Genomics For All: Healthcare, Pharmaceuticals and Beyond

Upasana Mangrolia, Shradha Khater and Gautam Das*

miBiome Therapeutics LLP, Flexcel Park, C Wing, SV Rd, Momin Nagar, Jogeshwari West, Mumbai, Maharashtra 400102
*gautam@mibiome.com

Introduction

Genomics is the study of an organism's genome, its structure, its function, and its interaction with the surrounding environment. A succession of nucleic acids in the form of deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) stores the information required for an organism's structure and its function. To understand the content of the genetic information, it needs to be sequenced. Increased accuracy, increased throughput combined with plummeting cost of DNA sequencing has made genomics an essential tool for probing biological systems. Genomics is routinely used in areas of genotyping, epigenetics, gene editing, structural genomics, metagenomics, epidemiology, and functional genomics. In the field of human genetics and medicine alone the genomics revolution has helped us in answering questions such as what makes us unique, how are we different than other people or living beings around us, tracing our ancestry, and sketching the path of human evolution, genetic bases of a number of diseases like cancer and genome driven drug discovery. The vast trove of data generated using genomics has helped in its application to fields such as medicine and public health, forensics, agriculture, bioengineering, diagnostics, and consumer products.[1] Unquestionably genomics was the most essential tool for epidemiological surveillance during the SARS-CoV-2 pandemic.

In this chapter, we take a sneak peek at the evolution of genomics and then discuss its application in infectious disease, in the current pandemic, in oncology and its application in metagenomics.

BOX / Glossary

What is DNA?

Deoxyribonucleic acid (DNA) is a contiguous sequence of four nucleotides - Adenine (A), cytosine (C), guanine (G), and thymine (T). Two such polynucleotide chains wrap around each other to form a double-helical DNA molecule. In nearly all organisms DNA carries genetic information.

What is a gene?

Gene is a segment of DNA coding for the synthesis of either RNA or a protein. The information in the gene is copied into RNA via a process called transcription. The RNA could be functional by itself or could serve as a template for protein synthesis. Proteins form the body structure of an organism, carry out the metabolic reaction and coordinate all the functions.

What is Genome?

Genome is the entire set of an organism's DNA/RNA. The genome contains all the information required for the organism to function. The human genome is made up of approximately 3 billion base pairs and contains about 30000 genes.

What is DNA Sequencing?

DNA sequencing determines the order of occurrence of nucleotide bases (C,G,T and A) in a segment of DNA. Reading the sequence of bases helps determine the source of the DNA segment, whether the DNA segment codes for a gene or RNA or regulatory elements, and whether changes in DNA are disease-associated.

What is Metagenomics?

Metagenomics is the study of a collection of microbes collected from their natural habitat. As the name suggests (meta + genomics), it involves sequencing the genetic material (DNA or RNA) of many organisms together.

From base-pair to bedside - How did we get here?

The word "genomics" is more modern than the field of genomics itself.[2] The field of genomics was initiated with the discovery of genes as the inheritable material by Johannsen at the beginning of the 20th century.[3] It took several decades to recognize DNA as the basic heritable material and unravel its three-dimensional structure.[4] The later discovery by Watson and Crick in 1953 and sequencing of the first polypeptide chain by Sanger paved the way

for DNA sequencing. The approach of protein sequencing by Sanger included random fragmentation, reading the content of the fragment, overlapping the read fragments to yield a complete polypeptide chain. The basics of protein sequencing were retained for the first RNA sequencing and ultimately for the first DNA to be sequenced.[5] Sanger used the "plus and minus" method of DNA sequencing to sequence the first DNA genome. The same year Sanger introduced the chain termination method of DNA sequencing which became the most widely used DNA sequencing method for almost four decades. With an accuracy of 99.99%, chain termination or Sanger sequencing is still considered the gold standard.

Sanger sequencing

The Sanger method relies on spiking a small amount of fluorescently labeled dideoxynucleotide triphosphates (ddNTP - ddATP, ddTTP, ddCTP, and ddGTP) with normal dNTPs to the DNA synthesis mixture.[6] Each base is labeled with a different colored fluorescent dye. The general convention being A is labeled with green, T by red, G by black, and C by a blue dye. Unlike dNTP, ddNTPs lack the capability to form the phosphodiester bond with the next nucleotide leading to chain termination. During the reaction, DNA polymerase incorporates ddNTPs at random resulting in millions of copies of DNA fragments terminated at random lengths. The DNA fragments are sorted by size using gel electrophoresis. In the automated machine a laser excites the fluorescent dye and the computer detects the identity of each terminal ddNTP. The fluorescent intensity is translated into peaks along the length of the DNA being sequenced.

The continuous improvement and automation of Sanger sequencing aided in sequencing from genes to genomes of microorganisms and eventually some multicellular organisms. Sequencing created a massive amount and a need to store and share the data. GenBank was established in 1982 with just about half a million bases but grew exponentially to over 40 million bases by 1990.

