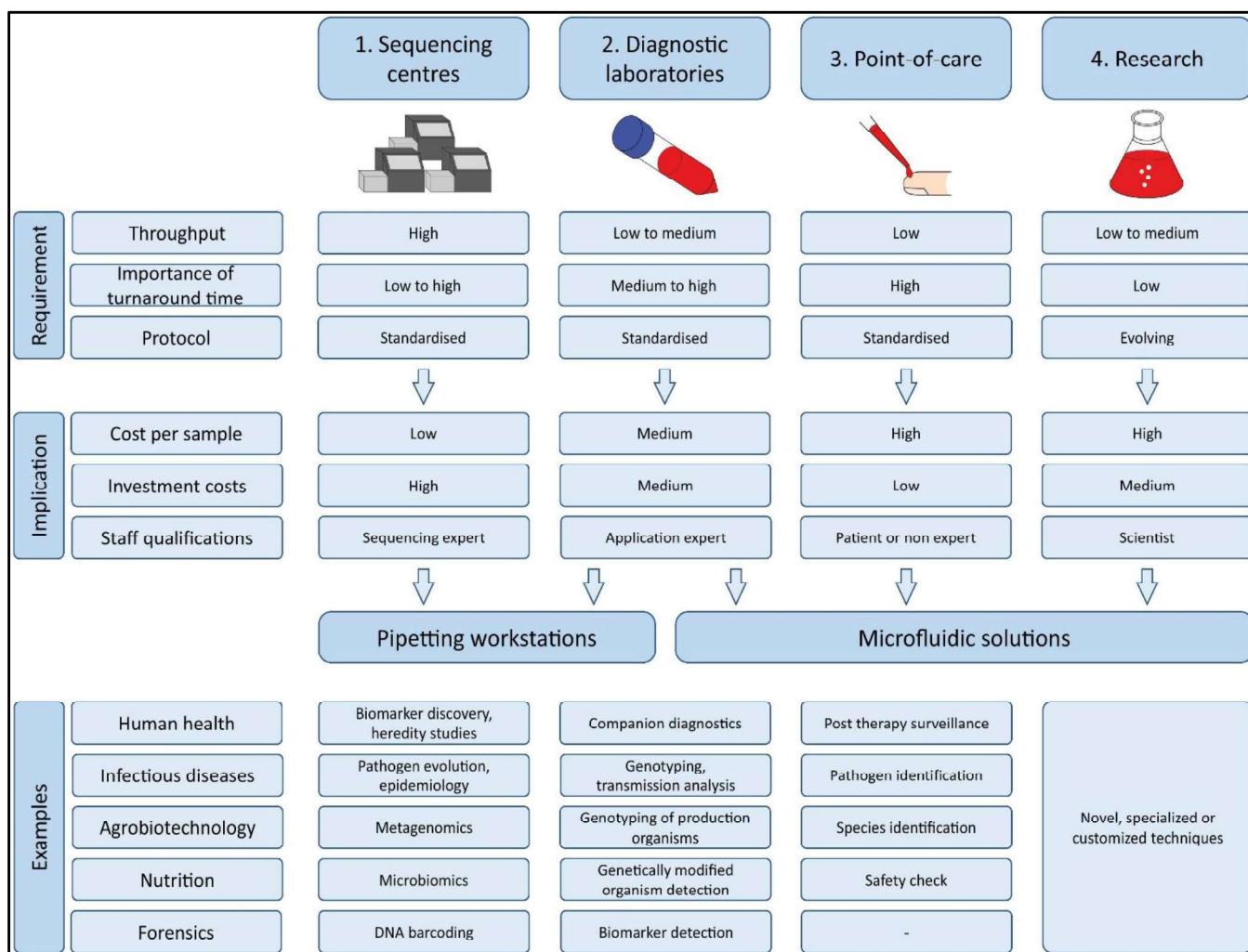


Figure 2: An illustrative overview of four different sequencing scenarios for library preparation for next generation sequencing (NGS). These scenarios differ in terms of associated requirements that include throughput, protocol and importance of turnaround time. Potential examples for different domains are also shown at the bottom of the figure [Source: *Biotechnology Advances* 2020].

