REVIEW ARTICLE
Volume - 11 | Issue - 02 | February - 2022 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

EVALUATION OF ORAL SQUAMOUS CELL CARCINOMA AND ORAL POTENTIALLY MALIGNANT DISORDERS USING NEXT GENERATION SEQUENCING: A REVIEW.

Dental Science

Dr. Madhura Bhatt* MDS Oral Pathology, Honorary dental surgeon, Satya Sai Institute of Higher Medical Sciences, Bengaluru, Karnataka, India. *Corresponding Author

Dr. Afshan Anjum MDS Oral Pathology Dental surgeon, Specialist Dental Clinic, RT Nagar, Bengaluru, Karnataka, India.

ABSTRACT

The need for rapid and accurate sequencing of human genome has resulted in the introduction of next generation sequencing (NGS) technology. NGS refers to the second-generation and third generation DNA sequencing technologies where millions of DNA can be sequenced simultaneously. NGS is useful to determine the individual genetic sequence, larger genetic regions, chromosomes as well as to sequence RNA and proteins. Oral squamous cell carcinoma (OSCC) is one of the most frequently occurring malignancies and most often it is preceded by potentially malignant disorders like oral leukoplakia, oral submucous fibrosis and oral lichen planus. NGS technique provides valuable insights in its study and management. This paper provides an overview of the basics of NGS and some of its applications in oral cancer and common potential malignant disorders.

KEYWORDS

DNA (Deoxy ribonucleic acid), RNA (Ribonucleic acid), NGS (Next generation sequencing), PMDs (Potentially malignant disorders), OSCC (Oral squamous cell carcinoma), sequencing, gene expression.

INTRODUCTION

The initial DNA sequencing methodology was introduced by Maxam and Gilbert in 1977, which was soon followed by the chain termination method of DNA sequencing developed by Sanger et al in 1977.[1] Sanger and Maxam – Gilbert technologies are referred to as the first generation of sequencing technology. These were used commonly until the introduction of high throughput technologies or next-generation sequencing (NGS) which opened new perspectives and new avenues in genome exploration and analysis in 2005.[2] Genomic technology has evolved and many next-generation sequencing technologies have been developed since then.[3] NGS can be used to sequence entire genomes or constrained to specific areas of interest, including all 22,000 coding genes (a whole exome) or small numbers of individual gene.[4] The NGS is divided in to second and third generation technologies. The basic characteristics of second generation sequencing technology are: (1) The generation of millions of short reads in parallel, (2) Speed of the sequencing process compared to the first generation, (3) the low cost of sequencing and (4) the sequencing output is directly detected without the need for electrophoresis. The basic characteristics of third generation sequencing is easy sample preparation without the need PCR amplification which also makes it significantly faster than the second generation sequencing technologies. Also these produce long reads exceeding several kilobases unlike short reads in second generation sequencing thus resolving the assembly problem and repetitive regions of complex genomes.[2]

Oral squamous cell carcinomas (OSCC) are a group of cancers which arise from the mucosal lining of the oral cavity.[5] OSCC corresponds to 95% of all oral cancers. It is associated with severe morbidity, recurrence and reduced survival rates.[6] Sometimes, oral cancers are preceded by lesions which are visible clinically, are noncancerous in the beginning and therefore have been termed precancerous. The most common oral potentially malignant disorders are leukoplakia, erythroplakia, oral lichen planus and oral submucous fibrosis. A large number of these oral mucosal lesions have a tendency to transform into malignancy.[7] Next-generation sequencing technologies provide a powerful tool for understanding the genetic basis of tumor initiation and progression, which makes it very useful to guide us in personalized precision medicine.[8]

DISCUSSION

NGS refers to the second-generation DNA sequencing technologies where millions of DNA can be sequenced simultaneously. Some of the next gen sequencing methods employed are Roche/454 life science, Illumina/Solexa, SOLiD system and HeliScope. Application of NGS in decoding the genomic database of various oral diseases may possess therapeutic and prognostic value. Since the first DNA sequencing in 1970s, there has been tremendous advancements in the technologies aimed to determine the entire human genome.

