IJPSR (2025), Volume 16, Issue 2 **(Research Article)**

INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES AND RESEARCH

Received on 22 August 2024; received in revised form, 23 September 2024; accepted, 25 October 2024; published 01 February 2025

GENDER-BASED DISPARITIES IN GENE EXPRESSION AND MUTATION PROFILES IN ORAL SQUAMOUS CELL CARCINOMA REVEALED BY THE WHOLE EXOME SEQUENCING AND RNA SEQUENCING

K. V. Lisina [*1] and Bharti Mittal [2]

DBT-Bioinformatics Centre [1], Computational Biology Laboratory, Department of Bioinformatics, Bharathiar University, Coimbatore - 641046, Tamil Nadu, India.
BDG Life Sciences [2], Senior Technical Research Associate, New Delhi - 110001, Delhi, India.

Keywords:

Drug design, Galaxy web server, Head and neck squamous cell carcinoma, Oral squamous cell carcinoma, RNA sequencing, Whole exome sequencing

Correspondence to Author:
K. V. Lisina

Research Scholar,
DBT-Bioinformatics Centre,
Computational Biology Laboratory,
Department of Bioinformatics,
Bharathiar University, Coimbatore - 641046, Tamil Nadu, India.

E-mail: lisina.kv@gmail.com

ABSTRACT: Head and Neck Squamous Cell Carcinoma (HNSC) is a fast-growing form of cancer. This cancer originates in the squamous cells of the mouth and throat and primarily affects individuals who engage in risky lifestyle choices such as tobacco and alcohol consumption. The poor prognosis and high mortality rate of HNSC underline the urgent need for extensive research and novel treatment approaches. This comprehensive study aims to determine the genetic and gene expression differences between male and female patients with Oral Squamous Cell Carcinoma (OSCC). It will also be examined whether differences in unique genes are common in both male and female patients. We used the Galaxy web server to conduct a study employing whole exome sequencing (WES) and RNA sequencing (RNA-Seq) data. The research we conduct requires accurate sample preparation that allows us to obtain useful insights. We found variations in the aligned reads within the WES data that provide insight into particular genetic variations associated with OSCC. At the same time, we discovered various gene expression patterns in the RNA-sequence data. In particular, we observed that the expression of some genes altered in the presence of associated genetic mutations, displaying an obvious connection between genetic variation and modulation of gene expression. In an attempt to identify a gender-independent treatment target, we found a variety of genes that were significantly upregulated or downregulated in both male and female OSCC samples. These genes shared dysregulation across genders, making them interesting treatment targets. Using this insight, we started drug design efforts to develop appropriate precision medicine options for OSCC patients of all genders. This ground-breaking research advances our understanding of OSCC. It paves the way for targeted, gender-neutral medications that provide enhanced outcomes and quality of life for people with this difficult condition.

INTRODUCTION: Head and neck cancer originates from various tissues of the lip, oral cavity (mouth), larynx (throat), salivary glands, nose, sinuses, or facial skin. The lips, mouth, and larynx are the areas most commonly affected by this type of cancer [1].

QUICK RESPONSE CODE

DOI: 10.13040/IJPSR.0975-8232.16(2).417-29

This article can be accessed online on www.ijpsr.com

DOI link: https://doi.org/10.13040/IJPSR.0975-8232.16(2).417-29

Head and neck cancer is primarily the result of alcohol and tobacco consumption, including smokeless tobacco. However, there is a growing number of cases associated with the human papillomavirus (HPV).

Early detection of this type of cancer can often lead to successful treatment, but the prognosis is usually poor if the cancer is diagnosed at a later stage. A combination of surgery, radiation therapy, chemotherapy, and targeted therapy are some of the treatments that work best for the affected [2, 3]. Head and neck cancer is a major global health concern,

with approximately 650,000 new cases and 330,000 deaths each year. In 2018, it ranked seventh among the world's most common cancers, with 890,000 newly diagnosed cases and 450,000 individuals losing their lives due to this disease [4]. Around 75% of cases are attributed to alcohol and tobacco consumption [5]. Squamous cell carcinoma is a type of cancer that develops from squamous cells, a type of epithelial cells found in the skin and mucous membranes. This form of cancer is responsible for more than 90% of all head and neck cancers [6]. Patients with head and neck cancer have seen improvements in their quality of life, and survival rates due to advancements in diagnosis, local management, and targeted therapy. These advances have enabled better treatment and care options for people with this type of cancer [7].

HNSC is widely known for its poor prognosis and high mortality rate. This highlights the urgent need for extensive research and innovative treatment approaches. This study aims to bridge the knowledge gap on genetic and gene expression variations between male and female patients with oral squamous cell carcinoma. Specifically, the study aims to address the lack of information on whether there are gender-specific differences in genetic and gene expression patterns in OSCC. This study aimed to identify dysregulated genes in both male and female patients with oral squamous cell carcinoma. These findings highlight the need for a better understanding of common therapeutic targets that could be utilized to develop treatment strategies that are not gender specific. Before this study, there was limited research on common treatment goals between genders.

The results of this research will be used to initiate drug development and establish precise medical approaches for patients with OSCC, regardless of gender. Previous studies may have lacked sufficient research in this area, especially regarding gender-neutral precision medicine techniques for patients with OSCC. The main objective of this study is to address the existing knowledge gaps on Oral Squamous Cell Carcinoma, with a particular focus on how genetic variations and gene expression patterns are associated with gender. By addressing these gaps, the study aims to provide significant insights that may contribute to the development of more effective and integrated therapeutic strategies for people with OSCC, irrespective of their gender. Advanced techniques such as Whole Exome Sequencing and RNA Sequencing can bridge the gap in our knowledge of genetic and gene expression differences between male and female patients with Oral Squamous Cell Carcinoma.

By using these techniques, we can better understand genetic mutations and gene expression patterns in patients with OSCC. This valuable information may help develop more effective treatments for individuals with OSCC. This study aims to propose the idea that identifying therapeutic targets that are not gender-specific and incorporating them into precision medicine strategies could significantly transform the treatment of Oral Squamous Cell Carcinoma. This approach has the potential to not only improve the effectiveness of treatment but also make it more integrative, thereby improving the quality of life and increasing survival rates for all patients with OSCC, regardless of gender.

MATERIALS AND METHODS:

Data Acquisition: The primary goal of this study was to identify gene variants responsible for regulating specific genes in male and female patients. To achieve this, we accessed Exome and RNA sequence data from two different sources: the NCBI-SRA database (Project ID-PRJEB24758) [8] and the data for this study were deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB24758 [9]. Subsequently, we created two separate projects in Galaxy, [10] a widely used bioinformatics platform, to analyze DNA and RNA data. The raw data was uploaded to Galaxy for further processing.

Quality Assessment: The quality of the sequencing data was assessed to ensure the reliability of subsequent analyses. This evaluation was carried out using the FastQC tool [11]. The FastQC results were then consolidated using MultiQC [12]. To improve the data quality, adaptor sequences were removed using Trimgalore [13]. We then re-evaluated the quality of the trimmed sequences using FastQC and consolidated the results again using MultiQC.

Exome Data Analysis: Exome sequences were aligned to the reference genome using the BWA-

MEM2 [14]. The aligned BAM files were visualized using the IGV (Integrative Genomics Viewer) tool [15, 16]. To ensure the accuracy of subsequent downstream analyses, duplicate reads were removed from the BAM files using the RmDup tool [17]. Mpileup files were then generated using the mpileup tool [18] in Galaxy, allowing a comprehensive comparison of alignment files with the reference positions and documenting the positions where base changes occurred. Variants in the alignment were detected using VarScan [19].