The Human Genome Project

The new era of genomics was ushered by the Human Genome Project (HGP) - an exploration of what lies within us. It started in the year 1990 and took a bit more than one decade to publish the first draft of the human genome. In addition to providing the first glimpses into complete human genome, HGP accelerated advances in sequencing technologies, newer companies entered the sequencing market and adoption of genomics took place at an unanticipated scale. At the start of HGP Applied Biosystems undisputatedly ruled the market. In contrast at the conclusion of HGP 454, Solexa, Illumina, Agencourt, Complete Genomics, Applied Biosystems and Ion Torrent were offering sequencing technologies with higher throughput and reduced cost.

Next Generation Sequencing

The first alternative to Sanger sequencing came in the form of pyrosenquecing produced by Pyrosequencing AB and marketed by 454 Roche Life Sciences. Pyrosequencing relies on detection of light when there is pyrophosphate release. To the reaction mixture of single stranded template DNA, DNA polymerase, ATP sulfurylase and firefly luciferease, one of the four nucleotides is added. The intensity of light emitted on nucleotide incorporation is measured to determine how many nucleotides have been incorporated indicating the presence of complementary nucleotides of template strand. The nucleotide mixture is removed and a new nucleotide is added to the mixture and the process is repeated. Pyroseqeuncing was the first next generation sequencer and could produce 400-500 bp long reads with 99% accuracy. The reason why pyrosequencing revolutionized the sequencing space was its throughput and reduced cost. Pyrosequencing could produce up to 25 million base pairs in a single run and it costs one-sixth compared to Sanger sequencing.[5]

Although a number of sequencers were launched after pyrosequecing, the next leap in genomics came with the introduction of Illumina sequencers. Illumina commercialized the 'sequencing by synthesis' technology that relied on reversible dye-termination chemistry. Similar to Sanger sequencing, a fluorescently labelled reversible terminator dNTP is imaged on incorporation. The terminator is then cleaved and the process is repeated with next reversible dye-terminators. In 2014, Illumina had captured 70% of the DNA sequencers market and more than 90% of sequencing data being generated was from Illumina machines. Illumina also helped realize the goal of $1000 per genome in 2017.

Short read technologies like Illumina and 454 had a great impact on the field of genomics but they have their drawbacks. By their very nature, short read technologies cannot characterize repetitive genomic regions, extreme GC content, structural variants or genomes with multiple homologous elements. Even advanced bioinformatics algorithms can help only to a certain level. Repeats create unresolvable loops in the assembly graph that leads to discontinuous genome. Short read technologies use PCR amplification during library preparation step.[7] PCR amplification are inherently biased against regions containing extreme GC content. These regions ultimately lead to assembly as incomplete smaller

fragments.

Third Generation sequencing technologies

The gap left by second generation sequencing technology has been partially overcome by third generation long-read sequencing.[8] In 2011 Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) started providing long read sequencing. ONT has produced reads upto 2 million base pairs in length, though 10-30 kB reads are more common.[9] Long reads combined with absence of PCR amplification circumvent the problem of fragmented genomes. However, the high error rates of long-read sequencing set a challenge for accurate genomic analyses.[10]

To overcome the limitations of short and long read technologies and using both to their utmost potential hybrid genome assemblies are being used. Long reads can be used to resolve ambiguous loops in the assembly graph created by highly accurate short read technologies. Long reads can also be used to sequence extreme GC regions leading to less fragmented genomes. Hybrid assembly has been used to sequence human, plants and other non-human model organisms.[5]

An integral part of genomics is analysing the huge amount of data that is generated from the sequencers.

Genomics in Infectious Diseases

While considering infectious diseases, Tuberculosis (TB), caused by the Mycobacterium tuberculosis complex (MTBC), is among the leading causes of mortality, i.e. about 2 million people each year globally.[11] In 2019, an estimated 10 million fell ill with tuberculosis globally. Around 1.5 million die of tuberculosis every year globally. Acting upon its survival instincts, similar to other life forms, MTB strategically adapts to develop into MDR-TB, primarily attributed to inadequate diagnostics, inconsistent/ partial treatment and antibiotic drug abuse.[12] MDR-TB is typically characterised by development of resistance in MTB against at least isoniazid and rifampicin, the two most powerful anti-TB drugs.[13] For diagnosis, Sputum-smear microscopy is a common choice of test, however it does not detect extra-pulmonary or smear-negative TB, let alone the screening for various MDR-TB strains.

