DNA AND RNA SEQUENCING TECHNIQUE – ADVANCE IN DIAGNOSIS OF DRUG RESISTANT INFECTIONS, CANCER MANAGEMENT, REPRODUCTIVE HEALTH & GUT MICROBIOME

Author

Rajal Dave
Greenarray genomic Research & Solution Pvt. Ltd.
Happy Colony, 1st Lane Kothrud, Pune.
rajaldave12@gmail.com
ORCID ID: 0000-0002-6478-4495
918177820784

I. INTRODUCTION

Next generation sequencing technology (NGS) is the latest sequencing platform that assesses the whole genome, either DNA or RNA in a single process in very short time. Technique is similar works like capillary electrophoresis in which enzyme DNA polymerase incorporates fluorescently labelled-deoxyribonucleotide triphosphate (dNTPs) into DNA strand. Whole process carried out during DNA synthesis. Added nucleotides are recognized through excited fluorophore. Compared to other traditional method, NGS can sequence many DNA fragments with high accuracy.

Since 1977, there is huge transformation in clinical research and in drug development. In 2019, 56.6% DNA sequencing accounted specially NGS technology. Moreover from 56.5 %, 24.4 % accounted for oncology. In 2020, IIumina Company has been received authority for COVID sequencing test. In 2020, US government has taken initiative in drug development, clinical research especially in Oncology after get approval for First Liquid Biopsy. There are many sequencing variant available in market. Major are Illumina, Roche, Perkin Elmer, QIAGEN, Thermo Fisher Scientific, Oxford Nanopore Technologies, Bio-Rad Laboratories, Myriad Genetics and Agilent technologies [1].

Table 1: Types and Basic Difference in NGS Generations
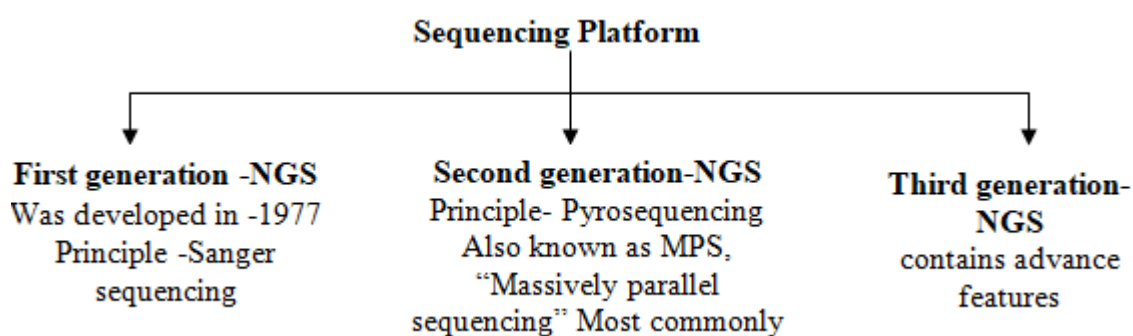
II. BASIC STEPS INVOLVED IN SEQUENCING

Next Generation sequencing is high throughput sequencing methodology which is used to determine nucleotide sequence of genome in a single biochemical reaction. Term high throughput means it can sequence millions of DNA fragment parallel. This sequencing reads are then assembled and mapped to the reference genome through software called "bioinformatics software". DNA sequencing includes sequencing by synthesis, Pyrosequencing and sequencing by ligation. Sequencing by synthesis includes the addition of reversible fluorescent and terminal nucleotides. While in Pyrosequencing, sequencing is begin with target sequence and then nucleotide are added and bind to DNA template and as a result pyrophosphate is release [1,2].

In sequencing, the enzyme DNA polymerase incorporates the fluorescently tagged deoxyribonucleotide triphosphate (dNTPs) in to the DNA strand during DNA synthesis. Added nucleotides identified through fluorophore excitation. NGS technology able to sequence millions of DNA fragments in very short time with high accuracy.
Second generation NGS is developed by Illumina which can sequence by two ways, sequencing by hybridization and sequencing by synthesis. In hybridization, all overlapped oligonucleotide sequences assembled together to identify DNA sequence. In synthesis, fluorescent tagged nucleotides are incorporated by using polymerase and ligase enzyme. While for long read of sequenced data third generation is best variant. All second generation NGS technology depends on amplification before analysis of sequence [2].

Before sequencing, there are three different DNA amplification techniques used to amplify DNA.

1. Emulsion PCR – Ion Torrent GenapSys
2. Bridge amplification - Illumina
3. DNA nanoball generation – BGI group (Beijing Genomic Institute).

