 DOI: 10.21608/adjalexu.2024.340500.1561

# IDENTIFICATION OF POTENTIAL BIOMARKERS AND THERAPEUTIC TARGETS FOR ORAL SQUAMOUS CELL CARCINOMA USING BIOINFORMATICS ANALYSIS

Nourhan A. Abou Madawi [1*], Zeinab E. Darwish [2], Radwa A. Mehanna [3,4], Nehal A. Khalil[5,6], Heba R. Mady [1], Enas M. Omar [7]

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

## ABSTRACT

**INTRODUCTION:** Oral squamous cell carcinoma (OSCC) is the most common cancer arising in the oral cavity and despite progress in its management, the survival rate has not markedly increased in the last few decades. Bioinformatics has helped in great knowledge about the genetic origins and functional mechanisms influencing human diseases specially cancer which has helped in identifying novel therapeutic drug targets. Thus, bioinformatics can pave the way towards successful management of OSCC.

**OBJECTIVES:** To identify genes driving oncogenesis in OSCC that can help in early diagnosis and developing new effective therapeutic drugs.

**METHODOLOGY:** The bioinformatics online tool UCSC Xena was used to extract RNA-Seq data of OSCC from head and neck cancer dataset in the The Cancer Genome Atlas (TCGA) database. Differential gene expression analysis between tumor and control samples was done by DESeq2 package in R. Functional enrichment analysis of differentially expressed genes (DEGs) was done using DAVID bioinformatics tool. Protein-protein interaction network was represented by STRING.

**RESULTS:** A total of 159 DEGs between OSCC and normal tissue samples were identified. 151 genes were downregulated and 8 genes were upregulated. GO annotation revealed that they were mainly enriched in keratinization and intermediate filament organization. Moreover, protein-protein interaction network defined intermediate filament organization to be functionally enriched with the genes KRT81, KRT83, KRT76 and KRT36 being highly correlated and significantly downregulated in oral cancer.

**CONCLUSIONS:** The key genes in OSCC identified from functional genomics using bioinformatics analysis can be used as new biomarkers and therapeutic targets for OSCC.

**KEYWORDS:** Oral squamous cell carcinoma, Biomarkers, Targeted therapy, Bioinformatics analysis

**RUNNING TITLE:** Bioinformatics-based identification of key oral carcinoma biomarkers

---

1-Assistant Lecturer at Oral Pathology Department, Faculty of Dentistry, Alexandria University
2-Professor at Oral Pathology Department, Faculty of Dentistry, Alexandria University.
3-Professor at Medical Physiology Department, Faculty of Medicine, Alexandria University.
4-Executive Manager, Stem Cells and Cell Culture laboratories coordinator, Center of Excellence for Research in Regenerative Medicine and Applications, CERRMA, Faculty of Medicine, Alexandria University.
5-Lecturer at Medical Biochemistry Department, Faculty of Medicine, Alexandria University
6-Member of Center of Excellence for Research in Regenerative Medicine and Applications, CERRMA, Faculty of Medicine, Alexandria University.
7-Assistant Professor at Oral Pathology Department, Faculty of Dentistry, Alexandria University.

***Corresponding author:***
nourhan.madawi@alexu.edu.eg

## INTRODUCTION

Oral squamous cell carcinoma (OSCC) is a type of head and neck cancer that is considered the most common cancer in the oral cavity (1). The incidence of OSCC is increasing worldwide and the 5-year survival rate has not greatly improved in the last few years despite progress in diagnostic and treatment methods (2).

Cancer is a complex genetic disease characterized by cumulative gene mutations and consequent abnormal gene expressions (3). Conventional treatment of cancer includes three methods: surgery, radiotherapy, and chemotherapy. However, these methods have side effects such as functional impairment and toxicity linked to non-specific targeting of cancer cells (4). Despite favorable initial response, a main hurdle affecting the effectiveness of chemotherapy is the development of drug resistance through the course of treatment (5).

Understanding molecular mechanisms of oral cancer is the key to its successful management. Biomarkers are molecules that orchestrate these mechanisms and studying them is crucial. The discovery of new biomarkers helps deciphering more about cellular events and signaling pathways that control oral cancer (6).