RNA Data Analysis: For RNA-sequence data, alignment to the reference genome was performed using the HISAT2 tool [20]. The overall alignment percentage was assessed and the number of read pairs that were uniquely mapped during the RNA sequence alignment was determined. The resulting aligned file was visualized using the IGV tool. To quantify specific gene expression from the RNA sequences, we used the feature count tool [21]. In galaxy differential gene expression analysis was performed using DESEQ2, [22] comparing male samples as control and female samples as treated, as well as female samples as control and male samples as treated. Annotated genes were identified using the DAVID bioinformatics resource [23].

Gene Enrichment Analysis: To identify common and unique genes between male and female samples, we created Venn diagrams using the web tool [24]. Gene enrichment analysis was performed using GSEA [25] to determine the involvement of these genes in immunological functions and gene ontology. Furthermore, functional enrichment analysis was performed using a G-Profiler [26] to understand how genes were functionally enriched in specific pathways.

Data Mining: For a broader perspective of data mining, we used the cBioPortal database [27, 28], and TCGA to check the presence of identified genes in larger datasets and also specifically for oral squamous cell carcinoma. This allowed us to assess the significance of the mutations identified in different individuals or samples.

Target Validation: After potential gene targets were identified, we proceeded with validation. This included examining the expression of the genes one by one, evaluating mutations, and confirming their presence in the GDC-TCGA database [29]. Further *in-silico* analyses were performed to collect more data for target validation. These validated targets could then be pursued for drug design and development.

Common Genes and Variant Analysis: Common genes of both genders were subjected to variant analysis using IGV tools against RNA bam files and DNA Variant Call Format (VCF) files. A gene, common in both genders was found to vary with different gene expression patterns in the RNA bam file. Associated protein networks were identified using the STRING database [30]. This finding provides a potential target for future research and drug development.

***In-vivo* Validation:** For future steps, the study may progress to *in-vivo* validation, where compounds could be synthesized to treat patients of both genders. This personalized medicine approach aims to minimize side effects and improve the effectiveness of treatment.

RESULTS: This study aimed to identify genetic mutations and differential gene expression between male and female oral squamous cell carcinoma patients. The WES data provided insights into the genetic changes associated with OSCC by examining aligned reads. Duplicate readings were removed to ensure accurate downstream analysis. Mpileup files were generated to compare the alignment with the reference sequence and genetic variants identified using the Varscan software tool. When analyzing the RNA data, sequences were aligned to the reference genome using the HISAT2 tool. Additionally, visualization of genetic variants and quantification of gene expression were performed. DESEQ2 was used to compare different gene expression variations in male and female OSCC patients. Annotated genes were identified using the online bioinformatics resource DAVID.

The MA plot was created when we performed differential gene expression analysis on RNA sequencing data using DESEQ2 and compared two conditions, “control” and “treated.” The MA plot is generally used to visualize differences in gene expression levels in “control” and “treated” conditions. The x-axis indicates the mean of the normalized counts over several orders of

magnitude. The y-axis represents the log fold change with values ranging from -2 to 2 and smaller values, given in scientific notation (1e-02 and 1e-01). In this graph, genes with altered gene expression under both control and treated conditions were represented with positive and negative values (highly upregulated and highly downregulated). A volcano plot **Fig. 1** was generated when we took male samples as controls and female samples as treated after RNA data analysis. A volcano plot **Fig. 2** was generated when we used female samples as controls and male samples as treated. Then we can compare them. This makes it easy to find out in which gender the gene expression is highly upregulated and downregulated. Then we can take it forward for experimental studies that are further validated in male and female samples. So that we can know which gender has highly varied genes. Therefore, in this study, we created a group of male and female patients with OSCC.

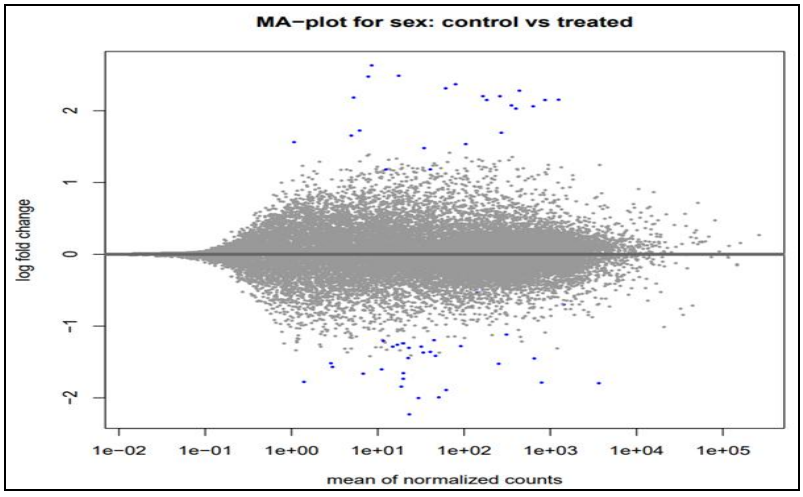


FIG. 1: VOLCANO PLOT FOR MALE AS CONTROL AND FEMALE AS TREATED

The MA plot generated by the DESEQ2 tool is used to analyse differential gene expression in genomics. This MA plot compares "control" and "treated" conditions in terms of sex-specific gene expression. In this context, positive values indicated increased gene expression, while negative values indicated decreased gene expression in the "treated" group compared to the "control" group. The x-axis represents the mean of the normalized counts spanning several orders of magnitude from 1e-02 to 1e+05. The y-axis represents log fold change values ranging from -4 to 4, with positive and negative values. MA-Plot helps to visualize different patterns of gene expression and discover the target genes that influence the treatment of the disease. These genes are expressed differently in both male and female patients.

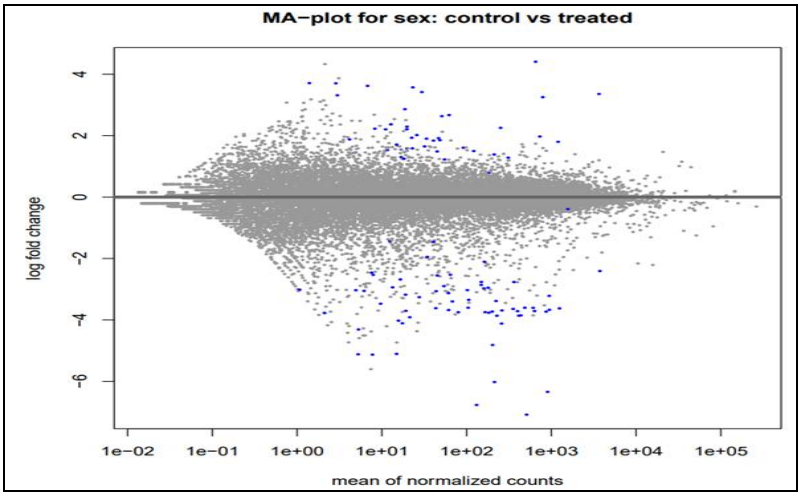


FIG. 2: VOLCANO PLOT FOR FEMALE AS CONTROL MALE AS TREATED

Using the online Venn diagram tool, we can find out how many genes are common and unique in male and female patients. In both conditions, 16 genes are in common. In the **Table 1** we see the total number of elements in male and female patients, as well as unique elements. The Venn diagram shown below **Fig. 3** relates to different conditions or groups in a biological context such as gene expression analysis. The numbers "35," "16," and "1" represent the number of unique features of each group and their intersection points. "Female versus male" and "male versus female" indicate comparisons between male and female patients. The Venn diagram usually visualizes the common and unique elements in both conditions. 35 elements unique to male patients, 16 genes common to both male and female patients, and 1 gene unique to female patients. This type of analysis is used to identify similarities and differences between different data sets or groups.