There are main four steps of next generation sequencing (NGS) which includes nucleic acid isolation, library preparation, sequencing (clonal amplification) and data analysis [2,4].
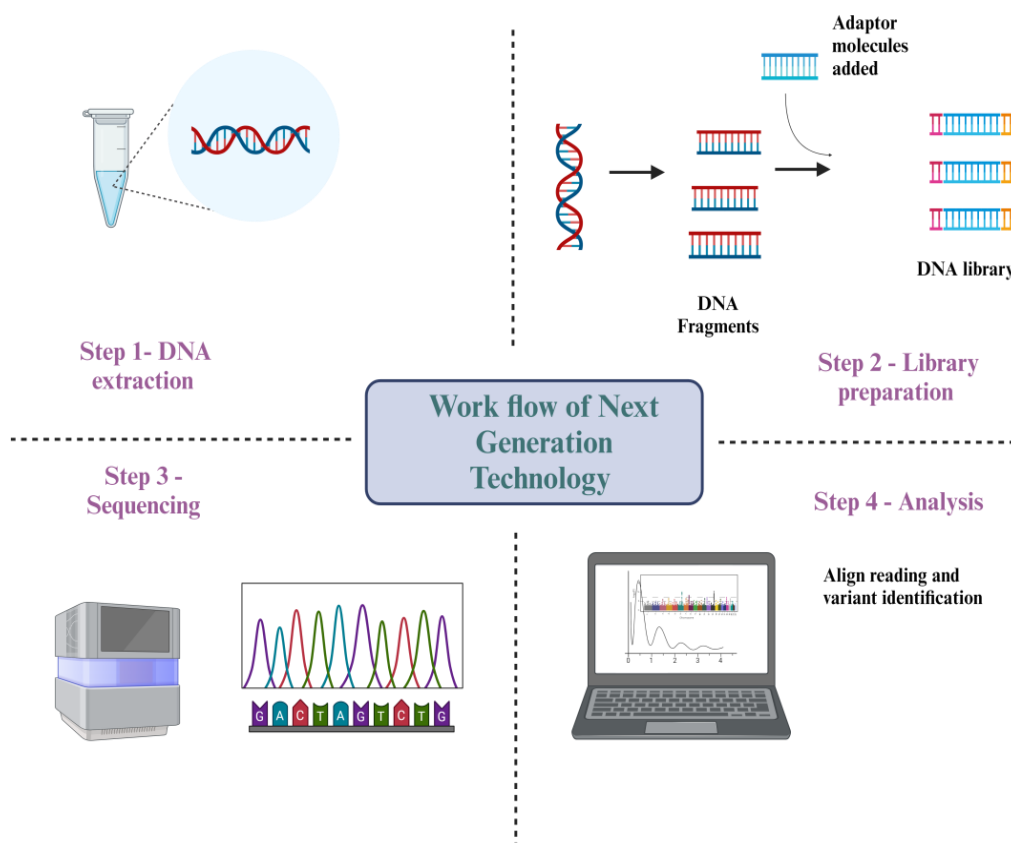
Figure 1: *Basic steps involve in Next generation Technology (NGS)* [16]

Step 1 – DNA / Nucleic Acid Extraction: It is the vital step in NGS with regardless of total RNA, genomic DNA or different RNA types. It is very crucial step to yield high amount of DNA/RNA from sample. The quality and yield of isolated nucleic acid are very important for successful of next step (Figure 1).

Step 2 – Library Preparation: It involves the preparation of DNA/RNA sample preparation. It achieves by fragment the each sample and adds adaptors at both the ends. So the DNA fragments and adaptors at both ends called NGS libraries (Figure 1).

Step 3 – Sequencing and Amplification of Clone: It involves the amplification of DNA fragments and then sequences it by binding of ions surfaces. This will allow developing strong signals to the sequencers. After clonal amplification, all NGS library loaded into sequencers which identify nucleotide one by one (Figure 1).

Step 4 – Data Analysis: It involves the three main stages, processing, analysis and interpretation of raw sequencing data through bioinformatics tool. At the end of process, huge data is generated. Interpretation of all data is major challenge. The primary data is collected in the form of a FASTQ file. It contains the information about raw sequence data. The quality of sequence is achieved by "Phred" score, which is accessing true variant. Each NGS platform has different metrics for calculation of Phred score. All sequenced are aligning in binary files which contain sequence information. These binary files are further analysed by integrated genome viewer software. Software also contains information, location of variant.

Data is interrelated in a form of percentage of desired sequence has been truly tested as a breadth of coverage and number of molecules sequenced from each desired point is in form of depth of coverage [5].

III. QUALITY CONTROL OF EACH NGS RUN

Z Score: It is approach to determine the difference in total cfDNA responsible for fetal Trisomy by determine the Z score. Z score means the number of standard deviations, proposition of reads of particular chromosome is above the mean. For example if the Z score is exceeds to >2.5, result is reported as positive for trisomy.

Z Score with GC Content: GC content represents the how much chromosome is amplified in sample. GC content count rate is used to remove GC bias.

Z Score with GC Content Using Internal Control: It is strategy of quality control that each chromosomes should accessed with internal reference chromosome. The optimal internal reference chromosome is one who has same GC content of clinical sample. This strategy helpful for identification of aneuploidy from small sample [7].

Normalized Chromosome Value (NCV – Algorithm): It allows the compare of read from chromosome of interest. It helps to minimize the intra and inter run sequencing variation [6].

IV. TYPES OF NEXT GENERATION SEQUENCING

NGS technology has vast used in field of research, biomarker analysis, and in therapy.

First Generation – NGS

Sanger sequencing technique

Second Generation – NGS

Fastest and cheapest form of NGS is second generation – NGS more commonly seen for short read sequence (SRS). It can produce 500-600 base pairs.

- Pyrosequencing,
- Sequencing by ligation,
- Sequencing by reversible terminator chemistry.

Box 1: What is difference between Sanger sequencing and NGS?

Both methods used to determine sequence of nucleotides in genome.