Gene sequencing has greatly helped in the field of cancer among many diseases allowing detection of pathologic genetic elements and the expression of specific genes related to cancer. With the development of this technology, genetic data obtained from different cancer types have been stored in electronic databases. Analysis of these data provides great benefits in identifying new biomarkers and better understanding of molecular mechanisms and pathways driving cancer which is the role of bioinformatics (7). Genetic data analysis has allowed for the development of therapeutic agents aimed at cancer; however, successful cancer management is still a major problem and needs extensive study.

Previous studies conducted to designate the genetic landscape in head and neck squamous cell carcinomas have detected abnormal expression of genes such as TP53, NOTCH1, CDKN2A, EGFR and PIK3CA. However, targeting these genes has not generated impactful results in OSCC management (8-11). Moreover, studies have revealed that genetic profiles of OSCC patients differ between various ethnic groups possibly owing to variation in etiological factors and propensity to certain anatomical sites which infer that patients may not benefit similarly from molecular targeted drugs (12-15). Thus, it is urgent to determine biomarkers and utilize them for improving the effectiveness of targeted therapy which is highly dependent on patient's genetic and molecular profiles. Establishment of a molecular classification of OSCC subtypes would markedly direct the selection of the most appropriate treatment modality as is the case with breast cancer. However, clinical application for OSCC is still in its initial steps (16).

The aim of the present study was to identify new biomarkers of OSCC based on bioinformatics analysis that provide avenues for advancing diagnosis, assessing prognosis and providing innovative treatment methods.

## MATERIALS AND METHODS

Data acquisition

In the present study, we used UCSC Xena which is an online exploration tool for publicly available data of multiple cancer types at The Cancer Genome Atlas (TCGA) project. HTSeq - Counts workflow of RNA-sequencing (RNA-Seq) of head and neck cancer dataset (n=546) were downloaded from TCGA. (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 dataset 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

(e) Sample type: primary tumor

After exclusion of the data related to extraoral sites, a total of 360 samples within the OSCC dataset were selected, among which 328 tumor samples and 32 matched normal tissue samples were detected.

Exploratory analysis of data

Screening of data was 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 out the lowest variance genes, the standard deviation score (SDS) was calculated for all the mapped genes using matrix state package (CRAN: Package matrixStats (r-project.org). Selection of the top 25000 sorted genes had been done into which differential expression analysis was done. To reveal the similarities between the two study groups, histogram and 3D principal component analysis (PCA) plots have been made to detect the outliers before the differentiation step.

Differential gene expression analysis

Differential gene expression analysis between tumor tissues and normal tissues across the OSCC 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, $\log_2$FC (fold change)> 2 as upregulated and $\log_2$FC < -2 as downregulated genes.

Functional enrichment analysis

Gene Ontology (GO) functional enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of the differentially expressed genes (DEGs) were done using DAVID online tool (http://david.ncifcrf.gov ). where, false discovery rate (FDR) less than 0.05 and minimum enrichment of 3 were used as the screening criteria. Graphical representations such as bar, chord and circ plots were constructed. This was accomplished using the "GOplot" package (https://bioconductor.org/packages/GOplot/ ) in R software (version 4.21), ensuring a clear and interpretable presentation of the outcomes. Go terms for DEGs were identified as well as the related KEGG pathways. Go terms were

categorized into three groups: biological processes (BP), molecular functions (MF), and cellular components (CC).

Protein-protein interaction network of DEGs

Protein-protein interaction (PPI) of the most significant 50 DEGs was done using STRING online tool (https:// www.string-db.org/ ), to obtain the PPI network.

A flowchart for the study design is illustrated in figure 1.

Statistical analysis

Data were analyzed using R programming language run by R studio with screening criteria of adj. $P < 0.01$ and $\log_2 FC > 2$, or $\log_2 FC < -2$ for statistical significance.

## RESULTS

Clinical results of OSCC dataset

In the present study, a total of 360 data files of OSCC have been selected from the downloaded head and neck cancer dataset. Among them, 328 files were true tumor samples while 32 files were related to matched normal tissue samples. Clinical analysis of data reveals that males represent (68.6%) of tumor cases (n=225), while females represent (31.4%) of tumor cases (n=103). 45.7% of tumor samples were dead (n=150) and 54.3% of tumor samples were live (n=178).

Concerning the anatomical site, the most common site of occurrence was the tongue representing 47.6% of the cases (n=156), followed by the other non-specified parts of the oral cavity representing 23% of the cases (n=76). The floor of the mouth accounted for 17% of the cases (n=56), the cheek mucosa 5.8% of the cases (n=19), the gums 3.6% of the cases (n=11). The hard palate, lips and mandible were represented by 1.5%&1%&0.5% respectively (n=5), (n=3), (n=2).