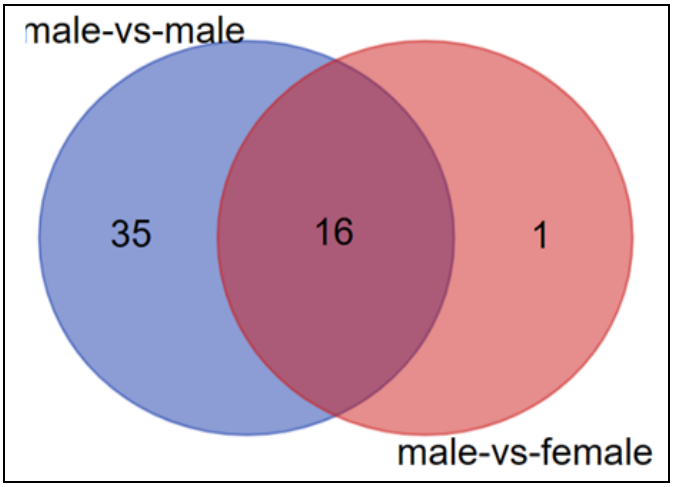


FIG. 3: VENN DIAGRAM FOR FEMALE-VS-MALE AND MALE-VS-FEMALE

TABLE 1: ELEMENT COUNTS IN GENDER COMPARISONS AND UNIQUENESS

List names	Number of elements	Number of unique elements
female-vs-male	51	51
male-vs-female	17	17
The overall number of unique elements		52

The table shows the number of elements in both female-vs-male and male-vs-female and the overall number of unique elements.

The role of these 16 common genes in immunological functions and pathways was assessed using GSEA and G-Profiler. Gene enrichment analysis was performed using GSEA **Table 2** and functional enrichment analysis was performed using G-Profiler for 16 unique genes in both genders. These Gene Ontology overlaps are shown here. There we see two macrophages and a T cell. The first column of the table shows the name of the gene set and the second column describes the up-and down-regulated genes, the number of overlapping genes, and the statistical significance. This report helps researchers understand the biological meaning and significance of different gene sets when analysing differential gene expression.

TABLE 2: GSEA-DATA MINING OF FUNCTIONAL GENE ENRICHMENT

Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	k/K	p-value?	FDRq-value?
GSE5099_CLASSICAL_M1_VS_ALTERNATIVE_M2 _M2_MACROPHAGE_DN [189]	Genes downregulated in macrophages: classical (M1) versus alternative (M2).	7		$1.92\ e^{-13}$	$2.96\ e^{-9}$
GSE5099_DAY3_VS_DAY7_MCSF_TREATED_MACR ACROPHAGE_DN [184]	Genes downregulated upon CSF1 [GeneID=1435] treatment: monocytes (3 days) versus macrophages (7 days).	5		$4.18\ e^{-9}$	$3.18\ e^{-5}$
GSE3982_MEMORY_CD4_TCELL_VS_BCELL_UP [199]	Genes up regulated in comparison of memory CD4 [GeneID=920] T cells versus B cells.	5		$6.19\ e^{-9}$	$3.18\ e^{-5}$

The "k/K" column quantifies the proportion of overlapping genes within each set, while the "p-value" evaluates the statistical significance of these overlaps. In addition, the "FDR-Q value" represents multiple tests. Lower p-values and FDR-q values indicate stronger statistical significance.

The importance of the ODAPH gene, which is present in both male and female patients shown in **Fig. 4**. This figure originates from Gene Set Enrichment Analysis (GSEA). The ODAPH gene is common in both male and female patients and plays an important role. The following list shows the various genes linked to the Y chromosome. However, ODAPH plays a special role in both genders. GSEA provides valuable information about molecular mechanisms common to both genders. Here we see the list of genes we have provided with the specific functions they perform and a description of the gene. Here we see the ODAPH gene and its description as a phosphoprotein associated with odontogenesis. Odonto means oral. We can therefore confirm that the ODAPH gene is related to Oral.

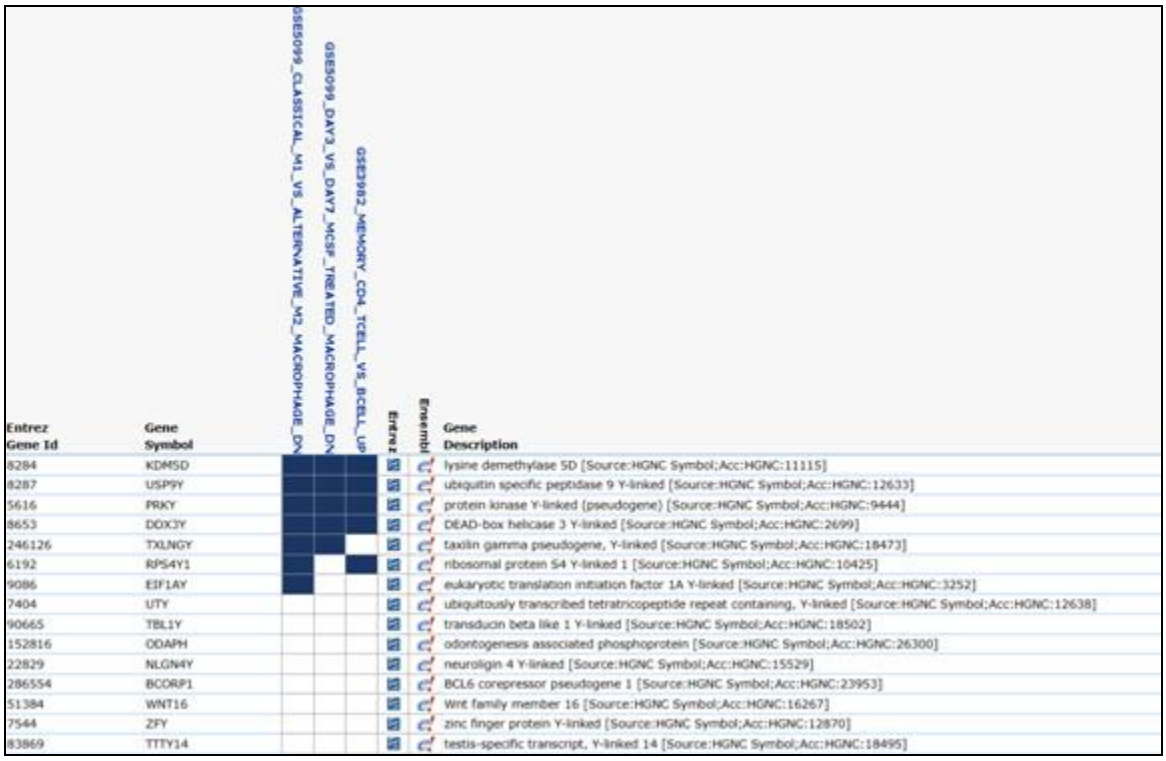


FIG. 4: ODAPH GENE: A PHOSPHOPROTEIN ASSOCIATED WITH ODONTOGENESIS

We see a list of different tissues and regions as well as gene identifiers and their descriptions derived from GSEA data **Fig. 5**. These tissues and regions represent different biological samples. The ODAPH gene is found in various tissues. Therefore, GSEA provides valuable information about differences in gene expression and the functional roles of these specific genes in different parts of the body. The expression profile of the selected genes is presented.