Sanger sequence is based on the incorporation of base specific chain terminating dideoxy-nucleotide by DNA polymerase during in vitro DNA replication. It has lower throughput, sequencing one longer DNA fragment at a time.

While In NGS, based on DNA fragmentation and library preparation with parallel sequencing. It has high throughput with more sensitivity.

It is also known as "First generation massively parallel sequencing (MPS) (Table 1). MPSS is allows the tens of million short sequence for rapid sequencing in one run. These fragments are categorized by chromosomes either maternal or fetal DNA. Sequence is compared with normal reference chromosomes. This procedure is called counting. If the amount of

chromosome specific sequence exceeds to the threshold that result reported as trisomy for chromosome while threshold considered as normal disomic chromosome. A trisomic embryo more three copy of chromosome so it has 50% more genetic material. So more cell free DNA number (cfDNA) quantity. Maternal DNA is always more in cfDNA sample [7,8].

Target Sequencing

Rather to amplify the random genomic fragments, target sequencing is selectively amplify the specific genomic region. It allows the selectively specific genomic DNA for amplification. It allows cfDNA to focus on clinically important chromosomes like 13, 18, and 21, X, Y which allows the higher sensitivity and specificity.

For example Illumina, that uses the within procedure innovation called bridge amplification. Silicon chip, which contains millions of amplified DNA with fluorescence tag incorporated. Fluorescence tag excited by laser and as a result fluorescence read from each nucleotide as DNA fragments. These amplified reads analysed and compared with reference reads by bioinformatics. It will identify changes and differences between assembled and reference genome [1, 6, 7].

Variants of Illumina available in market are, iSeq100, NextSeq and NovaSeq.

Third Generation - NGS - It is used in case of complex organisms. It can produce long read sequencing (LRS) data. It can produce >10 kilo bases on average.

It contains

- Single molecule fluorescent sequencing,
- Semiconductor sequencing,
- Single molecule real time sequencing and Nanopore sequencing.

Fourth generation is very rarely available and its aim is genome analysis directly from cell.

Oxford Nanopore technologies (ONT), is also known as single molecule or Nanopore sequencing. It identify nucleotide sequence by passing single long strand of DNA through Nanopore in flow cell which is in ion sequencer. The ONT MiniON and GridON instrument detects

Box 2: Checklist of resources and Quality assured NGS in diagnostics laboratory.

- Financial resources
- Site infrastructure
- Personnel
- Instrument and their service maintenance
- Test commodities & reagents
- Quality control and assurance mechanism
- Bioinformatics analyst
- Reporting staff.

Quality assurance of NGS

1. Establishment of quality assurance program
2. Test method validation & verification
3. Instrument validation and maintenance
4. Technical and quality documentation
5. Training of staff
6. Competency assessment of testing staff
7. Enrollment of External quality

the unique changes in electric current which is generated by A, T, G and C nucleotides. MiniON is one flow cell capacity while GridON has five flow cell capacity which can able to produce 10-30 kb sequencing reads [8].

How NGS works for Drug resistant tuberculosis

Once the sequencing reads assembled and mapped, bioinformatics analysis can identify the changes in the sequence of nucleotide, which was read and compared with reference sequence. Changes in mycobacterial DNA possible to relate with antimicrobial resistance. National TB program has role on elimination of drug resistant tuberculosis.

assessment system (EQAS)
8. Post market validation reagent in market

Turnaround Time (TAT) and Cost Consideration Criteria

TAT and cost are inversely related. Cost may be reducing by batching the samples where need to compromised in TAT.

For clinical use – TAT will be short as possible
For surveillance purpose – Long TAT acceptable than clinical.

Consider sequencing method, existing equipment availability, type, rate of reagent consumption

Application of tNGS, recommended by WHO for DR TB is "targeted" Next Generation Sequencing (tNGS) because it investigate whole gene instead of different regions. tNGS also targets the wider number of drug resistant genes associated with resistance to more anti-TB drugs which include bedaquiline, linezolid, clofazimine, pretomanid and delamanid. tNGS is more expensive than molecular WRDs (molecular WHO-recommended Molecular diagnostics) [11].

V. APPLICATION OF NGS IN DIAGNOSIS OF ANTIMICROBIAL RESISTANT

As the traditional method- culture of clinical specimen, NGS technology has unique application as in role of diagnosis of bacterial and fungal infections.

Antibiotic resistance gene has been detected and identified by application of real time metagenomics oxford Nanopore sequencing (mONS). It can detect lower respiratory tract infection Blood stream infection, prosthetic joint infection and urinary tract infection. Resistome identified by ABRicate, ResFinder, AMR FinderPlus, Ariba and ResFinder tools are used to unassembled illumina data.