Methodologies of clinical NGS:

1. DNA Extraction:

DNA extraction from the sample which is followed by DNA quantitation.

2. Library preparation:

It refers to the process of preparing DNA for use on a sequencer. It involves breaking DNA into fragments and adding adaptors to the ends. Adaptors may include molecular bar codes (to allow pooling of patient samples), universal polymerase chain reaction (PCR) primers, hybridization sequences to bind the DNA fragment to a surface, and recognition sites to initiate sequencing. The term library refers to these fragments of DNA with flanking adaptors that are ready for sequencing.[9]

3. Target Enrichment:

The resulting library undergoes enrichment for both whole exome analysis and targeted testing or is sequenced directly for whole genome analysis. Enrichment may be performed by hybridization to complementary sequences (sequence capture) or by PCR.[9]

4. Sequencing:

Clinical sequencing is commonly performed on 2 main types of instruments: Illumina sequencers (San Diego, California); or the Ion Torrent series of machines (Thermo Fisher Scientific).[9]

5. Bioinformatics:

The raw data reads from either type of instrument undergo a series of bioinformatics processes to deliver a variant call file (VCF).

These processes include demultiplexing, quality analysis, mapping of the reads to a reference genome (resequencing), and variant identification/ annotation. The use of bar codes to tag a specimen's DNA fragments allows multiple samples to be pooled and sequenced together, thus decreasing the sequencing cost. Demultiplexing process is where all reads are sorted by bar code/sample before further analysis. This demultiplexed file with raw reads is referred to as a FASTQ file. Following this step, the individual reads for a sample are mapped to a reference genome which is referred to as the BAM file (.bam) and any difference between the reference and the sequencing read is noted. Identical (duplicate) reads are discarded for whole genome sequencing but not for the amplicon-based sequencing. If multiple reads show the same difference, a variant is called (the threshold for the number or percentage of reads required is determined by the laboratory and should be validated. The quality of signal for an individual base read and the mapping quality are also factors considered when calling a variant. The output file that defines all the variants for a sample and their allelic fractions is referred to as a variant call file (VCF). The variant call file will contain all variants including common variants, and additional bioinformatics tools are used to filter out variants meeting certain criteria. Before implementation, clinical NGS requires end-to-end validation from DNA extraction through the bioinformatics pipeline.[9]

6. Interpretation of variants:

There are guidelines for the interpretation of germline variants put

forth jointly by the American College of Medical Genetics (ACMG) (now the American College of Medical Genetics and Genomics), the Association of Molecular Pathologists, and the College of American Pathologists (CAP).These guidelines assign strength of evidence for various criteria regarding a particular variant and rules for combining all the criteria to classify that variant as pathogenic, likely pathogenic, uncertain significance, likely benign, or benign.[9]

When we re-sequence the entire genome and map the sequence back to the human genome to identify mutations, it is called the whole genome sequencing. It provides the advantage of full coverage of the entire genome, including promoters and regulatory regions. Therefore, whole genome sequencing is mostly used to identify novel and rare mutations. On the other hand, whole exome sequencing involves sequencing all exons of all known genes at a relatively deeper depth. This is also more cost effective compared to whole genome sequencing. It is used to identify genes associated with cancer, diabetes, immunologic disorders etc. Transcriptome sequencing involves sequencing complementary DNA (cDNA) fragments generated by reverse transcription of RNA. Cancer epigenetics can also be analyzed using NGS.[10]

Applications of NGS in cancer diagnostics, therapy and research:

1. Diagnosis: Many tumor subtypes are now defined by genetic mutations they have. By the use of whole genome (or whole exome) sequencing, novel genetic aberrations and associated potential therapeutic targets are found in many cancers. This information also sheds light on the pathogenesis of these tumors.[10,11]
2. NGS for personalized cancer therapy: An increasing number of therapies are now tied to DNA sequencing results (Targeted therapy).[10,11]
3. When a patient stops responding to a targeted therapy with known resistance mutations.[11]In this scenario, patients can be redirected to other available targeted therapy or be enrolled to clinical research studies which offer them additional chance at a treatment.
4. NGS in hereditary cancer syndrome genetic testing.[10]

NGS in Oral squamous cell carcinoma (OSCC):

Several previous studies have documented that the most common genetic aberrations in OSCC involve the FAT1 (FAT Atypical Cadherin 1), CASP8 (cysteine-aspartic acid protease 8), TP53 (tumor protein p53), CDKN2A (cyclin-dependent kinase inhibitor 2A), NOTCH1 (Notch homolog 1), FBXW7 (F-Box And WD Repeat Domain Containing 7), HRAS (Harvey Rat sarcoma virus), and PIK3CA (Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha) genes.[12,5] TP53 is the most frequently mutated gene in OSCC. Among the different mutations observed in TP53 in OSCC, missense mutations are found to be the most common. In addition, it should be noted that most of the mutations in this gene are located near the DNA binding domain of the protein, which is associated with poor prognosis. PIK3CA is another commonly mutated gene in OSCC and these are commonly seen in HPV (Human Papilloma virus) positive oral cancers. Most often, the mutation is seen in the exon 20 region of this gene.[12]

Targeted therapy in OSCC:

EGFR (epidermal growth factor receptor) is a common target for the targeted therapy in OSCC. The drugs in this category are subdivided in to two depending on their targeting mechanism. Cetuximab and nimotuzumab function as monoclonal antibodies against EGFR. The EGFR tyrosine kinase inhibitors are gefitinib, erlotinib, and afatinib. Another important mechanism which is targeted by the targeted therapy is the tumor angiogenesis through Vascular endothelial growth factor (VEGF) and its receptors. Some of the anti-angiogenic drugs or inhibitors of VEGF include bevacizumab, sorafenib, aflibercept, and vandetanib. The mammalian target of the rapamycin (mTOR) signaling pathway plays a key role in regulating metabolic processes in cells.[12]

Immune check point inhibitors: (ICI)

Pembrolizumab and nivolumab are the FDA approved antibodies that target the programmed cell death protein 1 (PD-1). The newly developed durvalumab and atezolizumab also bind to PD-L1 (Programmed death-ligand 1) of tumor cells. Durvalumab as well as atezolizumab are currently under phase III clinical trials for HNSCC (Head and neck squamous cell carcinoma) treatment.[11]

Expression of miRNA

MicroRNAs (miRNAs) are small, non-coding RNA molecules, made up of 22 nucleotides and which regulates the gene expression in various organisms. There is differential expressions of miRNA among cancerous tissue and benign tissue which help in diagnosis. NGS is very useful for identifying novel miRNAs as it does not require knowledge of the miRNA sequences in advance. miR-21 was the most commonly identified miRNA using NGS and it is up-regulated in OSCC. Apart from being a promising biomarker, miRNA also acts as a potential therapeutic target.[12]

Cancer associated fibroblasts (CAF) in OSCC:

CAF are the most prominent non-immune cells cancers. The cell of origin of CAF or progenitor are local fibroblasts, pericyte, adipocytes, endothelial cells bone marrow derived mesenchymal stem cells, macrophages and cancer stem cells. With the help of single cell transcriptomic sequencing (scRNAseq) we are now beginning to characterise CAF heterogeneity within OSCC. We now know that myofibroblast activation in the stroma (CAF) is a promoter of the malignant transformation in oral potentially malignant lesions like leukoplakia, oral submucous fibrosis and erythroplakia. It is also a marker of tumour aggressiveness, including invasion, metastasis, absence of T-cells and resistance to treatment implying the multifactorial role of this cell.[13] The treatments directed to target CAF have been unsuccessful and it has been difficult to identify specific CAF targets due to a limited understanding of the molecular and functional phenotypes of this cell. But with the advancement of new technologies, especially the single-cell RNA-sequencing, is unraveling the complexity associated with CAF and it needs to be seen if this paves way for new advancements and success in therapies directed towards this cell.[13]

