 DOI: 10.21608/adjalexu.2025.342797.1570

BIOINFORMATICS ANALYSIS: A ROADMAP TO AN AGGRESSIVENESS MODEL FOR ORAL SQUAMOUS CELL CARCINOMA KEY DIFFERENTIATED GENES

Heba R.Mady[1*] MSc, Nefertiti Elnikhely[2] PhD, Samar El Achy[3] PhD, Enas M.Omar[4] PhD, Omneya R.Ramadan[5] PhD

This manuscript was presented as a poster at the Alexandria International Dental Congress 2024

ABSTRACT

INTRODUCTION: Bioinformatics is a cross-disciplinary research field that combines computer science with biological science. The most fundamental yet frequently most impactful application of bioinformatics analysis is identifying variations in gene expression across two or more circumstances in a process known as "differential expression". Identification of genes involved in oral squamous cell carcinoma initiation, progression and metastasis is a crucial breakthrough aiming at implementing targeted gene therapy as a promising oral cancer therapeutic modality.

OBJECTIVES: To identify key differentially expressed genes (DEGs) related to progression and aggressiveness of oral squamous cell carcinoma using bioinformatics analysis.

MATERIAL AND METHODS:

1. In the current study differential gene expression analysis has been done between low stage tumor cases (n=73) and high stages tumor cases (n=226) using the (DEseq2) package in R
2. GO enrichment and KEGG enrichment analysis of the significant DEGs were done using DAVID online tool and Enrich R online tool.
3. Selected hub genes of the protein-protein interaction network were obtained using the CytoHubba plugin MCC Algorithm

RESULTS:A total of 65 significant DEGs were detected

Functional enrichment analysis of the significant DEGs showed that:

The most enriched biological processes term was keratinization and defense response to bacteria

The most enriched cellular component term was the extracellular space

The most enriched molecular function term was cysteine-type endopeptidase inhibitor activity,

The most enriched KEGG terms was Salivary secretion.

CONCLUSION: In summary, hub genes discovered were the gene sequence of STATH, HTN3, HTN1, CST4, SMRB, LPO, BPIFA2.

KEY WORDS: Bioinformatics, OSCC, differentially expressed genes, hub genes

RUNNING TITTLE: Bioinformatics analysis: A roadmap to an aggressiveness model for Oral Squamous cell carcinoma key differentiated genes

---

1. Assistant Lecturer of Oral Pathology, Faculty of Dentistry, Alexandria University.
2. Assistant Professor of Biotechnology, Institute of Graduate Studies and Research, Alexandria University
3. Assistant Professor General Pathology, Faculty of Medicine, Alexandria University.
4. Assistant Professor of Oral Pathology, Faculty of Dentistry, Alexandria University.
5. Professor of Oral Pathology, Faculty of Dentistry, Alexandria University.

Corresponding author:
heba.ramadan.dent@alexu.edu.eg

INTRODUCTION

Oral squamous cell carcinoma (OSCC) remains a main cause deteriorating the quality of life of significant number of patients all over the world regardless of the massive advances in diagnostic and therapeutic strategies.(1) One of the main reasons of treatment resistance is the ability of the cancerous cells to express a countless number of mutated genes that play a role in their survival.(2) Oral squamous cell carcinoma patients have a 5-year survival rate of no more than 60% because of tumor growth, metastasis, and recurrence.(3)

For cancer patients, whole genome sequences have offered a fresh path for researchers to investigate these individuals at the genetic level. This knowledge has enabled the execution of high-throughput, genome-wide screening for gene functions. (4)

Bioinformatics is a cross-disciplinary research field that combines computer science with biological science; in other words, bioinformatics is a union of biology and informatics. It refers to computer technology used to store, retrieve, manipulate, and analyze the information of the biological macromolecules such as DNA, RNA, and proteins. (5)

The most fundamental yet frequently most impactful application of bioinformatics analysis is

identifying variations in gene expression across two or more circumstances in a process known as "differential expression". The results are commonly referred to as differentially expressed genes (DEGs), which may be either upregulated or downregulated. [6] As a result of such analysis, several research have been conducted to be of use in fields of cancer diagnosis, prognosis and in cancer treatment. [7-10] Identification of genes involved in carcinogenesis, progression and metastasis is a crucial breakthrough aiming at implementing targeted gene therapy as a promising oral cancer therapeutic modality. [11]

The null hypothesis of this study is that there are no key differentially expressed genes between low stages and high stages samples of oral squamous cell carcinoma.