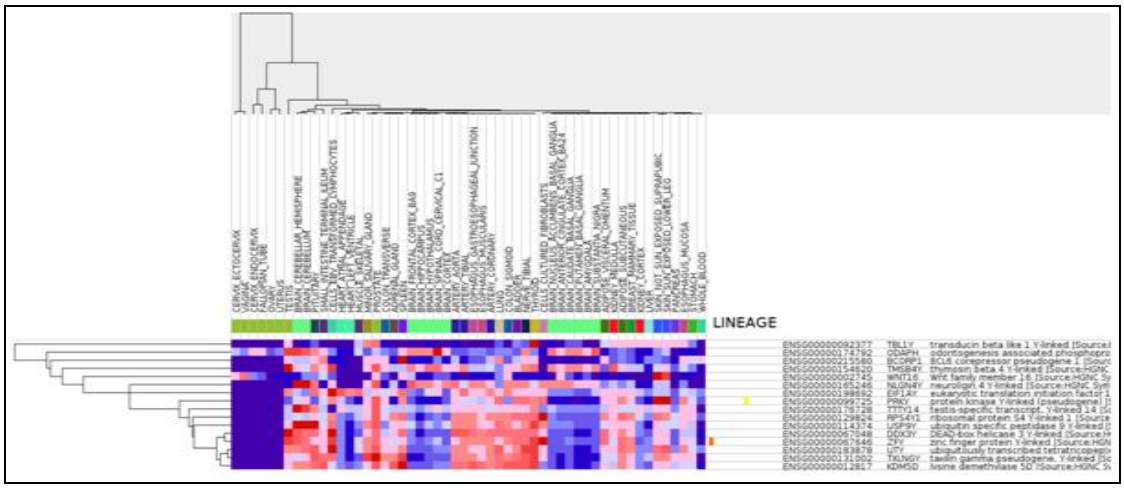


FIG. 5: EXPRESSION PROFILE FOR THE SELECTED GENES

G-Profiler is a tool for gene set enrichment studies. The following **Fig. 6**, generated by G-Profiler analysis, shows molecular function (MF) information from the Gene Ontology (GO) database. Each entry has an MF term name, a term size (i.e., the number of genes in the term), a unique term identifier, adjusted p-values, and a negative logarithm (base 10) of these p-values. They also represent statistical significance. In this way, genes are functionally enriched in specific biological pathways.

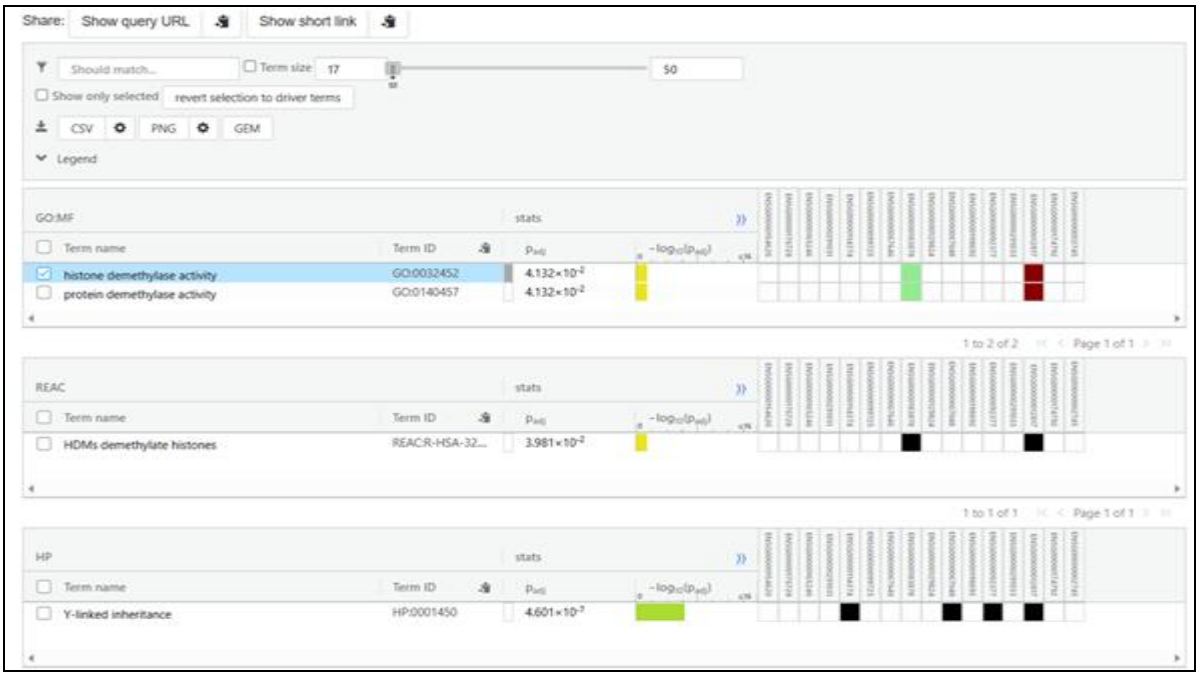


FIG. 6: MOLECULAR FUNCTION ANALYSIS BY G-PROFILER: GENE ONTOLOGY INSIGHTS

To validate the identified genes, data mining was performed using the cBioportal and TCGA databases. The following figure shows a list of cancer types and their associated genomic data. It provides access to various cancer data, including mutation data, variant data, and copy number variation data. This **Fig. 7** shows different types of cancer, such as E.g., uveal melanoma, testicular germ cell tumors, cholangiocarcinoma, and other types of cancer. This will then be useful for researchers to find genetic mutations in specific cancers. cBioportal is an important source of information for studying and understanding the specific characteristics of different types of cancer.