Microfluidics Based Automation Solutions

There are multiple advantages of using microfluidics. Microfluidics based operation represents an interesting alternative for automation of library preparation due to the fact that it enables miniaturization, automation and parallelization of biochemical analysis [4]. Studies have shown that the assignment of the genetic information to the cell of origin is a technical challenge within single cell analysis by sequencing. This can be overcome by the application of microfluidics. Especially, droplet microfluidics have been

shown to offer unique solutions when larger cell numbers are analyzed ($> 10^4$) [4]. To this end, researchers employed encapsulated single cells in single beads that provided unique barcodes. Two aqueous flows that contained the cells and the beads respectively were merged shortly before compartmentalization. After droplet generation, cells were lysed and mRNA hybridized on primers that were immobilized on the bead surface [4]. Beads were then recovered and washed manually and subsequently reverse transcriptase transferred the mRNA into cDNA that was amplified by PCR (Figure 3) [4].

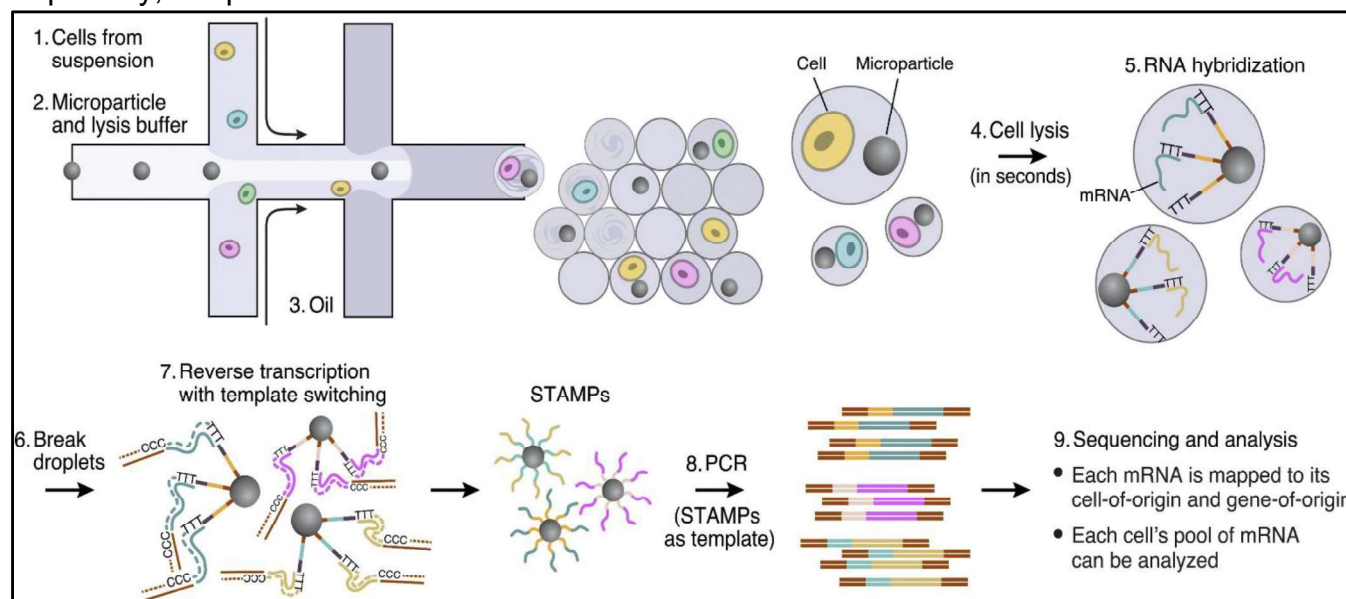


Figure 3: Illustrations showing single-cell barcoding with automated microfluidics solutions with additional features for library preparation [Source: *Biotechnology Advances*, 2020].

Next Generation Sequencing for Human Immunogenetics

Protein-coding sequences have been most widely studied that represent the best understood part of the human genome.