MATERIALS & METHODS

1. Data acquisition

This study represents differential gene expression analysis on publicly available data of GDC head and neck TCGA project, HTSeq - Counts workflow of RNA-seq of head and neck cancer data set (n=546) were downloaded from The Cancer Genome Atlas database (TCGA) Xena tool (https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Head%20and%20Neck%20Cancer%20(HNSC)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443).

With the recommended genome mapper https://gdc-hub.s3.us-east-1.amazonaws.com/download/gencode.v22.annotation.gene.probeMap; Full metadata

And associated phenotable https://gdc-hub.s3.us-east-1.amazonaws.com/download/TCGA-HNSC.GDC_phenotype.tsv.gz; Full metadata

OSCC RNA-seq were selected from the downloaded data set as follow:

(a) Primary sites: tongue, hard palate, lip, gum, floor of the mouth, cheek, base of the tongue, mandible, other unspecified parts of oral cavity

(b) Disease type: squamous cell neoplasms

(C)Workflow type: HTSeq – Counts

(D) sample type: primary tumor

Finally, a total of 328 OSCC primary tumor samples with gene expression data and corresponding clinical information were utilized for this study. Considering cases with stage I and stage II as low stage tumor and cases with stage III and stage IV as high stage tumors, data were further sorted as 226 cases with high stage, 73 cases with low stage, and 29 cases without declared stage information.

2. Exploratory analysis of data

Screening of the data has been done and we found that the downloaded HTSeq – Counts were processed to be (log2(count+1)) so the first step was to unlog (count)-1 to return the read count of HTSeq to be analyzed further by using (DESeq2) Bioconductor package in R (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) .

To filter the lowest variance genes, the Standard Deviation Score (SDS) were calculated for all the mapped gene Using matrix state package(CRAN: Package matrixStats (r-project.org) , selection of the top20000 sorted genes are done into which differentially expression analysis will be caried out.[12]

To reveal the similarities between the two study groups, data were transformed to tidy data and histogram was done using ggplot2 package in R https://ggplot2.tidyverse.org/ to detect the outliers in the study groups. [13]

PCA plot have been made using prcomp function in R to detect the outliers in the study groups before the differentiation step. [13]

3.Differential genes (DEGs) Expression Analysis

Differential gene analysis between Low stages samples and high stages samples in the selected dataset was done using the (DEseq2) package in R. The results were selected according to adjusted P-value <0.01 as screening criteria for differential genes, log FC> 3 as upregulated and log FC<3 as down regulated genes. A total of 65 DEGs were obtained. 55 genes were down regulated and only 10 genes were up regulated. [7]

4. Functional enrichment analysis

The Functional enrichment analysis of the 65 differentially expressed genes was done by DAVID online tool (http://david.ncifcrf.gov). where, FDR less than 0.05 and minimum enrichment of 3 were used as the screening criteria. graphical representations such as bar, dot, and chord plots were constructed. This was accomplished using the "GO plot" package (https://bioconductor.org/packages/GO_plot/) in R software (version 4.21), and on SR online tool http://www.bioinformatics.com.cn/srplot ensuring a clear and interpretable presentation of the outcomes.[10]

5. Protein interaction network of DEGs

Protein protein interaction analysis with screening of 65 DEGs was done into the String databasehttps://stringdb.org/cgi/input?sessionId=bt4DaRcSwMex&input_page_show_search=on to obtain the PPI network relationships of DEGs.[14]

6. Identification of hub genes

Selected hub genes of the protein-protein interaction network were obtained using the CytoHubba plugin MCC Algorithm. [15]