Out of total 10 million TB positive patients, only one third patients aware about their diagnosis and treatment. To end epidemic of TB by 2030, we have to expand drug resistant testing facilities. For quick and accurate diagnosis of drug – resistant tuberculosis, NGS is the best option which provide fast and error free genetic analysis compared to culture based methods. In 2023, WHO has recommended to use NGS in drug resistant tuberculosis [11].
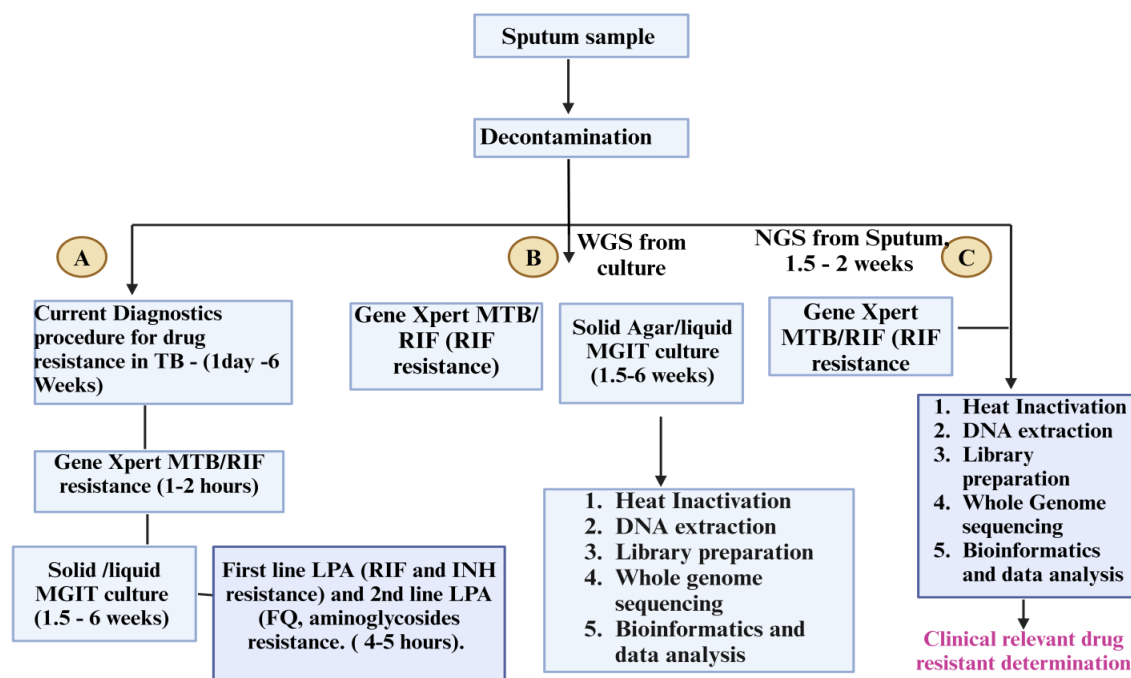
Figure 2: Workflow of drug resistant MTB. *A. Current workflow of diagnosis of TB.
B. Work flow of whole genome sequencing from culture. 3. NGS directly from sample.
TB- Tuberculosis; WGS – Whole genome sequencing; RIF- Rifampicin; INH – Isoniazide;
LPA – Line Probe Assay; FQ – fluoroquinolones [17].*

VI. APPLICATION OF NEXT GENERATION SEQUENCING IN MALIGNANCIES

NGS has significantly increase values in diagnosis of cancer, in identifying mutation in the novel cancer and in improved the diagnosis. It is widely used in the cancer prevention, diagnosis and in treatment. It can sequence the cancer genes and identify its therapeutic vulnerabilities of cancer. ONCOTARGET is the panel of genomic profiling. It contains 1229 probes which is targeting to TP53 gene and 10 oncogenes [9].

NGS - In Early Diagnosis of Pancreatic Cancer

One of recent study researcher had sampled the ovarian and breast tissues to check presence of BRCA1 and BRCA 2 genes. If patient having germ line mutation, that get good benefit from ADP ribose polymerase inhibitors in the time period of treatment. For the NGS application, Retinoblastoma is best example to find out mutation in RB1 gene.
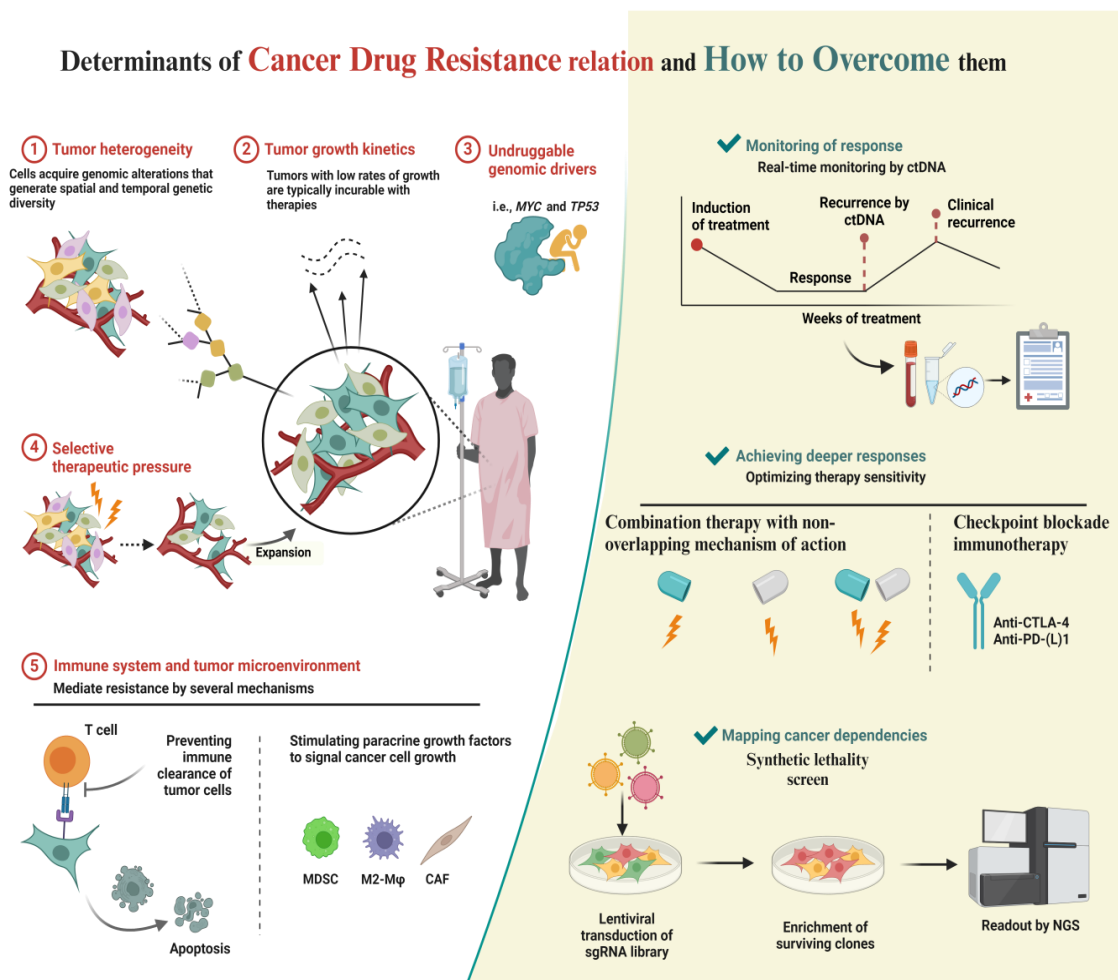
Figure 3: *Role of Next generation sequencing in cancer* [15]