Exploratory analysis and data processing results

To detect the similarities between the two study groups, data were transformed to tidy data and histogram was illustrated using ggplot2 package in R (https://ggplot2.tidyverse.org/). The histogram showed no outliers in the selected data between the 2 study groups (figure 2a). 3D PCA plot has been made using prcomp function in R and showed no outliers between the tumor data and soft tissue normal data (figure 2b).

Identification of differentially expressed genes

According to the analysis performed by DESeq2, a total of 1,506 DEGs were identified between tumor samples and normal tissue samples. DEGs with significance were selected according to adjusted P-value <0.01 as screening criteria for differential genes, $\log_2 FC > 2$ as upregulated and $\log_2 FC < -2$ as downregulated genes (figure 3a). 159 DEGs were obtained .151 were downregulated and only 8 were upregulated.

The obtained DEGS were ordered by adjusted P-value and $\log_2$ FC and selection of the most significant 50 DEGS has been done. All of them were downregulated except PCDHGA12. The expression of these 50 DEGs between the 2 groups of comparison is illustrated in figure 3b.

Functional annotation of the top 50 DEGs in Go term analysis

Functional enrichment analysis of DEGs revealed that they were enriched in keratinization, intermediate filament organization, biomineral tissue development and defense response to bacterium in the BP group; extracellular space, extracellular region, extracellular exosome and keratin filament in the CC group; structural constituent of skin epidermis in the MF group.

GO pathway enrichment analysis for the three categories of GO terms for top 50 DEGs is illustrated in bar plot (figure 4a) and circ plot (figure 4b).

Functional annotation of top 50 DEGs via KEGG analysis

KEGG pathway analysis of the top 50 DEGs revealed that the most enriched KEGG pathway for OSCC was salivary secretion with all the 9 genes associated were downregulated indicating that decreased salivary secretion might be considered as a risk factor for progression of oral cancer (figure 5).

Protein-protein interaction network

The analysis of the PPI network between differentially co-expressed genes was established through the STRING database tool, generating 40 nodes and 61 edges and resulting in 2 significant modules using MCL clustering method which are salivary secretion illustrated in red and intermediate filament organization illustrated in yellow. The hub genes of the PPI network were obtained using the CytoHubba plugin BottleNeck Algorithm provided through Cytoscape software. According to the BottleNeck Algorithm, the top 8 genes in the network included STATH, SMR3B, CST4, KRT83, PRB3, ZG16B, BPIFA2, MUC7. STATH, SMR3B, CST4 and KRT83 were the most significant (figure 6).

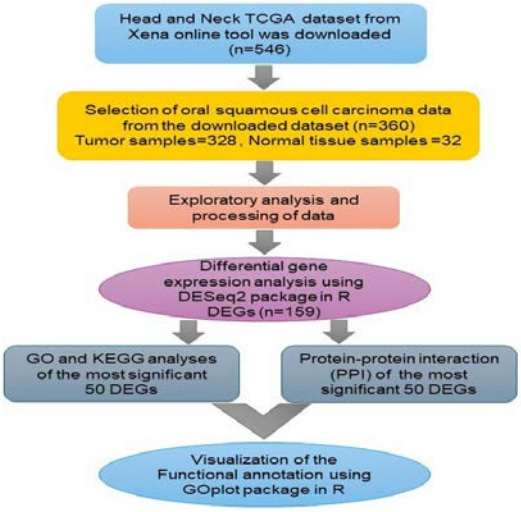

**Figure 1:** Flow chart showing the study design.

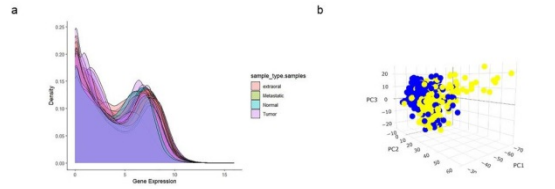

**Figure 2:** Histogram (a) and 3D PCA (b) revealing no outliers between the tumor data and normal tissue data. Yellow dots represent tumor samples while blue dots represent normal tissue samples.