However, the non-coding sequence is known to be substantially larger that constitutes about 98% of the human genome [8]. Researchers have suggested difficulty in the interpretation of any genetic findings for non-coding DNA. Figure 4 shows the exome

sequencing focusing on the protein coding parts of genes, which is more widely used in human genomics than whole genome sequencing [4, 8]. This approach has been shown to require exome enrichment of the sequencing library that is essentially the capture of the DNA sequences containing the protein-coding regions [4]. Because the sequencing output is limited in this step, it reduces the sequencing cost along with the amount of effort required to analyze and interpret the data [4]. Diagnostic applications

are shown in Figure 4. It uses restricted approaches that involve targeted 'panel' sequencing. In this configuration, only certain genes or smaller gene regions of interest are captured and sequenced, usually with a high coverage [4]. These advances have led to the development of NGS-based gene panels that are currently regularly used in diagnostic laboratories and a number of companies and institutions provide panel tests for different types of diseases [4].

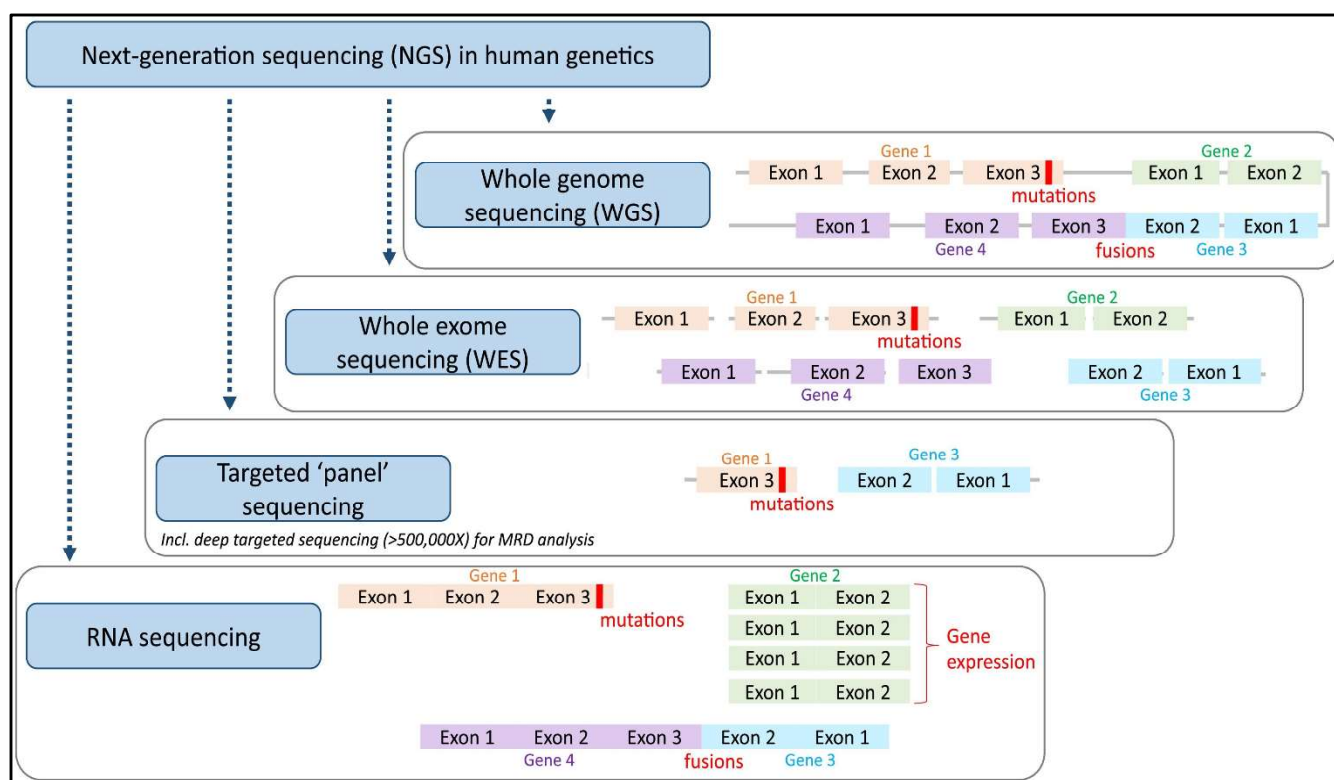


Figure 4: Next generation sequencing diagnostic applications in human genetics showing different sequencing targets for sequencing in human genetics. These involve whole genome sequencing, whole exome sequencing, targeted panel sequencing and RNA sequencing [Source: Biotechnology Advances, 2020].