Different methods for drug resistance profiling in tuberculosis include:

- Phenotypic (DST): First line drugs: R,H,E,Z; Second line drugs: S, Lfx,Mfx, Km,Cm, Am; Other drugs: Lzd, Cfz, Bdq, Dlm, PAS etc
- Growth based drug susceptibility testing
- Rapid drug resistance testing (Genotypic: DRT)
- Nucleic acid amplification based (NAAT): Cartridge (GeneXpert) and Chip (TruNAAT)
- Line probe assay (LPA): First line (R:Rifampicin, H:Isoniazid); Second line (Lfx, Mfx, : Fluroquinolones, Km, Cm, Am: Aminoglycosides)

Genomic techniques play a dynamic role in the control of infectious diseases such as MTB by providing a rapid yet accurate and comprehensive microbial-pathogen analysis. The options include whole-genome sequencing (WGS) and targeted NGS (tNGS). While, WGS in clinical settings would require initial TB culture step so as to generate sufficient bacterial load, tNGS has emerged as a feasible option for faster, comprehensive, and importantly direct sequencing form patient samples.[11]

Application of Next-Generation Sequencing in tuberculosis drug resistance management carried out at miBiome Therapeutics LLP include-

- Cultureless method, faster DR profiling from sputum, Bronchoalveolar (BAL), Cerebrospinal fluid (CSF), ascitic fluid, biopsy or any clinical sample
- Querying Novel Indian isolate specific mutations
- Amplicon based sequencing including targets that distinguish non-MTB from MTB in addition to targets for drug resistance markers
- MDR, pre-XDR and XDR known biomarker panels and novel biomarker panels designed in-house

Therefore, the early detection of Pre-XDR/XDR-TB could guide clinicians in the appropriate adjustment of MDR-TB treatment regimen with effective drugs to prevent treatment failure

How genome surveillance is shaping the SARS-CoV-2 pandemic:

The application of Next-Generation Sequencing during the SARS-CoV-2 pandemic has been the classic example of the advantages of NGS which helped guide the public health response to a pandemic in near-real time. This is inclusive of first distinction and identification of the novel corona virus, screening and prediction of mutations (Genotyping variant analysis), and furthermore, developing screening test for infected population. NGS has played a major role during the pandemic in diagnosis which eventually is aiding in vaccine and drug target selection.

Whole genome sequencing (WGS) of SARS Cov2 is being used for genome surveillance. Although being an excellent tool, application of WGS for public health programs can be very expensive. To overcome such challenges, our team at miBiome Therapeutic LLP have come up with a unique solution for SARS-CoV-2 surveillance. This is inclusive but not limited to the following-

- Using evolutionarily unique regions and targeted

amplicon sequencing with in-house-designed unique/specific primer sets which target only SARS-Cov-2.

- Being targeted sequencing (only 14% of the genome) and not WGS, our method is considerably lower in cost than WGS.
- Interestingly, the required predominant information (>75%) to track current SARS Cov-2 variants is retained in the evolutionarily unique regions, thus not compromising on the genome information required for tracking.
- The lesser data output needed per sample increases the power of multiplexing and endows performing genome surveillance at a cheaper cost.
- To aid in mass screening, 7680 samples can be sequenced in one run using Juno.

Being a part of service providers in the field of NGS, in addition to our above mentioned solution in SARS-CoV-2 and MTb, the other services are presented in the Figure 1.

Genomics in Cancer

Cancer accounted for approximately 10 million deaths in 2021.[14] Genomic aberration such as point mutations, insertions and deletions, variation in copy number are hallmarks of many cancers. NGS can be used to detect these mutations in cancerous genes aiding in diagnosis, prognosis and advance personalized treatment.[15] Whole Genome Sequencing (WGS), Whole Exome Sequencing (WES) and targeted gene panels are being used both in cancer research and clinical settings. Although WGS gives a broader perspective, in clinical setting it might become time and cost prohibitive. WES sequences only the coding region which is a bit more than 2% of the total human genome. Hence, WES in the method of choice when probing mutations in genes coding for proteins. Another time and cost effective method is targeted gene panels. As the name suggests targeted gene panels focusses on a selection of genes specific to a type of cancer. A number of targeted gene panels marketed by Illumina and LifeScience Technologies has seen a rapid uptake in the market in recent years.

As the NGS approach is novel, its use in diagnosis has not been fully realized neither the guidelines for its application in clinical setting has been formalized. In a wide variety of cancers. NGS has started to emerge as a diagnostic tool and for many more NGS-based diagnostic tools are in development. BRAC gene test, which tests for mutation in BRAC1 and BRAC2 genes, are recommended for early onset breast cancer patients or patients with family history of breast cancer. Sequencing of BRAC1 and BRAC2 genes using the Sanger method requires longer times and higher costs as these genes are made up of more than 20 exons each.[16] In addition to reducing cost, NGS helps in reducing the turnaround time of analysis and hence reducing the clinical reporting time. More number of early onset associated genes and high risk genes are being discovered and sequencing panels targeting these genes have been developed. Lin et al., developed a panel of 68 high risk breast cancer genes.