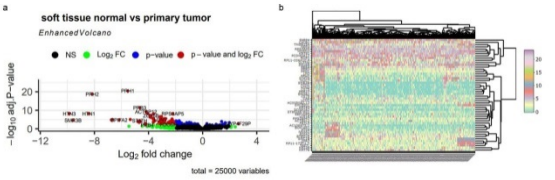


**Figure 3:** (a) Volcano plot showing adjusted P-value versus fold change for differential gene expression analysis (b) Heatmap illustrating the top 50 differentially expressed genes.

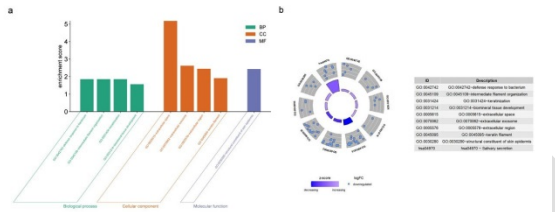

**Figure 4:** Gene ontology (GO) and KEGG analyses of differentially expressed genes. (a) Bar plot illustrating enriched GO terms for biological processes, cellular components, molecular functions. (b) Circ plot illustrating enriched GO terms for biological processes, cellular components, molecular function and KEGG pathway. Each dot represents an individual gene

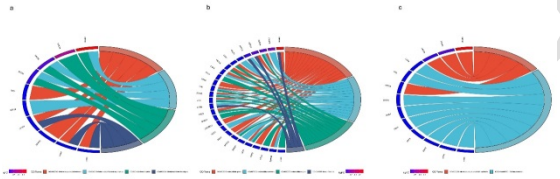

**Figure 5:** Chord plots illustrating relationships between genes and (a) biological processes terms (b) Cellular components terms (c) Molecular function term and KEGG pathway.

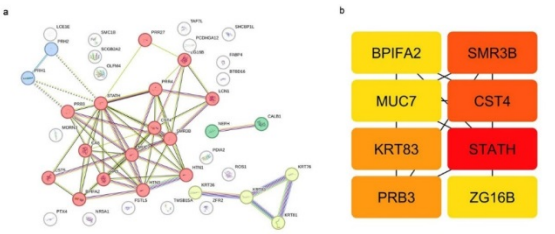


**Figure 6:** (a) Protein-protein interaction network of the 50 significant differentially expressed genes revealing 2 significant clusters: salivary secretion illustrated in red and intermediate filament organization illustrated in yellow. (b) Hub genes from the 2 significant clusters. Darker color indicates more significance.

## DISCUSSION

Despite progress in OSCC treatment, the survival rate of patient has not greatly improved. Recurrence is a common outcome following treatment that decreases survival (17). Late diagnosis is implicated in poor prognosis of OSCC (18). Patients diagnosed at an advanced stage may not benefit from the wide range of recent targeted therapeutic drugs. Thus, the discovery of new biomarkers can help in the early diagnosis of OSCC and selection of the most appropriate targeted drug tailored to each patient. Also, finding more about the genetic profile of oral cancer, enlighten the way in the journey of monitoring cancer progression and successful treatment (16).

In the present study, we used bioinformatics tools to identify DEGs in OSCC and highlight their functions. DEGs can be hopefully exploited and used as biomarkers to help in monitoring OSCC patients and allow innovation for developing effective anticancer drugs.

The PPI of the obtained top 50 DEGs accentuated one enriched pathway (salivary secretion) and a significant biological process (intermediate filament organization). cytoHubba plugin identified the most significant 8 genes (hub genes). All were downregulated. STATH, SMR3B, CST4, PRB3, ZG16B, BPIFA2, and MUC7 were correlated and share a common pathway which is salivary secretion. Downregulated salivary secretion genes in OSCC indicate low salivary secretion accompanying oral carcinogenesis. Interestingly, KRT83 was one of the top 8 genes but linked to a different cluster which is intermediate filament organization.

KRT83, KRT81, KRT76 and KRT36 were from the most significant 50 DEGs in OSCC. They are strongly interacting genes enriched in intermediate filament organization, keratinization and structural constituent of skin epidermis.

KRT83, KRT81, and KRT76 belong to keratin type II family of keratin genes responsible for formation of epithelium. They are located on chromosome 12 (4). Keratin filaments are intermediate filaments: one of the three main constituents of the cytoskeleton, the other being microfilaments and microtubules (19). Keratin filaments help structural organization of epithelial cells, making networks that allow for mechanical support and resisting cellular stress (20). In addition, they participate in cellular processes such as cell motility and apoptosis. Keratin genes play important roles in carcinogenesis such as invasion and metastasis (4).