In malignancy, pancreatic cancer is not showing any symptoms till late stage. The survival rate is 3.5 years and historically not possible to detect pancreatic cancer in early stage. There are two types of genetic mutation, hereditary and somatic. In hereditary mutation or germ line mutation is passes from parent while somatic mutation is caused by external environmental factors and it present in all cell of our body but not express in all until some influence of external factors. So for pancreatic cancer, 90% cases are due to somatic mutation of gene either BRCA1 or BRCA2 and KRAS gene. These genes are also responsible for breast and ovarian cancers too. Apart from these, MADH4, ARMET, TP53 and CDKN2A are also responsible for pancreatic cancer. So, NGS can detect all genetic variants in early stage of cancer. In one of study published in October 2021, researchers can identify 44 mutations out of 52 pancreatic cancer cases [10].

VII. NGS IN REPRODUCTIVE GENETIC TESTING

Infertility is major global health issue that affects millions of people. According to WHO's 2023 report, 1 in 6 people are affected with infertility globally. It means 17.5% of adult experiences infertility globally at any point of their life.

Non - Invasive Prenatal Testing (NIPT)

There are two mechanisms responsible to cause genetic diseases. First is genetic change in human genome; variation in copy number either insertion or deletion. Second is epigenic alteration or defect. Next generation technology is the advance technology for analysis of whole genome through sequencing. NGS allows sequencing of DNA fragments and this target sequence is compared with reference. It allow to analyse the many inherited disease in fetal allele such as β-thalassemia. Congenital abnormalities are one of highest causes of infant death in fetus. For example aneuploidy, define as gain and loss of chromosome from normal number. Down syndrome or trisomy 21 is one of the most found with frequency of 1 in 800 live births.
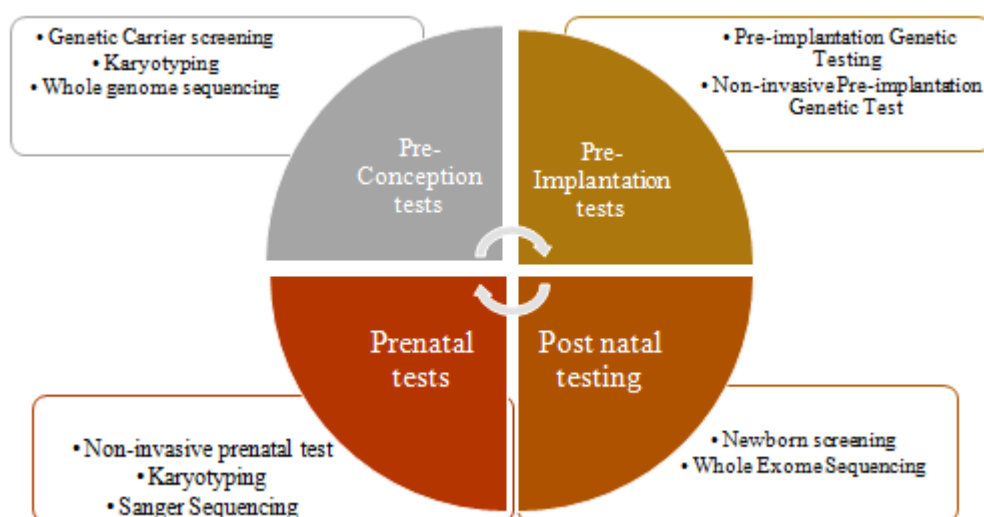

Figure 4: Application of Genome sequencing in reproductive health

After this patau syndrome or trisomy 13 and Edward syndrome or trisomy 18 is found 1 in 50000 live birth. In 1997, first time known that from cell free DNA, can possible from plasma and this finding allow to develop more clinical application. The amount of Cell free DNA is increases with increases of gestation age. From 2011, American college of Obstetrician and gynaecologist has recommended NIPT – Non-invasive prenatal testing as one of most sensitive and accurate test of pregnant high risk women. Prenatal diagnosis involves the investigation of health status of embryo before its birth (Figure 3).