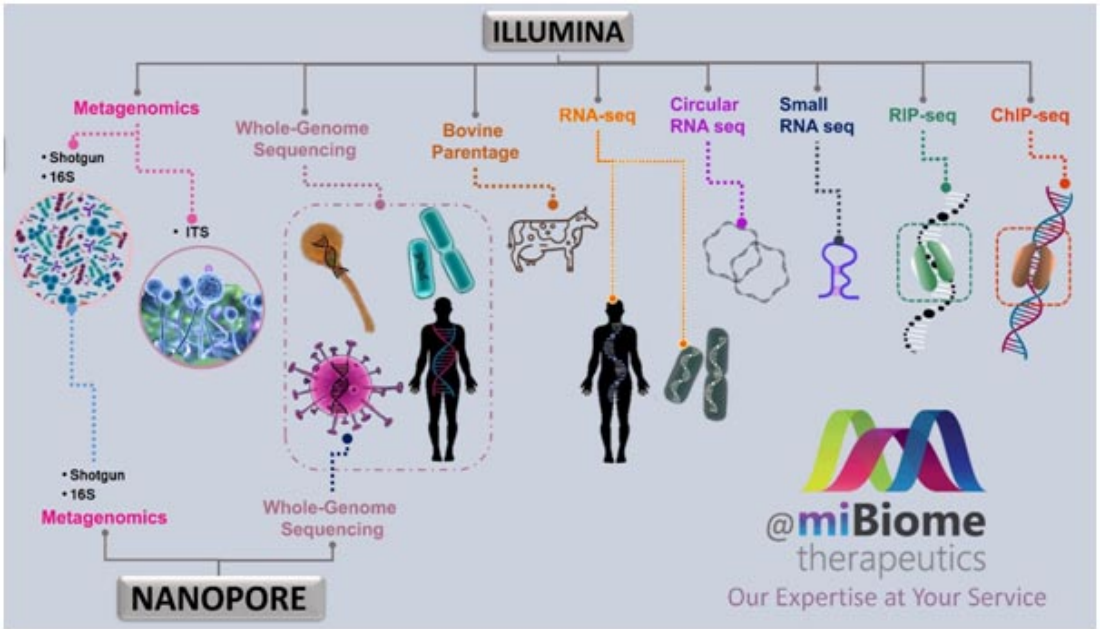

Figure 1: Services at miBiome Therapeutics: miBiome therapeutics is a leadingservice provider for NGS - based metagenomics, Whole genome sequencing, bovine parentage determination, RNA-seq, circular RNA-seq, small RNA-seq, RIP RNA-seq and ChIP-seq. Available with Illumina and Nanopore platforms.

NGS is being used for diagnosis of other cancers such as Lung Cancer, Colorectal Cancer and is being used in selection of therapy. For example, EGFR is targeted by several drugs in colorectal cancer. Relatively a smaller number of CRC patients can benefit from treatments targeting EGFR. Hence, prediction of treatment efficacy, reduction of side effects and cost can be achieved through detection of mutations in KRAS gene.

Other NGS based techniques like RNA-seq is being used to detect gene-fusion events, alternatively spliced transcripts, changes in expression of gene in cancer cells with respect to control or normal

cells. Few targeted gene panels to detect gene fusion events are based on RNA-seq in hematologic tumors. NGS are also being used to understand epigenetic changes, detect microRNAs and other small RNAs.

Genomics has made personalized treatment for cancer patients a realty (Figure 2). From treatment that was organ-centric to analysing the molecular details of cancer for treatment choice, cancer genomics is changing how we look at cancer diagnostics and treatment.[17]

Metagenomics in human health and Disease

Microbes are ubiquitous in nature and their specific composition is reflected in their ecosystem functionality.[18,19] The human body, as a host, has co-evolved with its microbiome, where some commensals evolved to be symbionts and some as pathobionts.[20] Additionally, microbiome architecture varies within different hosts and is primarily impacted by the physical environment.[21] In turn, various biochemical and bio-physical activities expressed by the microbial community, impacts the host metabolic functioning. Thereby understanding these microbes as influencing factors in host heath is assimilating attention. However, the intricacy of the microbial consortia/network makes its identification, profile prediction and mechanism of interaction challenging.[22] Since the traditional approaches of cultivation-based techniques provides a spurious overview on the microbiota, the culture-independent methods, based on PCR have proven to be an effective alternative and paved way for metagenomics. Metagenomic sequencing options have enabled unbiased systemic characterization of microbial communities, furthering therapeutic strategies in near future.[23]