Metagenomics in Oral cancer:

The human body is is a habitat for over 100 trillion microbial cells which live in symbiosis with their host. Microbiome is a term coined to describe "the collective genomes and gene products of all microbes residing within an organism.[14] Their role in health and disease is not well understood because of difficulties in isolating and culturing the organisms. With the recent development of metagenomic sequencing, which allows the discovery of all the DNA sequences in a specimen, it is now possible to investigate the microbiome in more detail than ever.[15]We know that specific bacteria strongly correlate with OSCCs, such as Streptococcus, Peptostreptococcus, Prevotella, Porphyromonas gingivalis and Capnocytophaga gingivalis. So far, we have gathered some knowledge about the bacterial community as they are the predominant species in oral cavity but there is a growing interest in the less abundant microbial communities such as fungi and viruses.[16]Sequencing studies are revealing new organisms and are discovering many differences between the microbiome of healthy persons and patients with oral cancer. Also, sequencing studies show the existence of DNA sequences that may be from novel microbes but are actually of unknown origin and therefore are referred to as the dark matter. All the sequencing data must be further studied to reveal novel pathogens and new pathways in the development of oral cancer and new targets to treat oral cancer.[15]

Application of NGS in Oral Lichen Planus

Oral lichen planus (OLP) is a immune mediated disorder and the immune system is affected by a lot of things including drugs, systemic and metabolic diseases, physical and mental stress, and also the oral microbiome or pathogens. In attempts to improve the treatment outcomes, various molecular pathways active in oral cavity cells of OLP patients and the related pathogens were identified using NGS techniques and in the process, it was found that HNF4A (hepatocyte nuclear factor alpha) gene network activation was present in the oral cavity cells of OLP patients, and several periodontopathogens, including Prevotella denticola, dominated in patients with OLP. This bacteria could activate the HNF4A network in cells of the oral cavity. Prevotella denticola may serve as a target for future therapies in this context. [17]

Application of NGS in Oral Leukoplakia

In a study, whole exome sequencing was performed to differentiate between progressive and non-progressive oral leukoplakia using the frequency of exomic variants, particularly in DNA damage repair pathway genes like BRCA1 (BReast CAncer gene 1), BRCA2 (BReast CAncer gene 2) and other double strand break (DSB) repair Fanconi anaemia (FA)/BRCA pathway genes. [18]

Evasion of apoptosis is one of the important hallmarks of cancer. This could be due to inactivation of CASP8 (which encodes Caspase 8

protein) through various pathways such as mutations, epigenetic modifications, altered transcription, alternative splicing and post translational changes. In a study, using NGS, 56% cancer and 30% leukoplakia tissues were found to have CASP8 somatic mutations and it was suggested that CASP8 mutation could serve as a potential signature for progression of oral cancer from leukoplakia. It was also implied that although clinically diagnosed leukoplakia did not show any sign of invasion but these were not free from molecular alteration. [19]

CONCLUSION

NGS is being developed as an important research means in assessment of genomic alterations in various human diseases. The advantage is that most of the available NGS platforms share a common parallel sequencing process of clonally amplified DNA molecules. With ever improving knowledge regarding its utility, NGS can have a wider role in clinical practice provided some of the limitations are addressed. The need of the hour is to educate the current and future clinicians regarding its applications, the availability of accurate bioinformatics tools to assess the enormous data generated; and to improvise the technical skill and expertise of the laboratory operators. Overall, NGS is a significant discovery to help in disease diagnosis and implementation of appropriate therapy with minimal complications.