RESULTS

1.Clinical results

1.1 Concerning demographic results

In the present study, A total of 328 OSCC data files were chosen from the head and neck cancer dataset that was downloaded. Clinical analysis of the data show that gender distribution to be as follows: 69%

of the cases included were from males (n=225) and 31% were females (n=103)

1.2 Concerning the anatomical site

The tongue was the most frequently observed site of incidence representing 48% of the cases (n=159), followed by the other non-specified part of oral cavity representing 26% of the cases (n=88). Followed by the floor of the mouth representing 15% of the cases (n=53). Following was the cheek mucosa 5% of the cases (n=19), the gum representing 2.5% of the cases (n=9). the hard palate, lips and mandible are represented by 1.5%&1%&1% respectively (n=5), (n=3), (n=2).

1.3 Concerning the Pathological stages of the tumor samples,

According to pathological staging of the tumor, we sorted the data as follow:

Stage I and stage II as low stage tumor sample

Stage III and stage IV as high stage tumor sample

This resulting in 68% of the cases (n=226) were categorized as high stage tumor samples while 22% of the cases (n=73) were categorized as low stage tumor samples.

2 Exploratory analysis and data processing results

To reveal the similarities between the two study groups, data were transformed to tidy data and histogram was done using ggplot2 package in R https://ggplot2.tidyverse.org/. 3D PCA plot has been made using prcomp function in R. Histogram & PCA show no outliers between the 2 groups of the study (Figure.1)

3. Identification of differentially expressed genes

The differential gene analysis between low stage samples versus high stage samples across the OSCC samples in the selected dataset was done using the (DEseq2) package in R, resulting in a total of 1520 differentially expressed genes. DEGs. A total of 65 DEGs were obtained. 55 genes were down regulated and only10 genes were up regulated. (Figure.2) Circular heatmap showing the expression of the significant 65 DEGs between the 2 groups of comparison is illustrated in (Figure.3)

4. Identification and functional annotation of DEGs in Go term analysis

Using the DAVID web tool (http://david.ncifcrf.gov), the functional enrichment analysis of the 65 differentially expressed genes was completed. where the screening criteria were a minimum enrichment of three genes and a false discovery rate (FDR) of less than 0.05.

Concerning biological processes, four GO terms were significant

GO:0031424~keratinization

GO:0042742~defense response to bacterium

GO:2000117~negative regulation of cysteine-type endopeptidase activity

GO:0001580~detection of chemical stimulus involved in sensory bitter taste precipitation

Concerning cellular component, two GO terms were significant

GO:0005615~extracellular space

GO:0005576~extracellular region

And Concerning Molecular function, two GO terms were significant

GO:0004869~cysteine-type endopeptidase inhibitor activity

GO:0004866~Endopeptidase Inhibitor Activity

Chord plot displaying gene-pathway interactions can help identify relevant genes and their roles in important biological processes, cellular components, and molecular function Go terms is shown in (Figure.4)

GO pathway enrichment analysis for the three terms for significant DEGs is illustrated in bar plot and bubble plot respectively in (Figure.5).

5.Functional analysis of significant DEGs via pathway analysis.

KEGG pathway analysis of the significant differentially expressed genes was done by EnrichR online tool (https://maayanlab.cloud/Enrichr/enrich). where, adjusted p-value less than 0.05 and minimum enrichment of 3 used as screening criteria show that the most enriched KEGG terms was Salivary secretion with all the 8 genes associated were downregulated indicating that decrease salivary secretion might be considered as a risk factor for progression of oral cancer

6.Protein-protein interaction network

The STRING database program was used to analyze the protein-protein interaction network between differentially expressed genes, creating 55 nodes and 50 edges and yielding two significant clusters using the k-means clustering method: salivary secretion and keratinization.

7. Identification of hub genes

The CytoHubba plugin's MCC Algorithm was used to extract a subset of the protein-protein interaction network's hub genes. The MCC Algorithm indicates that the gene sequences of STATH, HTN3, HTN1, CST4, SMR3B, LPO, and BPIFA2 are among the top 7 genes in the network. Where STATH, HTN3, HTN1 are the 3 most representative hub genes respectively. (Figure.6)

DISCUSSION

In The present study, the differential gene analysis was done using the (DEseq2) package in R between low stage samples versus high stage samples across the 328 OSCC samples, resulting in a total of 65 significant DEGs where 55 genes were down regulated and only 10 genes were up regulated.