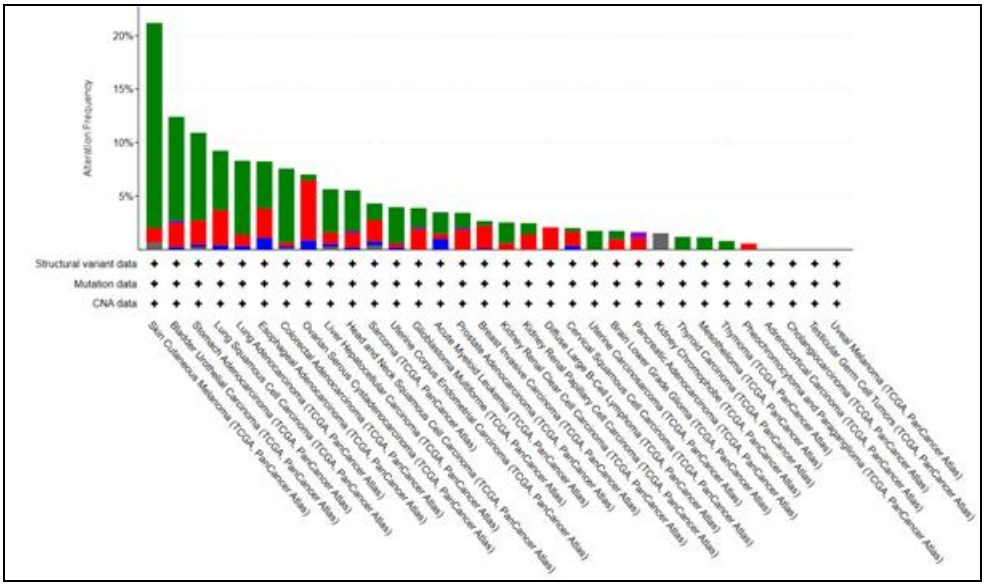


FIG. 7: VALIDATION OF IDENTIFIED GENES USING CBIOPORTAL: CANCER TYPES AND GENOMIC DATA

The **Fig. 8** generated from cBioportalis related to HNSCC mutation data. This shows frequency changes of 0.5% to 2.5%. This indicates the prevalence rates of specific genetic mutations or alterations in the HNSCC dataset. This alteration may include changes in DNA sequences that are associated with HNSCC progression. cBioportal is a valuable resource for researchers and clinicians to access and analyze genetic data. It helps them better understand the genetic factors that contribute to cancer and targeted treatment strategies. All of the genes listed below are found in head and neck squamous cell carcinoma. That is why they are mutated in head and neck squamous cell carcinoma.

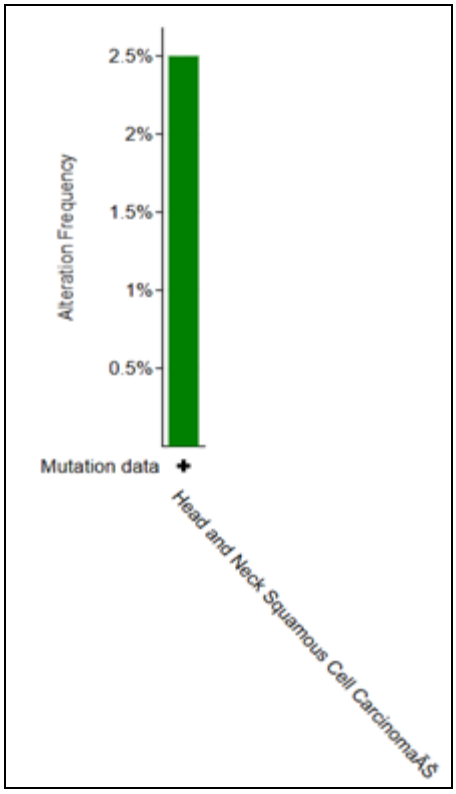


FIG. 8: HNSCC MUTATION DATA ANALYSIS FROM CBIOPORTAL

All potential gene targets were validated by checking gene expression patterns, evaluating mutated genes, and, most importantly, confirming their presence in the GDC-TCGA database. *In silico* data analysis was performed to obtain additional data for target validation. The following **Fig. 9** shows the distribution of cancer data from the GDC-TCGA database. It highlights 58 cases affected by 50 mutations and represents a specific type of cancer associated with specific genetic mutations. These mutations were observed in 18 different research projects. The "% of cases affected" column represents the proportion of cases within each project affected by these mutations. It provides information on the genetic profile of certain cancers and understands their complexity and molecular characteristics in various research studies.

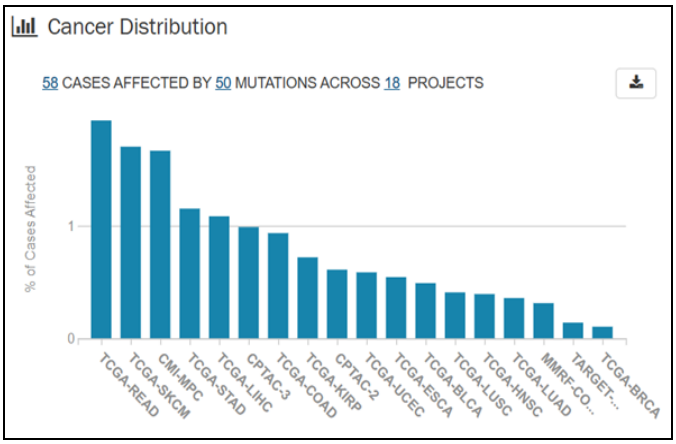
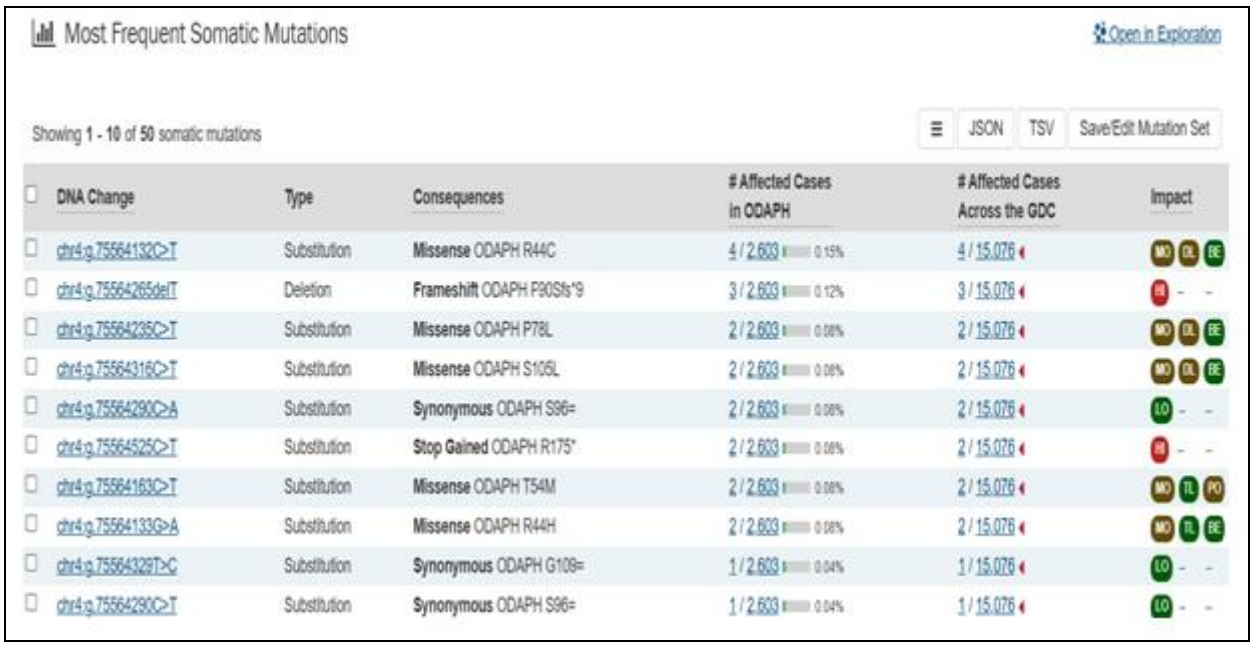


FIG. 9: DISTRIBUTION OF CANCER DATA FROM GDC-TCGA DATABASE

The below set of genes tends to mutate where all it appears like skin cancer, bronchus, lungs, bladder, stomach, kidney, brain, etc. The following figure, generated from the GDC-TCGA database, shows somatic mutations in a genomic dataset **Fig. 10**. It lists 10 out of 50 mutations, describing the type of mutations and their consequences, such as missense or frame shift mutations, as well as the number of cases in which the ODAPH gene is affected. Numbers such as “4/2,603” indicate the mutation prevalence of the affected cases in the ODAPH dataset, and “0.19%” represents their frequency. It all depends on the frequency of mutations.