KRT83 functions in cell-cell adhesion. Its abnormal expression is linked to cancer (21). Change in gene expression levels in cancer may be due to mutation affecting the DNA sequence or errors in mechanisms altering the transcriptional process known as epigenetic dysregulation (22). Epigenetic alteration is a common cause driving cancer apart from genetic mutation. KRT83 was found to be methylated in metastatic breast cancer patients to the brain (21).

Miyata et al. (23) reported that the expression of KRT83 in metastatic cholangiocarcinoma was higher than in primary cholangiocarcinoma using cDNA microarray (23).
In a study concerned with osteosarcoma, mRNA sequencing data and clinical data of pediatric osteosarcoma patients were obtained from TARGET database available at the Genomic Data Commons (GDC). Gene expression analysis detected that high KRT83 expression was strongly associated with overall survival and high mortality (24).

Moreover, Cui et al. (25) conducted a study to explore the prognostic value of tumor mutation burden (TMB) in breast cancer utilizing bioinformatics analysis. KRT 83 was among the top 10 DEGs between low-TMB and high-TMB breast cancer patients. KRT83 was upregulated in the low-TMB group (25).

In the study of Yan et al. (20), the expression of KRT81 in triple negative breast cancer (TNBC) was high while it was low in other types of breast cancer. KRT81 expression increased in advanced stage of the TNBC. TNBC is an aggressive form of breast cancer and its treatment is challenging with lower response to drugs. KRT81 was found to be a prognostic factor of TNBC (20).

Nanashima et al. (26) carried out knock down of KRT81 in breast cancer cells which resulted in decreased cell migration and invasion by downregulating invasion-related genes such as TNF, MMP9 and LCN2. Although KRT81 is a hair keratin gene known to be expressed in hair and nails (hard keratin), it is expressed- in lower levels- in other tissues like normal breast tissues and breast cancer (26).

Similarly, knock down of KRT81 in melanoma cell lines decreased cell proliferation, migration and invasion by downregulating interleukin-8. In the same study, the expression of KRT81 was measured in melanoma tissues and adjacent normal tissues and was found to be higher in cancer (27).

KRT36 is a type I keratin gene positioned on chromosome 17. It belongs to the hair keratins and is a marker of differentiation of the hair cortex. KRT36 is normally expressed in the tongue but it was found to be downregulated in tongue squamous cell carcinoma (28). Furthermore, clinically normal adjacent tongue tissue exhibited low expression of KRT36 accentuating the field cancerization event (28,29). In contrast, KRT81 was upregulated in tongue cancer (28).

KRT76 is a type II keratin gene whose protein product is expressed in the suprabasal layers of a variety of human epithelia (30). In the oral cavity, KRT76 is particularly expressed in the suprabasal cell layers of the masticatory epithelium (epithelium lining gingiva and hard palate) (31,32). It was reported that KRT76 was downregulated in OSCC and correlated with poor prognosis (32). Also, the expression of KRT76 was diminished in oral premalignant lesions and was even lower in gingivobuccal carcinomas as compared to normal oral tissue (32). In a study carried out on mice with different expression states of KRT76 and induction of carcinomas, the development of carcinomas was faster in homozygous mice (KRT76−/−) than homozygous mice (KRT76+/+) and heterozygous mice (KRT76+/−) suggesting that loss of KRT76 participates in oral cancer progression (30).

KRT76 is important in maintenance of epithelium and was found to be actively expressed in sites of wound healing. Mice models with Knockout of KRT76 experienced non-healing skin wounds (33).

A study conducted by Lee et al. (34) to compare gene expression between the gingival and dental pulp revealed that C-MYC and KRT76 among other genes were more expressed in the gingiva than the dental pulp reflecting the role of gingiva in proliferation and differentiation. It was reported that oral carcinomas have downregulation of KRT76 and upregulation of C-MYC (32,35) and targeting these molecules synchronously may provide favorable clinical results.

Abnormal expression of keratin genes in OSCC propose defect in differentiation which may be linked to change in cell cycle normal pattern (36).

The difference between our results and the results of previous studies is due to variation among various cancer types and even between patients of the same cancer type supporting the ongoing era of personalized medicine to ensure the best treatment outcome for each single patient.