It is crucial to emphasize that the highest enriched biological terms in functional enrichment analysis were keratinization and defense response to bacteria. where the expressed genes (LCE2B, LCE2C, LCE2A, KRT2, LCE2D) and (HTN3,

BPIFA2, LPO, STATH, HTN1) are the significant DEGS in each term respectively.

In oral epithelial cell lines, Sakamoto et al. and Morifuji et al. found that in vitro transfection of different keratins resulted in changes in cell shape and motility, suggesting that these changes are caused by dysregulation of cell differentiation and can be used to prognosticate oral cancers. [16, 17] Furthermore, Chen et al suggest that KRT2 downregulation could be looked at as a viable marker for identifying high-risk groups for gastric cancer by altering the microbiome. [18]

The preservation of physiological homeostasis, adaptation to environmental changes, and survival all depend on the proper balance between commensal bacteria and the host. [19] Recent research has shown the potential for pathogenic microorganisms linked to periodontal disease to have a significant impact in the aggressiveness and carcinogenesis of oral cancer. Pignatelli et al demonstrated that *T. denticola*, *P. gingivalis*, and *F. nucleatum* promote OSCC migration, invasion, and tumor sphere formation through integrin alpha V/focal adhesion kinase signaling. Additionally, the authors showed that T. denticola contributes to the aggressive nature of the oral squamous cell carcinoma phenotype by triggering cell migration through crosstalk between toll-like receptors 2 and 4 and integrin alpha V/focal adhesion kinase signaling. [20]

Within the enrichment of cellular components, we found that 19 genes were mainly involved in the extracellular space, these genes include sequences of (IL22, ZG16B, KRT2, GP2, PRH2, LPO, IGFL1, CST5, DLK1, CST4, PRB4, CST2, SMR3B, BPIFA2, PRB3, FGF19, SCGB3A2, SCGB2A2, CA6). The extracellular matrix (ECM) is a complex network structure consisting of macromolecules produced by tumor cells into the extracellular space. The regulatory abnormality of ECM has a prominent role in orchestrating the Tumor microenvironment. Numerous distinct signaling pathways that aid in cell invasion and proliferation can be activated by the interaction between tumor cells and extracellular matrix. Thus, it is helpful to have a sufficient understanding of ECM dysregulation in order to discover possible targets for tumor treatment. [21]

The analysis of molecular functions showed that 3 genes are involved in the regulation of cysteine-type endopeptidase inhibitor activity. These genes include sequences of (CST2, CST4, CST5). The cystatin (CST) superfamily is a kind of thiol proteinase inhibitor (TPI) that is extensively prevalent in human tissues and body fluids and functions as competitive and reversible inhibitors of cysteine proteases. Numerous studies have demonstrated that CST2, CST4, and CST5 are essential for numerous pathological processes and impact different phases of carcinogenesis and tumor development, such as angiogenesis, invasion, metastasis, apoptosis, and proliferation in breast, colon, and stomach cancer. [22-24] CST5 expression is stimulated by vitamin D, which has anticancer effect and is downregulated in human colon cancer cells. CST5 extends the cell cycle and decreases the proliferation, migration, and invasiveness of colorectal tumor cells by blocking the Wnt/β-catenin signaling pathway and oncogenic c-MYC expression. [25,26]

On the other hand, the only significant KEGG pathways in this study was salivary secretion with all the 9 genes associated being downregulated indicating that decrease salivary secretion might be considered as a risk factor for progression of oral cancer. Yuan et al. explained that, because high-stage patients are more likely to receive radiation therapy these patients have significantly lower salivary secretion levels. This is due to a combination of factors, including chronic inflammation and cytokine effects that disrupt normal secretory signaling in the salivary glands, beside the destruction of salivary gland secretory cells concluding that this is a research area requiring further studies. [27]

The top seven hub genes in the protein-protein interaction network generated by the CytoHubba plugin MCC Algorithm were STATH, HTN3, HTN1, CST4, SMR3B, LPO, and BPIFA2. STATH, HTN3, CST4, and HTN1 are the most prominent hub genes, respectively.