It helps researchers provide information about common genetic changes in the data. Shown here is the mutation on chromosome 4 of the ODAPH gene. In addition, we can validate a small number of data in the laboratory or create a mutation in the oral cell line and evaluate the changes caused by this mutation. Then the normal cell becomes the oral squamous cell. Once the potential target is found, we can develop a drug against that target. This gene is common to both genders. In addition, there is a mutation in DNA that modulates the expression of genes in RNA. Therefore, we can validate this target for future experimental studies.



Most Frequent Somatic Mutations — Open in Exploration

Showing 1 - 10 of 50 somatic mutations — JSON | TSV | Save/Edit Mutation Set

DNA Change	Type	Consequences	# Affected Cases in ODAPH	# Affected Cases Across the GDC	Impact
chr4:g.75564132C>T	Substitution	**Missense** ODAPH R44C	4 / 2,603 0.15%	4 / 15,076	MO DL BE
chr4:g.75564265delT	Deletion	**Frameshift** ODAPH P90Sfs*9	3 / 2,603 0.12%	3 / 15,076	HI - -
chr4:g.75564235C>T	Substitution	**Missense** ODAPH P78L	2 / 2,603 0.08%	2 / 15,076	MO DL BE
chr4:g.75564316C>T	Substitution	**Missense** ODAPH S105L	2 / 2,603 0.08%	2 / 15,076	MO DL BE
chr4:g.75564290C>A	Substitution	**Synonymous** ODAPH S96=	2 / 2,603 0.08%	2 / 15,076	LO - -
chr4:g.75564525C>T	Substitution	**Stop Gained** ODAPH R175*	2 / 2,603 0.08%	2 / 15,076	HI - -
chr4:g.75564163C>T	Substitution	**Missense** ODAPH T54M	2 / 2,603 0.08%	2 / 15,076	MO TL PO
chr4:g.75564133G>A	Substitution	**Missense** ODAPH R44H	2 / 2,603 0.08%	2 / 15,076	MO TL BE
chr4:g.75564329T>C	Substitution	**Synonymous** ODAPH G109=	1 / 2,603 0.04%	1 / 15,076	LO - -
chr4:g.75564290C>T	Substitution	**Synonymous** ODAPH S96=	1 / 2,603 0.04%	1 / 15,076	LO - -

FIG. 10: SOMATIC MUTATIONS IN GENOMIC DATASET FROM GDC-TCGA DATABASE

Genes present in both genders were allowed to perform variant analysis using the Integrative Genomics Viewer tool. A gene with a common genetic mutation and varying gene expression patterns was identified. Bam files are visualized along with variants using IGV tools. We can see that the ODAPH gene is mutated in female patients. The RNA bam file is only from the coding regions, so this variation is valid in the coding regions. The ODAPH gene is located in the regulatory regions. They have no immune functions. In the figure below **Fig. 11** we can see that in the variant gene, the base is T and in the reference it is C. Then we can see variations in alignment. The following figure visualizes the genomes using the Integrative Genomics Viewer (IGV) tool. It focuses on human chromosome 4 (GRCh38/hg38) and displays the variant calling file from the Galaxy web server. It checks the position (ch4:75,564,112-75,564,155) in the genomic region of chromosome 4 with details including genetic variant, reference, alternate alleles, quality, type, and allele frequency. In this figure, we see the uploaded DNA and RNA Bam files. Overall, this figure helps researchers visualize and analyse genomic mutations and sequence data.

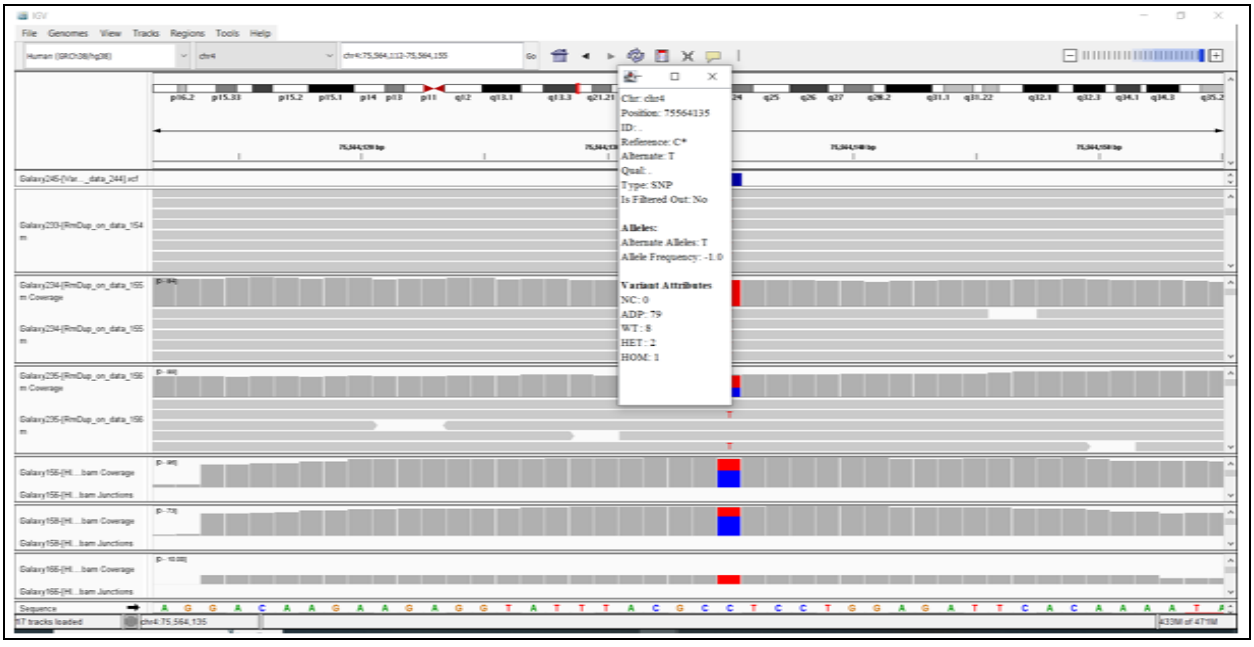


FIG. 11: FEMALE RNA AND DNA BAM FILES AND VCF FILE WITH ODAPH GENE

The following **Fig. 12** visualizes the genomic regions in male data using the IGV tool. It focuses on human chromosome 4 (GRCh38/hg38). The uploaded DNA and RNA bam files and variant files were also viewed. It focused on chromosome position chr4:75,564,116-75, 564, and 159. It highlighted the genetic mutation of a gene. It helps researchers learn more about genetic mutations in genomic regions. In summary, these data describe a comprehensive genome analysis with multiple data components.