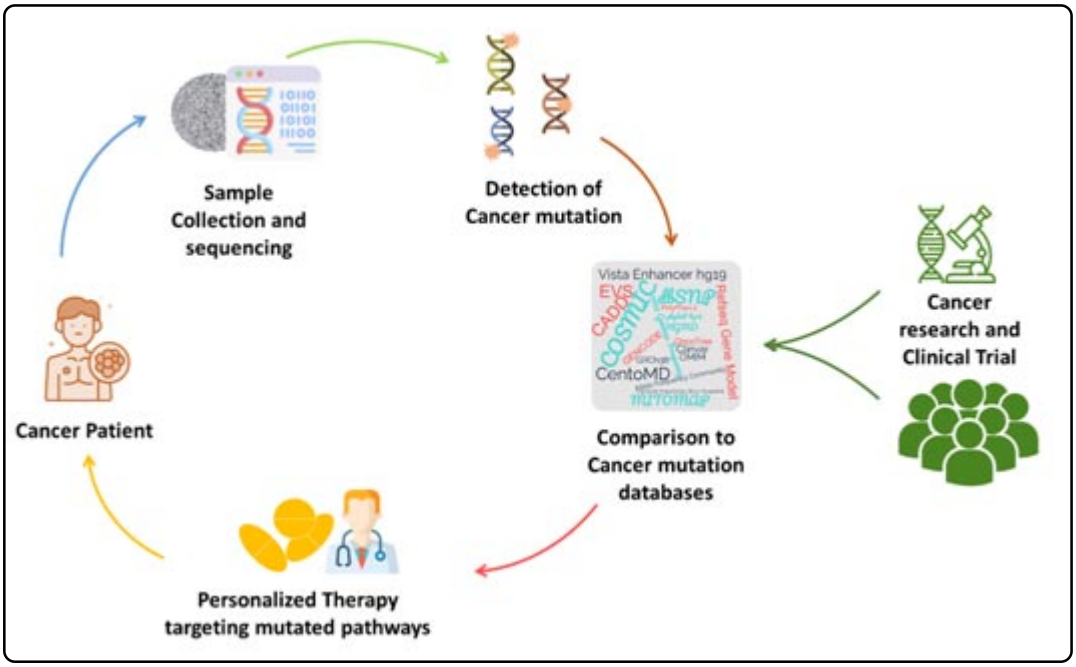

Figure 2: Personalized medicine in Cancer Genomics: The genetic material (DNA/ RNA) from cancer samples are sequenced to detect specific alternations in a cancer patient. Based on the detected aberrations personalized medicines are recommended by specialists to stop the cancer growth.

Metagenomics - Infectious diseases

Considering infectious disease surveillance in public health, knowledge on the causative agent (pathogen) of interest is important. Under standard laboratory practices, this pathogen must be known with a validated test for detection. In case of unknown and unusual infectious diseases due to emerging pathogens, they can be screened against the known suspected pathogens. However, conventional laboratory practices are not always reliable due to evolving pathogenic traits at genetic levels. At such critical exigency, metagenomic investigational methods provide pathogen-agnostic approach to include culturable and non-culturable microbes. Interestingly, studies have reported the efficiency of metagenomics as a diagnostic tool in cases where traditional detection methods for pathogen detection have failed. Foremost application of metagenomics (shotgun) was explored by, Willson et al. and detected presence of Leptospirosis in cerebrospinal fluid (CSF).[24] In another case, shotgun metagenomics identified presence of Abiotrophia defective in culture negative blood and valve samples of endocarditis patient.[25] For other similar studies, investigators employed metagenomics approach in unsolved (culture negative) cases.[26-29] Although pathogen specific molecular tools and serology can confirm the diagnosis/causative agent, the atypical clinical manifestations call for metagenomics approach.

These studies suggest the advantage of metagenomics in identifying non-culturable microbes in presumed sterile samples. This can further be deployed in developing new diagnostic tests, and moreover in developing algorithms to predict the future genetic and etiological instances for the pathogen under consideration.

Detection of antibiotic resistance is the next

engaging application of NGS since gut microbiota majorly comprises of uncultivable bacteria,[30] using metagenomic approach, demonstrated the presence of higher number of antibiotic resistance genes in patients when compared to control, which was otherwise overlooked by culture based methods.[30] Thus metagenomics can be crucial in simultaneous pathogens detection and presence of antibiotic resistance genes.The advantages of NGS are represented in Figure 3.