## CONCLUSION

Since cancer is a polygenic disease driven by a huge number of genetic abnormalities, targeting a number of biomolecules linked together in a signaling pathway or a molecular function can offer superior therapeutic results and encourage the production of new effective targeted drugs.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

FUNDING

The authors did not receive any fund for this study.

## REFERENCES


1. Ahmad P, Nawaz R, Qurban M, Shaikh GM, Mohamed RN, Nagarajappa AK, et al. Risk factors associated with the mortality rate of oral squamous cell carcinoma patients: A 10-year retrospective study. Medicine. 2021;100: e27127.
2. Obrenović M, Šupić G, Miyabe S, Mladenović I, Kozomara R, Jović S, et al. Impact of TLR9 and TLR7 gene polymorphisms on prognosis and survival of patients with oral squamous cell carcinoma. Biomol Biomed. 2024; 24:1682-91.
3. Capellini A, Williams M, Onel K, Huang KL. The functional hallmarks of cancer predisposition genes. Cancer Manag Res. 2021; 13:4351-7.
4. Takan I, Karakülah G, Louka A, Pavlopoulou A. "In the light of evolution:" keratins as exceptional tumor biomarkers. PeerJ. 2023;11: e15099.
5. Emran TB, Shahriar A, Mahmud AR, Rahman T, Abir MH, Siddiquee MF, et al. Multidrug resistance in cancer: understanding molecular mechanisms, immunoprevention and therapeutic approaches. Front Oncol. 2022; 12:891652.
6. Liu L, Chen J, Cai X, Yao Z, Huang J. Progress in targeted therapeutic drugs for oral squamous cell carcinoma. Surg Oncol. 2019; 31:90-7.
7. Huang L, Chen J, Zhao Y, Gu L, Shao X, Li J, et al. Key candidate genes of STAT1 and CXCL10 in melanoma identified by integrated bioinformatical analysis. IUBMB life. 2019; 71:1634-44.
8. Kozaki K, Imoto I, Pimkhaokham A, Hasegawa S, Tsuda H, Omura K, et al. PIK3CA mutation is an oncogenic aberration at advanced stages of oral squamous cell carcinoma. Cancer Sci. 2006; 97:1351–8.
9. Taoudi Benchekroun M, Saintigny P, Thomas SM, El-Naggar AK, Papadimitrakopoulou V, Ren H, et al. Epidermal growth factor receptor expression and gene copy number in the risk of oral cancer. Cancer Prev Res (Philos). 2010; 3:800–9.
10. Cervino G, Fiorillo L, Herford AS, Romeo U, Bianchi A, Crimi S, et al. Molecular biomarkers related to oral carcinoma: clinical trial outcome evaluation in a literature review. Dis Markers. 2019; 2019:8040361.
11. Starzynska A, Adamska P, Sejda A, Sakowicz-Burkiewicz M, Adamski LJ, Marvaso G, et al. Any role of PIK3CA and PTEN biomarkers in the prognosis in oral squamous cell carcinoma? Life (Basel). 2020; 10:325.
12. Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. Cancer Discov. 2013; 3:770-81.
13. Song X, Xia R, Li J, Long Z, Ren H, Chen W, et al. Common and complex Notch1 mutations in Chinese oral squamous cell carcinoma. Clin Cancer Res. 2014; 20:701-10.
14. Al-hebshi NN, Li S, Nasher AT, El-Setouhy M, Alsanosi R, Blancato J, et al. Exome sequencing of oral squamous cell carcinoma in users of Arabian snuff reveals novel candidates for driver genes. Int J Cancer. 2016; 139:363-72.
15. Su SC, Lin CW, Liu YF, Fan WL, Chen MK, Yu CP, at al. Exome sequencing of oral squamous cell carcinoma reveals molecular subgroups and novel therapeutic opportunities. Theranostics. 2017; 7:1088-99.
16. Chai AWY, Lim KP, Cheong SC. Translational genomics and recent advances in oral squamous cell carcinoma. Semin Cancer Biol. 2020; 61: 71-83.
17. Zhu L, Yan D, Chen Y, Chen S, Chen N, Han J. The identification of autophagy-related genes in the prognosis of oral squamous cell carcinoma. Oral Dis. 2020; 26:1659–67.
18. Madhura MG, Rao RS, Patil S, Fageeh HN, Alhazmi A, Awan KH. Advanced diagnostic aids for oral cancer. Diease-a-Month. 2020; 66: 101034.
19. Jacob JT, Coulombe PA, Kwan R, Omary MB. Types I and II keratin intermediate filaments. Cold Spring Harb Perspectives Biol. 2018;10: a018275.
20. Yan Z, Zhong Z, Shi C, Feng M, Feng X, Liu T. The prognostic marker KRT81 is involved in suppressing CD8+ T cells and predicts immunotherapy response for triple-negative breast cancer. Cancer Biol Ther. 2024; 25:2355705.
21. Pangeni RP, Channathodiyil P, Huen DS, Eagles LW, Johal BK, Pasha D, et al. The GALNT9, BNC1 and CCDC8 genes are frequently epigenetically dysregulated in breast tumours that metastasise to the brain. Clin Epigenetics. 2015; 7:1-7.
22. Fymat AL. Genetics, epigenetics and cancer. Cancer Ther Oncol Int J. 2017; 4:555634.
23. Miyata T, Yamashita YI, Yoshizumi T, Shiraishi M, Ohta M, Eguchi S et al. CXCL 12 expression in intrahepatic cholangiocarcinoma is associated with metastasis and poor prognosis. Cancer Sci. 2019; 110:3197-203.
24. Cui T, Guo J, Jin C. Identification of a 7-mRNA signature as a prognostic biomarker in pediatric osteosarcoma. Transl Cancer Res. 2020; 9:6733.
25. Cui S, Feng J, Tang X, Lou S, Guo W, Xiao X, et al. The prognostic value of tumor mutation burden (TMB) and its relationship with immune infiltration in breast cancer patients. Eur J Med Res. 2023; 28:90.