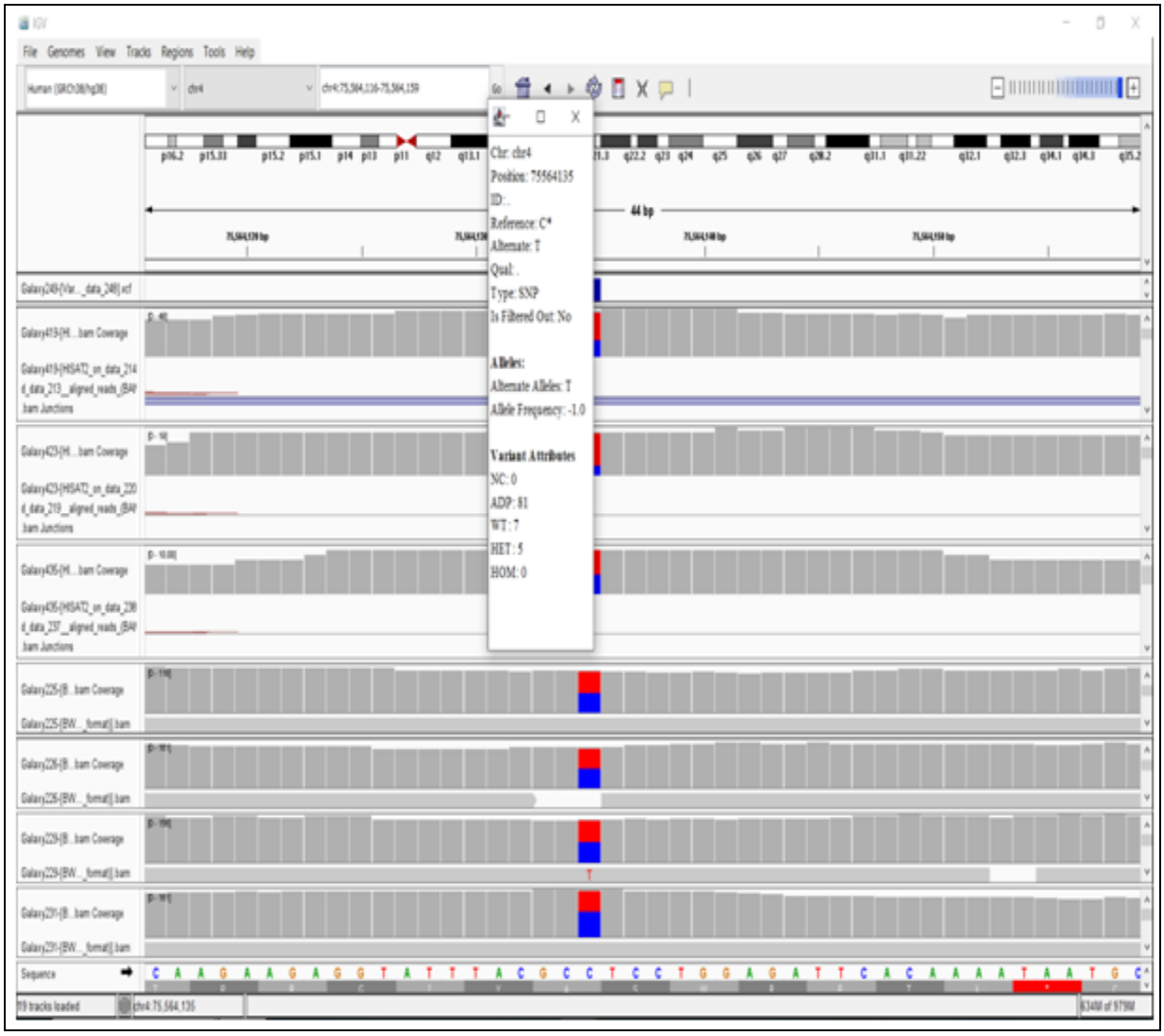


FIG. 12: MALE RNA AND DNA BAM FILES AND VCF FILE WITH ODAPH GENE

Using the STRING database, a protein network associated with this gene is identified **Fig. 13**. It shows a network of genes related to the ODAPH gene. Each gene name represents a protein. These genes are linked to ODAPH genes or related biological processes.

Other genes are functionally linked to or interact with ODAPH. Researchers use the STRING database to find protein-protein interactions and functional associations between related genes. It helps to identify the roles of genes and proteins in various diseases and biological pathways.

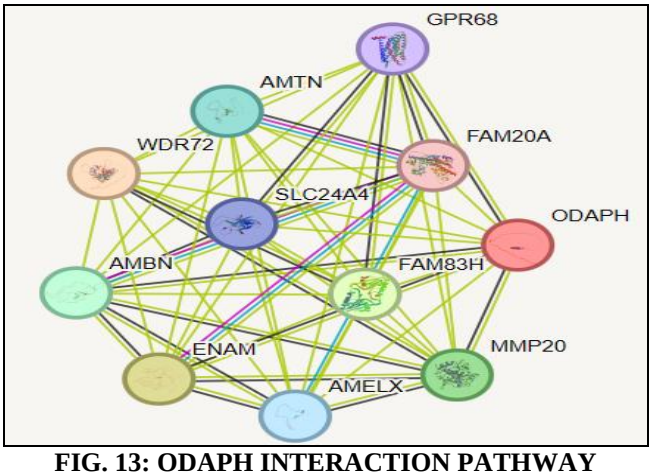


FIG. 13: ODAPH INTERACTION PATHWAY

The following **Fig. 14** provides insights into the ODAPH gene and its functional partners derived from the STRING database. The ODAPH gene is often associated with tooth enamel formation, particularly with the initiation of hydroxyapatite nucleation. Predicted functional partners include genes such as WDR72, ENAM, FAM83H, MMP20, AMBN, AMTN, AMELX, SLC24A4, GPR68 and FAM20A. All are involved in various aspects of enamel development, structural organization, and mineralization. These scores show the strength of these functional associations. This information sheds light on the molecular network involved in tooth enamel formation and provides valuable insights into dental biology and potential targets for dental health research.

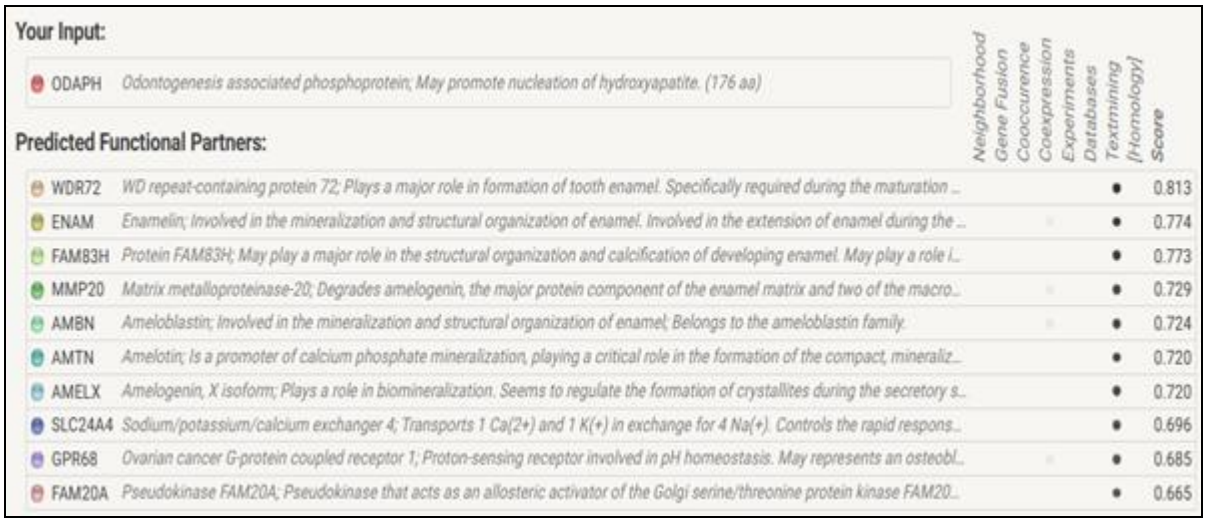


FIG. 14: ODAPH FUNCTIONAL PARTNERS

DISCUSSION: In future studies, we may opt for *in-vivo* validation, such as the synthesis of compounds for gender-neutral treatment options in patients with OSCC. The prevalence of Head and Neck Squamous Cell Carcinoma, particularly Oral Squamous Cell Carcinoma, requires extensive research and innovative treatment approaches to reduce its aggressiveness in the population. This study reviewed gene mutations and gene expression variations between male and female patients and aimed to provide insights into gender-neutral treatment options. Information from DNA and RNA data analysis is essential for understanding the molecular basis of OSCC progression.