Metagenomics - Non-infectious diseases

The application of NGS metagenomics goes beyond infectious agent detection. Clinical applications include disease correlation with the host microbiome, viral agent identification in oncology studies and subsequently developing bio-therapeutics. The incomplete clarity/understanding of the microbiome complexity and its involvement on disease pathogenesis has hindered the validation of microbiome-oriented tests in clinical practices.[31] For instance, the infection onset due to Clostridium difficile, an opportunistic pathogen, occurs under altered gut microbiome condition as reported by multiple studies. With applicative approach, management and treatment of Clostridium difficile-associated diseases could be among first clinical application of microbiome NGS.[32] Similarly, metagenomics can be a potential screening tool in discriminating the infectious from the non-infectious illnesses. Several disorders are associated with dysbiotic microbiome such as obesity, diabetes mellitus, cardiovascular diseases and inflammatory bowel disease, and manipulating the microbiome to patient's advantage would require an overview on the microbial landscape.[33]

Moreover, the advancement of NGS into RNA libraries for detection pathogens such as RNA viruses incidentally leads to generation of host gene expression data, i.e. transcriptome analyses (RNA-seq).[34] RNA-seq for gene expression profiling is used at present to characterise infections such as staphylococcal bacteraemia,[35] candidiasis,[36] tuberculosis[37] and Lyme disease.[38] To further this, subjecting the RNA-seq data to machine learning based analyses can be applied for early cancer detection and classification,[39] 2017). Parentheti-cally, WGS approach in identifying the mutated and/ or differentially expressed genes in cancer, data can be looked up for uncovering the viral association with cancer and the subsequent host-virus interactions.[40] Some of these viruses include herpes viruses, papilloma viruses and polyoma viruses, interest-ingly, Merkel cell polyomavirus, now believed to be the cause of Merkel cell carcinoma was first discovered with NGS technology.[41] Metagenomics

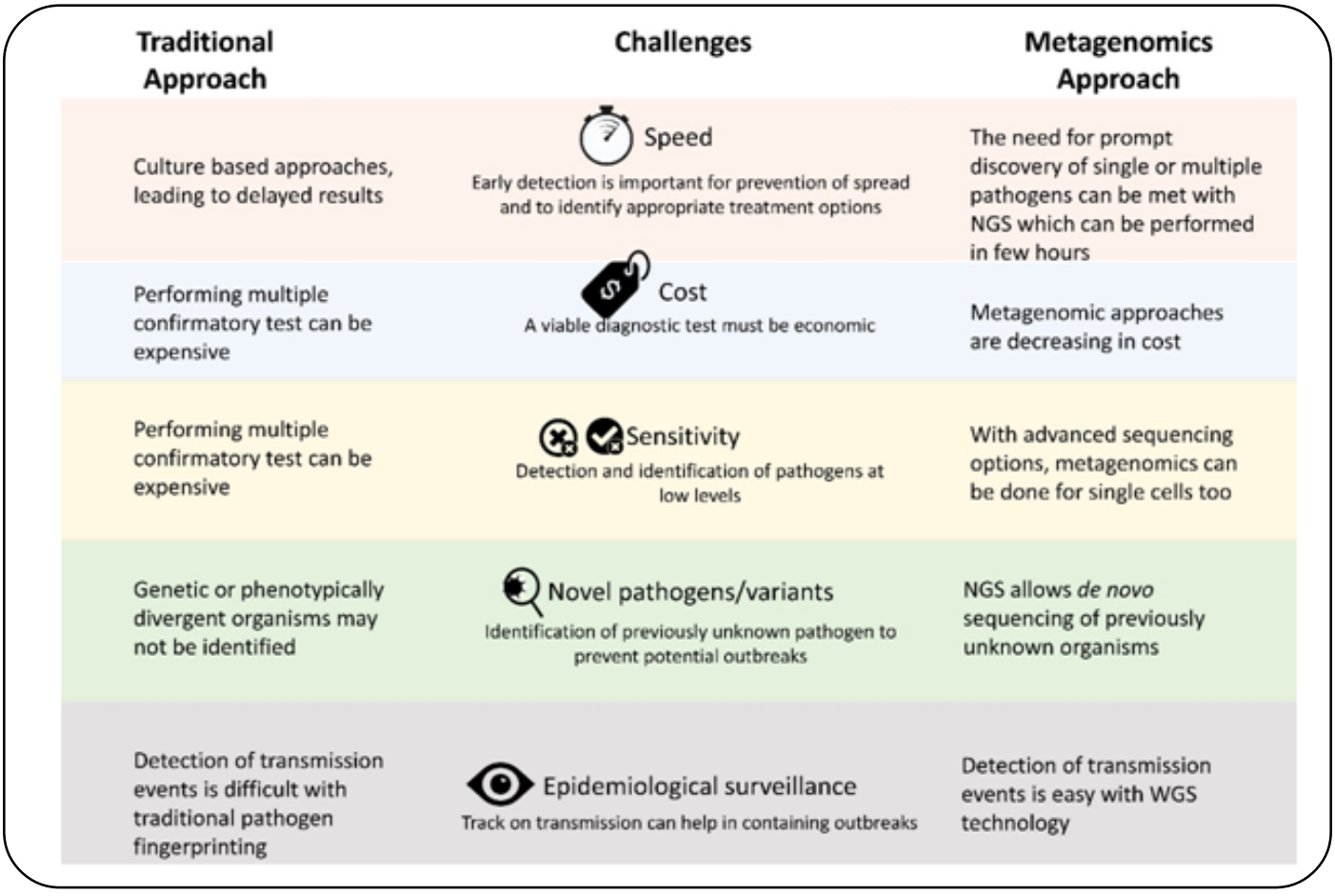

Figure 3: Metagenomics v/s traditional approaches: Challenges in the identification of pathogenic agents and advantages of the metagenomic approaches over traditional approaches.