The protein statherin is encoded by STATH, which is derived from a cluster of genes for secretory calcium-binding phosphoproteins. It was said to be in charge of keeping calcium phosphate from precipitating in saliva, which results in consistently elevated calcium and phosphate levels. [28] When compared to healthy controls, the amounts of statherin in saliva are considerably lower in cancerous and precancerous lesions. Reduced STATH expression indicates less free calcium, which may result in the onset of persistent proliferation, desmosome decrease, and changes in calcium-associated cellular processes. diminish antimicrobial activity against infections, emphasizing the significance of oral microbiota in the beginning and progression of oral cancer. [29]

Histatins comprise a group of natural anti microbial peptides that play significant roles in nu merous biological systems, including antifungal an d antibacterial activities, cancer development, imm une modulation, and wound repair. Histatin1 (Hst1) is essential for epithelial wound repair and cell mo vement, serving as the primary healing component found in saliva. There is limited understanding of t he downstream mechanisms that facilitate the impa ct of histatins on cancer development and progressi on [30] Toward an explanation, Dijk IA et al conclude that HTN1 protein strengthened the endothelial

barrier, reduced its permeability to large molecules, and prevented bacterial migration across epithelial cell layers through the action of the adherens junction protein E-cadherin (E-cad) and the tight junction protein zonula occludens1. [31] While Wassapol et al propose that HTN1 and HTN3 proteins counteracted the effects of epithelial-mesenchymal transition inducers on the formation of oral cancer cell spheroids. He proposed that the downregulation of both genes affects pathways associated with cancer development and that further research is needed to establish how both genes play a role in the onset and metastasis of OSCC. [32]

Cystatin S (CST4) is a secretory protein that plays a significant and multifaceted role in the process of tumor growth by serving as an inhibitor of cysteine protease activity. The inhibition exerted by CST4 has profound effects on the breakdown of the extracellular matrix (ECM), a crucial component that provides structural and biochemical support to surrounding cells. By influencing ECM degradation, CST4 modifies the tumor microenvironment in ways that can impact several key characteristics of tumor cells, including their adhesion to surfaces, their proliferation rate, and their ability to invade surrounding tissues through migratory processes. Despite the recognized potential of CST4 in contributing to tumor biology, it is noteworthy that the majority of the scientific research conducted thus far has placed emphasis on its potential diagnostic applications. This has led to a significant oversight in exploring the prognostic significance of CST4, which remains largely unexamined in the current literature. As such, understanding the full range of CST4's roles, particularly in relation to prognosis, warrants further investigation in order to better delineate its contributions to cancer progression and potential therapeutic implications. [33]

The LPO gene encodes for an enzyme found in saliva that exhibits bactericidal properties, playing an indispensable role in safeguarding the lactating as well as the intestinal tract of newborn infants against a variety of pathogenic microorganisms that could compromise their health. Beyond its antimicrobial properties, the lactoperoxidase enzyme has been reported in scientific studies to possess additional functionalities, including promoting cellular growth and exhibiting anti-tumor activity, thus highlighting its multifaceted roles within biological systems. [34] In a study conducted by Dhanya M, the expression levels of LPO were examined in normal tissue, in conditions characterized by oral premalignancy, and in cases of oral squamous cell carcinoma, revealing a significant elevation of LPO expression within the disease groups when compared to normal tissue samples; however, within the disease groups themselves, no significant differences in LPO levels were observed. Based on the findings of this particular study, it has been suggested that LPO could serve as a valuable biomarker as well as a screening tool to evaluate the extent and severity of cellular damage in patients suffering from potentially malignant oral disorders, thereby providing a useful diagnostic avenue. [35]