REFERENCES

1. Pattan V, Kashyap R, Bansal V, Candula N, Koritala T, Surani S. Genomics in medicine: A new era in medicine. World J Methodol. 2021 Sep 20;11(5):231-242.
2. Kchouk M, Gibrat JF, Elloumi M. Generations of Sequencing Technologies: From First to Next Generation. Biol Med. 2017:9(3).
3. Levy SE, Myers RM. Advancements in Next-Generation Sequencing. Annu Rev Genomics Hum Genet. 2016 Aug 31;17:95-115.
4. Behjati S, Tarpey PS. What is next generation sequencing? Arch Dis Child Educ Pract Ed. 2013 Dec;98(6):236-8.
5. Chai AWY, Lim KP, Cheong SC. Translational genomics and recent advances in oral squamous cell carcinoma. Semin Cancer Biol. 2020 Apr;61:71-83.
6. Varsha BK, Radhika MB, Makarla S, Kuriakose MA, Satya Kiran G, Padmalatha G. Perineural invasion in oral squamous cell carcinoma: Case series and review of literature. J Oral Maxillofac Pathol 2015;19:335-41.
7. Kumar S, Debnath N, Ismail MB, Kumar A, Kumar A, Badiyani BK, Dubey PK, Sukhtankar LV. Prevalence and Risk Factors for Oral Potentially Malignant Disorders in Indian Population. Advances in Preventive Medicine. 2015: Article ID 208519; 7 pages.
8. Nakagaki T, Tamura M, Kobashi K, Omori A, Koyama R, Idogawa M, Ogi K, Hiratsuka H, Tokino T, Sasaki Y. Targeted next-generation sequencing of 50 cancer-related genes in Japanese patients with oral squamous cell carcinoma. Tumour Biol. 2018 Sep;40(9).
9. Yohe S, Thyagarajan B. Review of Clinical Next-Generation Sequencing. Arch Pathol Lab Med. 2017 Nov;141(11):1544-1557
10. Guan YF, Li GR, Wang RJ, Yi YT, Yang L, Jiang D, Zhang XP, Peng Y. Application of next-generation sequencing in clinical oncology to advance personalized treatment of cancer. Chin J Cancer. 2012 Oct;31(10):463-70.
11. Gagan J, Van Allen E.M. Next-generation sequencing to guide cancer therapy. Genome Med.2015: (**7**)**:** 80.
12. Kim S, Lee JW, Park YS. The Application of Next-Generation Sequencing to Define Factors Related to Oral Cancer and Discover Novel Biomarkers. Life (Basel). 2020 Oct 2;10(10):228.
13. Bienkowska K, Hanley C.J, Thomas G.J. Cancer-associated fibroblasts in oral cancer: A current perspective on function and potential for therapeutic targeting. Frontiers in Oral Health. 2021:(2);1-11.
14. Zhao H, Chu M, Huang Z, Yang X, Ran S, Hu B, Zhang C, Liang J. Variations in oral microbiota associated with oral cancer. Sci Rep. 2017 Sep 18;7(1):11773.
15. Shillitoe EJ. The Microbiome of Oral Cancer. Crit Rev Oncog. 2018;23(3-4):153-160.
16. Deo PN, Deshmukh R. Oral microbiome and oral cancer - The probable nexus. J Oral Maxillofac Pathol. 2020;24(2):361-367.
17. Zhong EF, Chang A, Stucky A, Chen X, Mundluru T, Khalifeh M, Sedghizadeh PP. Genomic Analysis of Oral Lichen Planus and Related Oral Microbiome Pathogens. Pathogens. 2020 Nov 16;9(11):952
18. Farah CS, Jessri M, Bennett NC, Dalley AJ, Shearston KD, Fox SA. Exome sequencing of oral leukoplakia and oral squamous cell carcinoma implicates DNA damage repair gene defects in malignant transformation. Oral Oncol. 2019 Sep;96:42-50.
19. Singh R, Das S, Datta S, Mazumdar A, Biswas NK, Maitra A, et al. (2020) Study of Caspase 8 mutation in oral cancer and adjacent precancer tissues and implication in progression. PLoS ONE 15(6): e0233058.