using shotgun sequencing is a common choice of NGS in clinical metagenomics since it sequences all of the DNA and/or RNA in a given clinical sample. Clinical samples vary significantly in terms of their cellularity (cell-free body fluids or tissues). NGS is an interesting tool in prenatal testing and in identifying mutations to promote diagnosis in pre-symptomatic stages.[42]

Conclusion

With a rise in the urgency of identification of biological causative-agent, NGS has aided DNA/RNA sequencing and enhanced variant/mutation identification in near real time. This allows parallel sequencing of DNA or RNA samples with different lengths as well as whole genome sequencing. Moreover, NGS is applicable for both, cellular and cell-free nucleic acid sequencing - a major advantage over traditional culture-based approaches of pathogen detection. This also broadens the sample type requirement for sequencing, enabling tumor detection using mere liquid biopsy. Additionally, the landscape of microbial population in a given biological niche on the disease status can be explored with metagenomic approach, an important application of NGS. This has furthered our understanding on the role of microbial composition in different infectious as well as systemic diseases. Thus, NGS can provide a comprehensive overview on the disease at genetic level (host), identify causative agent (microbial) and also aid in uncovering the role of the host microbiome in disease pathology.

References:

1. Borry, P. et al. The challenges of the expanded availability of genomic information: an agenda-setting paper. J. Community Genet.9, 103-116 (2018).
2. McKusick, V. A. & Ruddle, F. H. A new discipline, a new name, a new journal. Genomics1, 1-2 (1987).
3. Weissenbach, J. The rise of genomics. C. R. Biol.339, 231-239 (2016).
4. WATSON, J. D. & CRICK, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. Nature171, 737-738 (1953).
5. Giani, A. M., Gallo, G. R., Gianfranceschi, L. & Formenti, G. Long walk to genomics: History and current approaches to genome sequencing and assembly. Comput. Struct. Biotechnol. J.18, 9-19 (2020).
6. Gomes, A. & Korf, B. Genetic Testing Techniques. in Pediatric Cancer Genetics 47-64 (Elsevier, 2018). doi:10.1016/B978-0-323-48555-5.00005-3
7. Chen, Y.-C., Liu, T., Yu, C.-H., Chiang, T.-Y. & Hwang, C.-C. Effects of GC Bias in Next-Generation-Sequencing Data on De Novo Genome Assembly. PLoS One8, e62856 (2013).
8. Klassen, J. L. & Currie, C. R. Gene fragmentation in bacterial draft genomes: extent, consequences and mitigation. BMC Genomics13, 14 (2012).
9. Payne, A., Holmes, N., Rakyan, V. & Loose, M. BulkVis: a graphical viewer for Oxford nanopore bulk FAST5 files. Bioinformatics35, 2193-2198 (2019).
10. Xiao, T. & Zhou, W. The third generation sequencing: the advanced approach to genetic diseases. Transl. Pediatr.9, 163-173 (2020).
11. Katale, B. Z. et al. Whole genome sequencing of Mycobacterium tuberculosis isolates and clinical outcomes of patients treated for multidrug-resistant tuberculosis in Tanzania. BMC Genomics21, 174 (2020).
12. Murray, S. Challenges of tuberculosis control. Can. Med. Assoc. J.174, 33-34 (2006).
13. Zhang, P. et al. Challenges Faced by Multidrug-Resistant Tuberculosis Patients in Three Financially Affluent Chinese Cities. Risk Manag. Healthc. PolicyVolume 13, 2387-2394 (2020).
14. WHO. Cancer. WHO (2022). Available at: https://www.who.int/news-room/fact-sheets/detail/cancer. (Accessed: 15th February 2022)
15. Sabour, L., Sabour, M. & Ghorbian, S. Clinical Applications of Next-Generation Sequencing in Cancer Diagnosis. Pathol. Oncol. Res.23, 225-234 (2017).
16. Shen, T., Pajaro-Van de Stadt, S. H., Yeat, N. C. & Lin, J. C.-H. Clinical applications of next generation sequencing in cancer: from panels, to exomes, to genomes. Front. Genet.6, (2015).
17. Gambardella, V. et al. Personalized Medicine: Recent Progress in Cancer Therapy. Cancers (Basel).12, 1009 (2020).
18. Fierer, N. Embracing the unknown: disentangling the complexities of the soil microbiome. Nat. Rev. Microbiol.15, 579-590 (2017).
19. Bardgett, R. D. & van der Putten, W. H. Belowground biodiversity and ecosystem functioning. Nature515, 505-511 (2014).
20. Kahn, R. A., Fu, H. & Roy, C. R. Cellular hijacking: a common strategy for microbial infection. Trends Biochem. Sci.27, 308-314 (2002).
21. Weinstock, G. M. Genomic approaches to studying the human microbiota. Nature489, 250-256 (2012).