There exists a continually growing body of scientific literature that provides compelling evidence suggesting a significant relationship between the aberrant expression of genes responsible for encoding opiorphin—specifically, PROL1, SMR3A, and SMR3B—and the subsequent onset of various forms of malignancies, which is a matter of considerable importance in the field of oncology. This intriguing correlation was first posited in an extensive meta-analysis published in the year 2008, wherein rank aggregation techniques were employed to discern opiorphin-encoding genes among the top 50 genes that consistently exhibited altered expression levels across a wide array of cancer types, thereby highlighting the potential relevance of these genes in the pathogenesis of neoplastic diseases. In the years following this initial discovery, numerous studies have demonstrated that the altered expression of opiorphin-encoding genes is not only prevalent but is also significantly correlated with specific malignancies, including invasive breast carcinoma, head and neck adenoid cystic carcinoma, and oropharyngeal squamous cell carcinoma, thus reinforcing the importance of these genes in cancer research. [36]

Moreover, it has been observed that opiorphin-encoding genes function as master regulators of the hypoxic response, which is a critical adaptive mechanism for cells subjected to low oxygen levels; this regulation is essential as it activates various biological pathways that are associated with enhanced blood perfusion and the process of angiogenesis, both of which are fundamental to tumor growth and metastasis. Consequently, the implications of these findings may extend beyond mere correlation, suggesting that the dysregulation of opiorphin-encoding genes could play a pivotal role in the development and progression of cancer, thereby warranting further investigation into their functional mechanisms and potential as therapeutic targets. In conclusion, the expanding corpus of evidence surrounding the relationship between opiorphin-encoding genes and malignancies underscores the necessity for continued research in this domain to fully elucidate the complex interplay between gene expression and cancer pathophysiology. [37]

BPIFA2, a soluble salivary protein that is released in the extracellular space of the salivary glands, particularly the parotid gland, is a member of the PLUNC protein family. Human chromosome 20 has the encoding gene with 10 exons and a 750 bp CDS (protein coding region) region that codes for a 250 amino acid protein sequence. BPIFA2, A saliva surfactant regulates saliva's surface tension by

binding to lipoproteins and inhibiting bacterial growth. [38] Recent studies indicate that BPIFA2 may interact with the host microbiome, serving as a biomarker for various diseases. For instance, salivary BPIFA2 levels were lower in periodontitis patients, but higher in those with mycobacterial infections or CMV. In developing a diagnostic model for oral squamous cell carcinoma (OSCC) using principal component analysis (PCA) and differentially expressed genes (DEGs), Lingdu et al. found that the OSCC model, based on the expression of SMR3B, PRR27, HTN3, STATH, CST5, BPIFA2, PRH2, KRT35, HTN1, and AMY1B, demonstrated significant diagnostic potential. [39]

CONCLUSION

Using bioinformatics tools in conjunction with saliva transcriptomics, proteomics, metabolomics and salivary microbiome help to understand the tumor microenvironment of OSCC and unravels molecular signatures of certain genes, proteins, cytokines, bacteria and peptides toward initiation, progression and lymph node metastases in OSCC.

FUNDING

This research is a part of a project funded by the science, technology & innovation funding Authority (STDF), Project ID (44982)

CONFLICT OF INTEREST

The authors declared that they have no conflicts of interest.