The fetal fraction is key parameter considered in non-invasive prenatal testing (NIPT) means the percentage of cell-free DNA (cfDNA) present in a clinical sample that comes from the fetus. It analyses both maternal and fetal cfDNA to detect genetic conditions. The fetal fraction is usually around 10–15% between 10 and 20 gestational weeks. As the gestation age increases, fetal fraction is also increases. For the test, fetal fraction must be above 4 %, which typically occurs around in the 10th week of pregnancy. Low fetal fractions interpretate as false negative result.

Indication of Cell-free fetal DNA analysis for fetal aneuploidy are if the maternal age more than 35 years, USG report indicating the risk of aneuploidy, prior history of trisomy, positive

report of first trimester, sequential or quadruple screen and parental balanced robertsonian translocation with risk of fetal trisomy 13 or 21[13].

New Born Screening

As per Wilson and Jungner guideline for development and implementation of new born screening testing, WHO (World Health Organization) has introduce New born Screening Program (NBS). NBS is public health program with aim to identify the defects or medically unfit in paediatric population before it becomes the symptomatic or in carrier state. This program first time started on 1960s with successful screening of PKU. This program undertake the identification of certain inherited diseases,

- Phenylketonuria (PKU)
- Hemoglobinopathies to identify sickle cell disease
- Glucose -6-phosphate dehydrogenase deficiencies (G6PD)
- Congenital hypothyroidism (CAH)
- Tyrosinemia type I (HT1)
- Cystic Fibrosis (CF)
- Congenital; adrenalhyperplasia (CAH)

Hence the diagnostics scope of traditional prenatal cytogenetic analysis has been recently extended to new genomic technology [13].

VIII. METAGENOMIC SEQUENCING

Based on this technology, there are four different segments present in the market,

1. 16 S rRNA sequencing – Sequence the region of one conserve gene.
2. Shotgun Metagenomic sequencing
3. Whole genome sequencing
4. De Novo Assembly and Metatranscriptomics

Metagenomics is a modern technology for analysis the resistome present in the gut microbiome, as well as the in the entire genome of an unknown microorganism or from community of microorganisms. This approach include the sequencing process of the bacterial genome and it is map into a reference database which can identify their presence or distribution of resistant genes present within gut microbiome. It involves to understand the mechanisms and pathway of acquiring Antibiotic resistant Genes (ARGs) of gut microbiome.

16 S rRNA Sequencing

It is accounted for more applicability due to high accuracy. Now a days, more research going on Microbiome metagenomics in all country mainly in Europe. The Nanopore technology that is used for analysis gut resistome. Next generation technology present the data which is related to resistome present in the gut microbiome. Gut microbiome is only reservoir of antibiotic resistant bacteria, their gene. There are many bacteria present in the environment

that are non-cultulable and their taxonomical identification is also rare. In this case, metagenomics approaches uses several Antimicrobial resistant gene database (ARGD).

This approach is based on 3 different methods, Targeted metagenomics, functional metagenomics, and sequence-based metagenomics where Sequenced data also provides Resisto Map. With this advance feature, Map able to monitor the levels of ARGs. With NGS, application of bioinformatics tool like Comprehensive antibiotic resistance database (CARD). It gives rapid identification of the organism and their AMR genes data.

1. Targeted metagenomics applies multiple PCR-based methods to identify various different resistance genes. This technique also provides quantitative data to identify a significant number of genes. The cost-effective nature and high throughput analysis make it a significant tool for analysing resistome.

2. Functional metagenomics mainly involves EPIC-PCR technique. Here cloning of a specific gene of interest into a plasmid vector and analyse its transcriptional response in different hosts. This approach can identify novel drug resistance genes in the gut microbiota which will help to understand the scenario of resistance mechanism in the gut microbiome. The EPIC-PCR means Emulsion, Paired isolation and concatenation PCR. It is technique that can be used for analysis of the functional diversity of the gut microbiome.

3. Sequence-based metagenomics is another rapidly growing technique which is useful to characterize the human gut microbe. This method use advanced DNA sequencing method to bacterial genome for high throughput analysis of the resistance genes which is present in gut as a resistome. Also, it is helpful to identify the antimicrobial resistance, virulence genes and information of the entire gene contents.

Shotgun Metagenomic Sequencing

Metagenomic technology utilizes shotgun high throughput sequencing to read all bacterial DNA sequences within a given sample. Metagenomic sequencing can identify various bacterial genes and related their functional pathways within of community. Thus it offers insight into the more functional capacity where screening of the microbiota. For example, metagenomics techniques able to identify gut bacterial enzymes which is require to digest seaweed, which are found in *Bacteroides* species of the Japanese but not in the North American people's intestinal microbiota.