22. Zhao, S. et al. Novel Glycoside Hydrolases Identified by Screening a Chinese Holstein Dairy Cow Rumen-Derived Metagenome Library. Appl. Environ. Microbiol.76, 6701-6705 (2010).
23. Finkbeiner, S. R. et al. Metagenomic Analysis of Human Diarrhea: Viral Detection and Discovery. PLoS Pathog.4, e1000011 (2008).
24. Wilson, M. R. et al. Actionable Diagnosis of Neuroleptospirosis by Next-Generation Sequencing. N. Engl. J. Med.370, 2408-2417 (2014).
25. Fukui, Y. et al. Metagenomic analysis for detecting pathogens in culture-negative infective endocarditis. J. Infect. Chemother.21, 882-884 (2015).
26. Hilton, S. K. et al. Metataxonomic and Metagenomic Approaches vs. Culture-Based Techniques for Clinical Pathology. Front. Microbiol.7, (2016).
27. Pendleton, K. M. et al. Rapid Pathogen Identification in Bacterial Pneumonia Using Real-Time Metagenomics. Am. J. Respir. Crit. Care Med.196, 1610-1612 (2017).
28. Salzberg, S. L. et al. Next-generation sequencing in neuropathologic diagnosis of infections of the nervous system. Neurol. - Neuroimmunol. Neuroinflammation3, e251 (2016).
29. Mongkolrattanothai, K. et al. Neurobrucellosis: Unexpected Answer From Metagenomic Next-Generation Sequencing. J. Pediatric Infect. Dis. Soc. piw066 (2017). doi:10.1093/jpids/piw066
30. Andersen, H. et al. Use of Shotgun Metagenome Sequencing To Detect Fecal Colonization with Multidrug-Resistant Bacteria in Children. J. Clin. Microbiol.54, 1804-1813 (2016).
31. Chiu, C. Y. & Miller, S. A. Clinical metagenomics. Nat. Rev. Genet.20, 341-355 (2019).
32. Samarkos, M., Mastrogianni, E. & Kampouropoulou, O. The role of gut microbiota in Clostridium difficile infection. Eur. J. Intern. Med.50, 28-32 (2018).
33. Boulangé, C. L., Neves, A. L., Chilloux, J., Nicholson, J. K. & Dumas, M.-E. Impact of the gut microbiota on inflammation, obesity, and metabolic disease. Genome Med.8, 42 (2016).
34. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet.10, 57-63 (2009).
35. Ahn, S. H. et al. Gene Expression-Based Classifiers Identify Staphylococcus aureus Infection in Mice and Humans. PLoS One8, e48979 (2013).
36. Zaas, A. K., Aziz, H., Lucas, J., Perfect, J. R. & Ginsburg, G. S. Blood Gene Expression Signatures Predict Invasive Candidiasis. Sci. Transl. Med.2, (2010).
37. Singhania, A. et al. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. Nat. Commun.9, 2308 (2018).
38. Bouquet, J. et al. Longitudinal Transcriptome Analysis Reveals a Sustained Differential Gene Expression Signature in Patients Treated for Acute Lyme Disease. MBio7, (2016).
39. Zhang, Y.-H. et al. Identifying and analyzing different cancer subtypes using RNA-seq data of blood platelets. Oncotarget8, 87494-87511 (2017).
40. Nakagawa, H. & Fujita, M. Whole genome sequencing analysis for cancer genomics and precision medicine. Cancer Sci.109, 513-522 (2018).
41. Feng, H., Shuda, M., Chang, Y. & Moore, P. S. Clonal Integration of a Polyomavirus in Human Merkel Cell Carcinoma. Science (80-.).319, 1096-1100 (2008).
42. Babayan, A. & Pantel, K. Advances in liquid biopsy approaches for early detection and monitoring of cancer. Genome Med.10, 21 (2018).

Instruction for authors of CRIPS

The articles may be (i) reviews or (ii) original research articles. All the articles will be peer reviewed. The length of an article may be approximately 8 printed pages of CRIPS. The contents of the articles should deal with contemporary issues in the field of Pharmaceutical Sciences. The articles should be of high quality, informative and factual. The original research articles should be associated with Electronic Supportive Information (ESI), providing proof of the data presented in the main article and should support the concept being promoted. All the results should be reproducible. The authors should subject the articles to plagiarism check before submitting.