REFERENCES

1. Sarode G, Maniyar N, Sarode SC, Jafer M, Patil S, Awan KH. Epidemiologic aspects of oral cancer. Disease-a-Month. 2020 Jun 27:100988.
2. Warnakulasuriya S, Greenspan JS, editors. Textbook of Oral Cancer: Prevention, Diagnosis and Management. Springer Nature; 2020 Feb 22.
3. rimie AI, Ciocan C, Gulei D, Mehterov N, Atanasov AG, Dudea D, BerindanNeagoe I. Current Insights into oral cancer epigenetics. International journal of molecular sciences. 2018 Mar;19(3):670.
4. Cano-Gamez E, Trynka G. From GWAS to function: using functional genomics to identify the mechanisms underlying complex diseases. Frontiers in Genetics. 2020;11.
5. Li Y, Huang C, Ding L, Li Z, Pan Y, Gao X. Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. Methods. 2019 Aug 15;166:4-21.
6. Abbas SZ, Qadir MI, Muhammad SA. Systems-level differential gene expression analysis reveals new genetic variants of oral cancer. Scientific reports. 2020 Sep 4;10(1):1-9.
7. Lu W, Li N, Liao F. Identification of key genes and pathways in pancreatic cancer gene expression profile by integrative analysis. Genes. 2019 Aug;10(8):612.
8. Zhang L, Tan X, Lin Y, Han X, Wang H, Ming H, Li Q, Liu K, Feng G. Systematic analysis of genes involved in oral cancer metastasis to lymph nodes. Cellular & molecular biology letters. 2018 Dec;23(1):1-4.
9. Coon, Jeffery, Karl Kingsley, and Katherine M. Howard. "miR-365 (microRNA): Potential biomarker in oral squamous cell carcinoma exosomes and extracellular vesicles." International Journal of Molecular Sciences 21.15 (2020): 5317.
10. Qiao B, Zhao M, Wu J, Wu H, Zhao Y, Meng F, Tian Y, Wang S, Shi J, Zhang H. A Novel RNA-Seq-Based Model for Preoperative Prediction of Lymph Node Metastasis in Oral Squamous Cell Carcinoma. BioMed Research International. 2020 Aug 31;2020.
11. Gallagher MD, Chen-Plotkin AS. The post-GWAS era: from association to function. The American Journal of Human Genetics. 2018 May 3;102(5):717 30.
12. Berger, Alexander, and Markus Kiefer. "Comparison of different response time outlier exclusion methods: A simulation study." Frontiers in psychology 12 (2021): 675558.
13. Yu, Lu, et al. "Identification of SPRR3 as a novel diagnostic/prognostic biomarker for oral squamous cell carcinoma via RNA sequencing and bioinformatic analyses." PeerJ 8 (2020): e9393.
14. Zhang, Liru, et al. "Identification and validation of transcription factor-driven enhancers of genes related to lipid metabolism in metastatic oral squamous cell carcinomas." BMC Oral Health 22.1 (2022): 126.
15. Chin CH, Chen SH, Wu HH, et al. cytoHubba: Identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014, 8, 1–7.
16. Sakamoto, K.; Aragaki, T.; Morita, K.; Kawachi, H.; Kayamori, K.; Nakanishi, S.; Omura, K.; Miki, Y.; Okada, N.; Katsube, K.; et al. Down-regulation of keratin 4 and keratin 13 expression in oral squamous cell carcinoma and epithelial dysplasia: A clue for histopathogenesis. Histopathology 2011, 58, 531–542.
17. Vaidya, Milind, Crismita Dmello, and Saie Mogre. "Utility of keratins as biomarkers for human oral precancer and cancer." Life 12.3 (2022): 343.
18. Chen, Jiahui, et al. "In-depth metaproteomics analysis of tongue coating for gastric cancer: a multicenter diagnostic research study." Microbiome 12.1 (2024): 6.
19. Li, Z.; Liu, Y.; Zhang, L. Role of the Microbiome in Oral Cancer Occurrence, Progression and Therapy. Microb. Pathog. 2022, 169, 105638.
20. Pignatelli, Pamela, et al. "Microbiota and oral cancer as a complex and dynamic