In this technique, DNA extracted from all cells. But, instead of performing target a specific genomic locus for amplification process, all DNA under in subsequently sheared into very small fragments, then all fragments independently sequenced. As a result, DNA sequences (i.e., reads) that align to various different genomic locations for genomes present in the samples [14].

Limitation: Despite these advantages, metagenomic sequence data presents several limitations. First, metagenomic data is relatively complex and large in size. So its analysis is always complicated. Second, metagenomic analysis requires a large volume of data for identification of significant results because of sample contain vast amount of genomic

information. This requirement can pose computational challenge but software development is rapidly advancing and it is improving the efficiency of metagenomic analysis. Third, metagenomic may contains unwanted host DNA, special in the case of microbiome study. In this condition, host DNA can mixed with community DNA that hampers molecular test [12]. Third, contamination is a challenge for environmental sequencing studies and it is very difficult identify contaminants, in case of read was generated and remove it [12]. Finally, metagenomic techniques is expensive to generate compared to amplicon sequences, especially in complex communities or when host DNA greatly outnumbers microbial DNA. Though ongoing advances in DNA sequencing technology are improving its affordability.

Whole Genome Sequencing

Whole genome sequencing is a laboratory procedure that determines the order of bases in the genome of an organism in one process.
Whole genome sequencing carried out by following four main steps:

1. **Shearing of DNA:** Scientists begins the process by use of molecular scissors to cut the DNA into small pieces that are small enough for the sequencing machine to read. DNA is composed of millions of bases (A's, C's, T's and G's).

2. **Bar Coding of DNA:** Addition of small pieces of DNA tags, to identify piece of sheared DNA belongs to bacteria.

3. **Sequencing of DNA:** From multiple bacteria, tagged DNA is combined and put into a DNA sequencer instrument. The sequencer identify the A's, C's, T's, and G's, or bases, that make up each bacterial sequences. The sequencer uses the tag to keep track of which bases belong to which bacteria.

4. **Data Analysis:** Scientists use computer, bioinformatics analysis tools to compare sequences with reference from multiple bacteria and identify differences from reference. The number of differences highlights the closeness of organism, and how they are part of the outbreak.

PulseNet

Whole genome sequencing is the standard PulseNet method for detecting and investigating of foodborne outbreaks which is associated with bacteria like campylobacter, Salmonella spp, vibrio spp, Shiga toxin-producing E-coli and Listeria. This PulseNet method is group of people that compares the DNA fingerprints of bacteria from patients to identify clusters of disease from outbreaks. Techniques has improved the surveillance program in public health laboratories specific for foodborne diseases during outbreaks and trends antimicrobial resistance of food borne pathogen. Techniques provides rapid data.

PulseNet established the structure to support whole genome sequencing at state public health laboratories through:

- Training public health laboratory scientists to perform whole genome sequencing
- Purchasing equipment and supplies

- Updating data analysis systems and software

As the use of whole genome sequencing elaborates the Centre for Disease control's (CDC) national surveillance systems. Whole genome sequencing is a rapid and affordable way to obtain detailed information about bacteria using just one test. Together, we can ensure rapid and economical diagnoses for individuals and collect the evidence needed to quickly solve and prevent foodborne outbreaks.

Advantage and Disadvantages of NGS Technologies

Currently short read sequencing is most used. It gives higher quality data, high depth in lower cost. Short read sequencing (SRS) requires reference sequence. It involves higher equipment cost and more maintenance. While long read has lower accuracy compared top short read sequencing equipment. Both technology requires sufficient technical and maintenance to support unidirectional sample preparation and testing with data analysis and storage. All NGS based test can considered as screening test which should be always cross check with Sanger method.

De Novo Assembly and Metatranscriptomics – RNA sequencing

The study of mRNA with its extraction and their genome analysis, known as Metatranscriptomics. It gives the information of regulation and its expression of genes of complex microbial community. Metatranscriptomics is very useful tool in the research study of the gut microbiome. A core function allows metabolism of carbohydrate, production of energy and synthesis of cellular component as the main function of the gut microbiome. Metabolically active microorganisms and comparison of its activity between each other or between the gastro intestinal tracts gives information of its diet and activity. Metatranscriptomics also gives the information for dietary alterations its impact on health, drug administration effect on gastrointestinal tract, which results into improvement of therapy.

In this test, complementary-DNA (cDNA) synthesized by reverse transcriptase from RNA which was extracted from an environmental sample and used as template for massively parallel shotgun sequencing procedure. These sequences comprises the diverse RNA pools, including sRNA, rRNA, mRNA, and numerous regulatory small RNAs. Thus this data gives collective information on the metabolic activity and its taxonomically diversity, potentially revealing species-specific responses to environmental change as well as covariation in metabolism between distinct community members.

IX. CONCLUSION

NGS technology has more advantages in clinical diagnosis for their accurate and rapidly screening for gene targets which are closely related to the pathogens. There is huge requirement to develop NGS technology as a routine testing process. However, there are gap in a view of bioinformatics software, not able to distinguished live or dead pathogen. So by considering all factors, NGS technology considered as semi-quantitative as it is fail to identify the relation between pathogen and progress of the disease. In regardless of its some limitations, the disease such as tuberculosis and in cancer, NGS has shown outstanding

advantages to identify in early stage. Hence, its clinical application and future development is worth in consideration.