26. Nanashima N, Horie K, Yamada T, Shimizu T, Tsuchida S. Hair keratin KRT81 is expressed in normal and breast cancer cells and contributes to their invasiveness. Oncol Rep. 2017; 37:2964-70.
27. Zhang K, Liang Y, Zhang W, Zeng N, Tang S, Tian R. KRT81 knockdown inhibits malignant progression of melanoma through regulating interleukin-8. DNA Cell Biol. 2021; 40:1290-7.
28. Brychtova V, Coates PJ, Hrabal V, Boldrup L, Fabian P, Vojtesek B, et al. Keratin 36, a specific marker of tongue filiform papillae, is downregulated in squamous cell carcinoma of the mobile tongue. Mol Clin Oncol. 2020; 12:421-8.
29. Boldrup L, Gu X, Coates PJ, Norberg-Spaak L, Fahraeus R, Laurell G, et al. Gene expression changes in tumor free tongue tissue adjacent to tongue squamous cell carcinoma. Oncotarget. 2017; 8:19389.
30. Sequeira I, Neves JF, Carrero D, Peng Q, Palasz N, Liakath-Ali K, et al. Immunomodulatory role of Keratin 76 in oral and gastric cancer. Nat Commun. 2018; 9:3437.
31. Collin C, Ouhayoun JP, Grund C, Franke WW. Suprabasal marker proteins distinguishing keratinizing squamous epithelia: cytokeratin 2 polypeptides of oral masticatory epithelium and epidermis are different. Differentiation. 1992; 51:137-48.
32. Ambatipudi S, Bhosale PG, Heath E, Pandey M, Kumar G, Kane S, et al. Downregulation of keratin 76 expression during oral carcinogenesis of human, hamster and mouse. PloS one. 2013;8: e70688.
33. DiTommaso T, Cottle DL, Pearson HB, Schlüter H, Kaur P, Humbert PO, Smyth IM. Keratin 76 is required for tight junction function and maintenance of the skin barrier. PLoS genetics. 2014;10: e1004706.
34. Lee KE, Kang CM, Jeon M, Kim SO, Lee JH, Choi HJ. General gene expression patterns and stemness of the gingiva and dental pulp. J Dent Sci. 2022; 17:284-92.
35. 35.Marconi GD, Della Rocca Y, Fonticoli L, Melfi F, Rajan TS, Carradori S,et al. C-myc expression in oral squamous cell carcinoma: molecular mechanisms in cell survival and cancer progression. Pharmaceuticals. 2022; 15:890.
36. Akgül B, Ghali L, Davies D, Pfister H, Leigh IM, Storey A. HPV8 early genes modulate differentiation and cell cycle of primary human adult keratinocytes. Experimental dermatology. 2007; 16:590-9