WES data analysis helped to find genetic variations associated with OSCC, and RNA-seq analysis helped researchers study and understand gene expression patterns in OSCC. Studies have also shown the importance of genetic variations in cancer progression. Modulation of gene expression patterns is an important factor in the progression of OSCC. Changes in gene expression can drive cancer phenotypes. WES data analysis revealed specific genetic mutations associated with OSCC and shed light on the genetic basis of the disease. These mutations could serve as potential targets for therapeutic interventions. The presence of this gender-agnostic mutation suggests common molecular pathways in the pathogenesis of OSCC. One of the important findings of this research is the relationship between genetic mutation and changes in gene expression. Genetic alteration in response to specific genetic mutations demonstrates the interaction between genetic and epigenetic factors in the pathogenesis of OSCC. Understanding the change in gene expression patterns associated with genetic mutations is crucial for the development of targeted therapies.

A differential gene expression pattern analysis in male and female patients with OSCC revealed the specific altered gene that caused the altered gene expression. These genes may play a crucial role in the development and progression of OSCC. Differences in gene expression patterns in patients with OSCC provide valuable information for identifying biological mechanisms in OSCC. More importantly, we identified genes that exhibit high upregulation or downregulation across genders.

These genes, which are highly upregulated or downregulated in both genders, will be potential targets for treatment. This target will be an attractive target for future research and drug development efforts. Identification of common and unique genes in male and female patients with OSCC contributes to elucidating gender-specific aspects of the disease. Gene enrichment analysis further reveals the functional roles of these genes and highlights their involvement in immunological functions and specific signaling pathways relevant to OSCC. Data mining on larger datasets helps to validate the significance of the identified genes in OSCC. These findings strengthen research and treatment of OSCC. These genes are selected as potential targets through target validation, data analysis, and database confirmation. The identification of different gene expression patterns and their association with the protein network is a crucial finding. This gene represents a critical node in the molecular network of OSCC and is a promising target for future research and drug development. We can further translate these findings into clinical applications through *in vivo* validation. The development of compounds to treat OSCC in a gender-neutral manner will improve the lives of patients suffering from this difficult disease. This research not only advances the understanding of OSCC at the genetic and molecular levels but also opens the door to the development of gender-neutral precision medicine. We are getting closer to identifying altered genes in male and female patients and enhancing the lives of OSCC patients by developing personalized treatment strategies.

CONCLUSION: In summary, this study provides valuable insights into the treatment of OSCC by examining genetic alterations and differences in gene expression between male and female patients. The identified gene, potential therapeutic targets, and *in-vivo* validation provide hope for personalized medicine options for patients. This study represents a key step forward in improving the quality of life of OSCC patients worldwide.

ACKNOWLEDGEMENT: Nil

Funding: The authors have received no specific financial support from any funding agency or institution.

CONFLICT OF INTEREST: The authors declared no conflicts of interest.

REFERENCES:

1. Auperin A: Epidemiology of head and neck cancers: an update. Current Opinion in Oncology 2020; 32: 178–86.
2. "Head and neck cancers." National Cancer Institute. Published March 29, 2017. Accessed February 7, 2021.
3. World Health Organization. World cancer report 2014. Geneva: World Health Organization; 2014. Chapter 5.8. ISBN 978-9283204299.
4. Chow LQ: Head and neck cancer. The New England Journal of Medicine 2020; 382: 60–72.
5. National Cancer Institute. Oropharyngeal cancer treatment (adult) (PDQ®)–patient version. Published November 22, 2019. Accessed November 28, 2019.
6. Haines GK: Pathology of head and neck cancers I: epithelial and related tumors. In: Radosevich JA, editor. Head & neck cancer: current perspectives, advances, and challenges. New York: Springer Science & Business Media 2013; 257–87.
7. Al-Sarraf M: Treatment of locally advanced head and neck cancer: historical and critical review. Cancer Control 2002; 9: 387–99.
8. National Center for Biotechnology Information. Home - SRA.
9. European Bioinformatics Institute. Project page: PRJEB24758. Available from: https://www.ebi.ac.uk/ena/browser/view/PRJEB24758
10. Afgan E, Baker D and van den Beek M: The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 2018; 46(1): 544.
11. Andrews S and Fast QC: A quality control tool for high throughput sequence data. Accessed August 22, 2024. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
12. Ewels P, Magnusson M, Lundin S, Källers M and Multi QC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016; 32(19): 3047–3048.
13. Krueger F: Trim Galore. In: GitHub repository. GitHub; 2021. Available from: https://https://github.com/FelixKrueger/TrimGalore.com/fenderglass/Flye.
14. Li H: Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM tool. Accessed August 22, 2024. Available from: http://arxiv.org/abs/1303.3997
15. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G and Mesirov JP: Integrative Genomics Viewer. Nature Biotechnology 2011; 29: 24–26.
16. Robinson JT, Thorvaldsdottir H, Wenger AM, Zehir A and Mesirov JP: Variant review with the Integrative Genomics Viewer (IGV). Cancer Research 2017; 77(21): 31–34.
17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R: The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009; 25: 2078–2079.
18. Li H: A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. Bioinformatics 2011; 27: 2987–2993.
19. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L and Wilson RK: VarScan 2: Somatic mutation and copy number

alteration discovery in cancer by exome sequencing. Genome Research 2012; 22(3): 568–576.
20. Kim D, Langmead B and Salzberg SL: HISAT: a fast spliced aligner with low memory requirements. Nature Methods 2015; 12: 357–360.
21. Liao Y, Smyth GK and Shi W: feature Counts: an efficient general-purpose program for assigning sequence reads to genomic features. Bioinformatics 2013; 30: 923–930.
22. Love MI, Huber W and Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 2014; 15(12): 550.
23. Nature Protocols 2009; 4(1): 44 & Nucleic Acids Res. 2009; 37(1):1.
24. Draw Venn diagram. Ghent University. Available from: https://www.ugent.be/EN/venn-diagram
25. Subramanian, Tamayo: (2005, PNAS) and Mootha, Lindgren, et al. 2003, Nature Genetics.
26. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H and Vilo JG: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res 2019; 47(1): 98.
27. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO and Aksoy BA: The cBio Cancer Genomics Portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discovery 2012; 2(5): 401–404.
28. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C and Schultz N: Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science Signaling 2013; 6(269).
29. Genomic Data Commons. National Cancer Institute. Available from: https://gdc.cancer.gov/
30. STRING: functional protein association networks. Available from: https://string-db.org/

How to cite this article:
Lisina KV and Mittal B: Gender-based disparities in gene expression and mutation profiles in oral squamous cell carcinoma revealed by the whole exome sequencing and RNA sequencing. Int J Pharm Sci & Res 2025; 16(2): 417-29. doi: 10.13040/IJPSR.0975-8232.16(2).417-29.