ABBREVIATION

NGS	-	Next Generation Technology
DNA	-	Deoxyribonucleic acid
dNTPs	-	deoxyribonucleotide triphosphate
BGI	-	Beijing Genomic Institute
MPS	-	Massively Parallel Sequencing
TB	-	Tuberculosis
WGS	-	Whole genome sequencing
RIF	-	Rifampicin
INH	-	Isoniazide
LPA	-	Line Probe Assay
FQ	-	fluoroquinolones
mONS	-	Metagenomics Oxford Nanopore Sequencing
RB1	-	Retinoblastoma
SRS	-	short read sequencing
LRS	-	Long read sequencing
ONT	-	Oxford Nanopore technologies
ARG	-	Antibiotic resistant Genes
ARGD	-	Antibiotic Resistant gene database
CARD	-	Comprehensive antibiotic resistance database

REFERENCES

[1] Chaorattanakawee, S., Wofford, R. N., Takhampunya, R., Katherine Poole-Smith,B., Boldbaatar, B., Lkhagvatseren, S., et al. (2022). Tracking tick-borne diseases in Mongolian livestock using next generation sequencing (NGS). Ticks Tick Borne Dis. 13, 101845. doi: 10.1016/j.ttbdis.2021.101845

[2] Gardy, J. L., and Loman, N. J. (2018). Towards a genomics-informed, real-time, global pathogen surveillance system. Nat. Rev. Genet. 19, 9–20. doi: 10.1038/nrg.2017.88

[3] Govender, K. N., Street, T. L., Sanderson, N. D., and Eyre, D. W. (2021). Metagenomic sequencing as a pathogen-agnostic clinical diagnostic tool for infectious diseases: a systematic review and meta-analysis of diagnostic test accuracy studies. J. Clin. Microbiol. 59, e0291620. doi: 10.1128/JCM.02916-20

[4] Ko, K. K. K., Chang, K. R., and Nagarajan, N. (2022). Metagenomics enabled microbial surveillance. Nat. Microbiol. 7, 486–496. doi: 10.1038/s41564-022- 01089-w

[5] von Fricken, M. E., Qurollo, B. A., Boldbaatar, B., Wang, Y. W., Jiang, R. R., Lkhagvatseren, S., et al. (2020). Genetic diversity of Anaplasma and Ehrlichia bacteria found in Dermacentor and Ixodes ticks in Mongolia. Ticks Tick Borne Dis. 11, 101316. doi: 10.1016/j.ttbdis.2019.101316

[6] Worsley-Tonks, K. E. L., Bender, J. B., Deem, S. L., Ferguson, A. W., Fèvre, E. M., Martins, D. J., et al. (2022). Strengthening global health security by improving disease surveillance in remote rural areas of low-income and middle-income countries. Lancet Glob. Health. 10, e579–e584. doi: 10.1016/S2214-109X (22)00031-6.

[7] Schuster SC Next-generation sequencing transforms today's biology. Nat Methods. 2008; 5:16–8. doi: 10.1038/nmeth1156.

[8] Bau S, Schracke N, Kränzle M, Wu HG, Stähler PF, Hoheisel JD, et al Targeted next-generation sequencing by specific capture of multiple genomic loci using low-volume microfluidic DNA arrays. Anal Bioanal Chem. 2009; 393:171–5. doi: 10.1007/s00216-008-2460-7.

[9] Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the association

for molecular pathology and college of American pathologists. J Mol Diagn. 2017; 19:341–65. doi: 10.1016/j.jmoldx.2017.01.011.

[10] Gomy I, Del Pilar Estevez Diz M Hereditary cancer risk assessment: insights and perspectives for the Next-Generation Sequencing era. Genet Mol Biol. 2016; 39:184–8. doi: 10.1590/1678-4685-gmb-2014-0346.

[11] World Health Organization (WHO), 2023, "Use of targeted next – generation sequencing to detect drug-resistant tuberculosis: rapid communication". ISBN 978-92-4-007637-2.

[12] Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. Front Plant Sci. 2014 Jun 16; 5:209. doi: 10.3389/fpls.2014.00209. PMID: 24982662; PMCID: PMC4059276.

[13] Hook EB, Chromosome abnormalities: prevalence, risks and recurrence. In: Brock DJH, Rodeck CH, Ferguson –Smith MA, editors. Prenatal Diagnosis and screening. Edinburgh, Scotland: Churchill Livingstone; 1992.pp.351-392.

[14] Venter, J.C.et al. Environmental genome shotgun sequencing of the Sargasso Sea.Science 304,66-74 (2004).

[15] R. Dave., "Determinants of Cancer Drug resistant and how to overcome them", by BioRender.com (2024), 49 Spadina Ave. Suite 200 Toronto ON M5V 2J1 Canada. www.biorender.com. Retrieved from https://app.biorender.com/biorender-templates

[16] R. Dave., "Work flow of Next Generation Technology", by BioRender.com (2024), 49 Spadina Ave. Suite 200 Toronto ON M5V 2J1 Canada. www.biorender.com. Retrieved from https://app.biorender.com/biorender-templates

[17] R. Dave., "Work flow of Drug resistant MTB", by BioRender.com (2024), 49 Spadina Ave. Suite 200 Toronto ON M5V 2J1 Canada. www.biorender.com. Retrieved from https://app.biorender.com/biorender-templates