Conclusion and Outlook

New generation sequencing (NGS) based assays are finding increasing applications in a vast array of diverse fields. However, there

are challenges associated with the NGS technology that include complex and labor-intensive library preparation. To overcome the challenges, different automation solutions

have been suggested that can be considered to address the needs of each specific setting. To this end, the well-established pipetting workstations are usually suggested for commercially available systems for high throughput applications.

Researchers have also paid attention to microfluidic devices that are emerging as an enabling technology for single cell analysis and as an alternative for lower throughput applications. These technologies are associated with lower capital costs that simplify their entry into the automation arena. However, before these technologies make into the commercialization stages, there are still scopes for further improvement in the standardization and reproducibility factors. It is anticipated that future developments for diagnostic laboratories and point-of-care applications would employ microfluidic systems that could considerably improve performance and ease-of-use of sequencing approaches for various advanced biomedical applications.

References:

1. E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, et al., Initial sequencing and analysis of the human genome, *Nature*, 409 (6822), 860-921 (2001), DOI: <https://doi.org/10.1038/35057062>
2. Johannes R. Lemke, Erik Riesch, Tim Scheurenbrand, Max Schubach, Christian Wilhelm, Isabelle Steiner, Jörg Hansen, Carolina Courage, Sabina Gallati, Sarah Bürki, Susi Strozzi, Targeted next generation sequencing as a diagnostic tool in epileptic disorders, 53, 8, 1387-1398 (2012), DOI: <https://doi.org/10.1111/j.1528-1167.2012.03516.x>
3. Sarah E. Calvo, Alison G. Compton, Steven G. Hershman, Sze Chern Lim, Daniel S. Lieber, Elena J. Tucke et. al., Molecular Diagnosis of Infantile Mitochondrial Disease with Targeted Next-Generation Sequencing. *Science Translational Medicine*, 4, 118, pp. 118ra10 (2012), DOI: [10.1126/scitranslmed.3003310](https://doi.org/10.1126/scitranslmed.3003310)
4. J.F. Hess, T.A. Kohl, M. Kotrová, K. Rönsch, T. Paprotka, V. Mohr, T. Hutzenlaub, M. Brüggemann, R. Zengerle, S. Niemann, N. Paust, *Biotechnology Advances*, 41, 107537 (2020), DOI: <https://doi.org/10.1016/j.biotechadv.2020.107537>
5. Jeffrey Gagan & Eliezer M. Van Allen, Next-generation sequencing to guide cancer therapy, *Genome Medicine*, 7, 80 (2015), DOI: <https://doi.org/10.1186/s13073-015-0203-x>
6. Michaela Kotrova, Jan Trka, Michael Kneba, Monika Brüggemann, Is Next-Generation Sequencing the way to go for Residual Disease Monitoring in Acute Lymphoblastic Leukemia? *Molecular Diagnostics and Therapy*, 21:481–492 (2017), DOI: <https://doi.org/10.1007/s40291-017-0277-9>
7. H. Fleischer, K. Thurow, *Automation Solutions for Analytical Measurements: Concepts and Applications*, Wiley-VCH,

Weinheim, Germany (2018).
[DOI:10.1002/9783527805297](https://doi.org/10.1002/9783527805297)

<https://doi.org/10.1038/s12276-018-0087-0>

8. [Brian S. Gloss & Marcel E. Dinger, Realizing the significance of noncoding functionality in clinical genomics, *Experimental & Molecular Medicine* 50, 97 (2018), DOI:
9. Pelegrin M, Naranjo-Gomez M, Piechaczyk M. Antiviral Monoclonal Antibodies: Can They Be More Than Simple Neutralizing Agents? *Trends Microbiol.* 2015;23(10):653-665. DOI: [10.1016/j.tim.2015.07.005](https://doi.org/10.1016/j.tim.2015.07.005)