microenvironment: a narrative review from etiology to prognosis." International journal of molecular sciences 23.15 (2022): 8323.
21. Eble, Johannes A., and Stephan Niland. "The extracellular matrix in tumor progression and metastasis." Clinical & experimental metastasis 36.3 (2019): 171-198.
22. .Zhang, W.P.; Wang, Y.; Tan, D.; Xing, C.G. Cystatin 2 leads to a worse prognosis in patients with gastric cancer. J. Biol. Regul. Homeost. Agents 2020, 34, 2059–2067
23. 23.Song, F.; Zhang, Y.; Pan, Z.; Hu, X.; Yi, Y.; Zheng, X.; Wei, H.; Huang, P. Identification of novel key genes associated with the metastasis of prostate cancer based on bioinformatics prediction and validation. Cancer Cell. Int. 2021, 21, 559.
24. Shi, D.; Zhou, Z.; Zhang, S. miRNA-6715-5p Inhibits Cellular Proliferation and Invasion in Colorectal Cancer by Directly Targeting CST4. J. Oncol. 2021, 2021, 7615712.
25. Berger, M.D.; Stintzing, S.; Heinemann, V.; Cao, S.; Yang, D.; Sunakawa, Y.; Matsusaka, S.; Ning, Y.; Okazaki, S.; Miyamoto, Y.; et al. A Polymorphism within the Vitamin D Transporter Gene Predicts Outcome in Metastatic Colorectal Cancer Patients Treated with FOLFIRI/Bevacizumab or FOLFIRI/Cetuximab. Clin. Cancer Res. 2018, 24, 784–793.
26. Munetsuna, E.; Kawanami, R.; Nishikawa, M.; Ikeda, S.; Nakabayashi, S.; Yasuda, K.; Ohta, M.; Kamakura, M.; Ikushiro, S.; Sakaki, T. Anti-proliferative activity of 25-hydroxyvitamin D3 in human prostate cells. Mol. Cell Endocrinol. 2014, 382, 960–970.
27. Yuan, Hao, et al. "Efficacy of kangfuxin liquid on radiotherapy-induced oral mucositis for patients with head and neck squamous cell carcinoma and its effect on salivary glands and immune function." American Journal of Translational Research 14.9 (2022): 6792.
28. Wu Q, Cao R, Chen J, Xie X. Screening and identification of biomarkers associated with clinicopathological parameters and prognosis in oral squamous cell carcinoma. Exp Ther Med. 2019 Nov;18(5):3579-3587.
29. Sehrawat, Kartik. "New perspectives of human saliva as genetic biomarker in early disease diagnostics." Biomedical and Therapeutics Letters 11.1 (2024): 907-907.
30. Son KN, Lee H, Shah D, Kalmodia S, Miller RC, Ali M, Balasubramaniam A, Cologna SM, Kong H, Shukla D, Aakalu VK. Histatin-1 is an endogenous ligand of the sigma-2 receptor. FEBS J. 2021 Dec;288(23):6815-6827.
31. van Dijk IA, Ferrando ML, van der Wijk AE, Hoebe RA, Nazmi K, de Jonge WJ, Krawczyk PM, Bolscher JGM, Veerman ECI, Stap J. Human salivary peptide histatin-1 stimulates epithelial and endothelial cell adhesion and barrier function. FASEB J. 2017 Sep;31(9):3922-3933.
32. Wassapol Wongpanuwich, Somchai Yodsanga et al. Association Between PD-L1 and Histatin1, 3 Expression in Advanced Head and Neck Squamous Cell Carcinoma, Anticancer-Research May2022, 42 (5) 2689-2699.
33. Zhang, Zijun, and Fenghuang Zhan. "Type 2 Cystatins and Their Roles in the Regulation of Human Immune Response and Cancer Progression." Cancers 15.22 (2023): 5363.
34. Magacz M, Kędziora K, Sapa J, Krzyściak W. The Significance of Lactoperoxidase System in Oral Health: Application and Efficacy in Oral Hygiene Products. International Journal of Molecular Sciences. 2019;20(6):1443.
35. M D, Kumar Vadivel J, T NU, Jayaraman S. Comparative Assessment of Lipid Peroxidase in Oral Cancer and Oral Potentially Malignant Disorders. Cureus. 2024;16(8):e66474.
36. Mukherjee, Amarnath, et al. "Role of opiorphin genes in prostate cancer growth and progression." Future Oncology 17.17 (2021): 2209-2223.
37. Wu, Qiqi, et al. "Screening and identification of biomarkers associated with clinicopathological parameters and prognosis in oral squamous cell carcinoma." Experimental and Therapeutic Medicine 18.5 (2019): 3579-3587.
38. He, Lexin, et al. "BPIFA2 as a Novel Early Biomarker to Identify Fatal Radiation Injury After Radiation Exposure." Dose-Response 20.1 (2022): 15593258221086478.
39. WEN, Lingdu, et al. "Construction of a diagnostic model for oral squamous cell carcinoma based on principal component analysis and differentially expressed genes." Journal of Prevention and Treatment for Stomatological Diseases (2022): 251-257.