



## Original Research

# Identification of novel prognostic indicators for oral squamous cell carcinoma based on proteomics and metabolomics

Zhitao Yao<sup>a,b</sup>, Wei An<sup>a,b</sup>, Maimaitixun Tuerdi<sup>a,b</sup>, Jin Zhao<sup>a,b,\*</sup>

<sup>a</sup> Department of Trauma and Orthopedics, the First Affiliated Hospital of Xinjiang Medical University, No. 137 South Liyushan Road, Urumqi 830054, China

<sup>b</sup> Oral Disease Institute of Xinjiang Uyghur Autonomous Region, No.137 South Liyushan Road, Urumqi 830054, China

## ARTICLE INFO

## Keywords:

Oral squamous cell carcinoma  
Proteome  
Metabolome  
Bioinformatics  
Prognostic risk score

## ABSTRACT

**Background:** The low 5-year survival rate of oral squamous cell carcinoma (OSCC) suggests that new prognostic indicators need to be identified to aid the clinical management of patients.

**Methods:** Saliva samples from OSCC patients and healthy controls were collected for proteomic and metabolomic sequencing. Gene expressed profiling was downloaded from TCGA and GEO databases. After the differential analysis, proteins with a significant impact on the prognosis of OSCC patients were screened. Correlation analysis was performed with metabolites and core proteins were identified. Cox regression analysis was utilized to stratify OSCC samples based on core proteins. The prognostic predictive ability of the core protein was then evaluated. Differences in infiltration of immune cells between the different strata were identified.

**Results:** There were 678 differentially expressed proteins (DEPs), 94 intersected DEPs among them by intersecting with differentially expressed genes in TCGA and GSE30784 dataset. Seven core proteins were identified that significantly affected OSCC patient survival and strongly correlated with differential metabolites ( $R^2 > 0.8$ ). The samples were divided into high- and low-risk groups according to median risk score. The risk score and core proteins were well prognostic factor in OSCC patients. Genes in high-risk group were enriched in Notch signaling pathway, epithelial mesenchymal transition (EMT), and angiogenesis. Core proteins were strongly associated with the immune status of OSCC patients.

**Conclusions:** The results established a 7-protein signatures with the hope of early detection and the capacity for risk assessment of OSCC patient prognosis. Further providing more potential targets for the treatment of OSCC.

## Introduction

Approximately 95% of head and neck cancers are squamous cell carcinomas. According to GLOBOCAN data up to 2020, head and neck squamous cell carcinoma (HNSCC) is the fifth leading cancer type worldwide [1]. Oral cancer is the most common type of HNSCC. Global statistics indicate 375,000 new cases of oral cancer and over 175,000 deaths by 2020 [2]. Oral squamous cell carcinoma (OSCC) is an aggressive epithelial tumor of varying degrees of differentiation that accounts for approximately 90% of oral cancers [3]. Alcohol abuse, smoking, and human papillomavirus (HPV) infection are the main risk factors for OSCC [4].

The main treatments for OSCC are surgery, radiotherapy, and chemotherapy [5,6]. However, despite more diverse and more advanced treatment options, the 5-year overall survival (OS) rate of OSCC has

remained at 50–60% in recent years [7]. Clinically OSCC can metastasize to cervical lymph nodes via lymphatic vessels, and the incidence of metastasis is up to 40%, which is an important prognostic factor affecting survival [8]. OSCC also has a high recurrence rate, and patients with late recurrence have a worse 5-year OS rate [9]. Therefore, exploring the targets and tumor stratification for the prognosis and treatment of OSCC is one of the urgent issues for basic research and the clinic.

Recent studies confirmed a close link between OSCC and metabolites, which may provide new perspectives and new potential targets for the diagnosis and treatment of OSCC [10,11]. Tumor metabolism is associated with tumor cell proliferation, survival, and antitumor immunity in the tumor microenvironment (TME) [12,13]. Accumulating evidence suggests that in-depth analysis of the immune response, emphasizing tumor infiltrating immune cells, is a widely accepted

\* Corresponding author at: Department of Trauma and Orthopedics, the First Affiliated Hospital of Xinjiang Medical University, No. 137 South Liyushan Road, Urumqi 830054, China.

E-mail address: [zhaojinxiang@163.com](mailto:zhaojinxiang@163.com) (J. Zhao).

<https://doi.org/10.1016/j.tranon.2023.101672>

Received 17 January 2023; Received in revised form 27 February 2023; Accepted 9 April 2023

1936-5233/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

prognostic tool that may significantly improve the development of novel immunotherapeutic approaches for OSCC patients [14]. Proteins are common molecules involved in cellular functions, and when the functional state is altered in cancer, proteins with connections to various pathways may alter the balance of signaling networks to enhance survival or apoptosis of affected cells [15].

Indeed, large-scale datasets generated using high-throughput analytical techniques for transcriptomics, proteomics, and metabolomics have helped to identify potential diagnostic and therapeutic targets in human cancer. Saliva has attracted extensive attention as one useful sample for early detection or monitoring of oral cancer. Expanding the selection of treatment options for OSCC in saliva specimens is an effective strategy to improve survival.

In this study, we investigated protein biomarkers and prognostic stratification of OSCC from the perspective of transcriptomics, proteomics as well as metabolomics, contributing to early diagnosis, treatment and prognosis of patients.

## Materials and methods

### Sample collection

Saliva samples from nine OSCC patients and three healthy controls were randomly selected from the First Affiliated Hospital of Xinjiang Medical University. All study protocols were approved by the Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University, Urumqi, China (No. K202203–24). Informed consent was known and signed by all participants.

### Salivary proteome

Lysate (1.5% SDS / 100 mM Tris-Cl) was added to saliva samples on ice, mixed well and centrifuged at 4 °C to obtain supernatant. Proteins in the supernatant were precipitated by acetone precipitation and reconstituted using the reconstitution solution (8 M Urea/100 mM Tris-Cl). Subsequently, dithiothreitol and iodoacetamide were added to block sulfhydryl. The protein concentration was determined by Bradford assay, and the protein bands were visualized by SDS-PAGE. Enzymatic digestion with trypsin after diluting the urea concentration below 2 M was followed by Sep-Pak C18 desalting. Equal amounts of samples were taken for TMT labeling, which was performed according to the manufacturer's instructions. After equal mixing of the labeled samples, the pooled samples were fractionated by high pH reverse chromatography and checked on the machine.

Mass spectrometry data were acquired using a liquid-mass coupled system with an Orbitrap Exploris 480 mass spectrometer in tandem with an EASY-nLC 1200 liquid phase. Peptide samples were loaded and incorporated into the analytical column for separation, Mass spectra were acquired in DDA mode with one MS full scan in each scan cycle ( $R = 60$  K, AGC = 300%, max IT = 20 ms, scan range = 350–1500  $m/z$ ), and 20 subsequent MS / MS scans ( $R = 15$  K, AGC = 100%, max IT = auto, cycle time = 2 s). The dynamic exclusion time for ion repeat acquisition was set to 35 s.

Mass spectrometric data were searched by MaxQuant (v1.6.6) software and the adopted database searching algorithm was Andromeda. The database used for the search was Swissprot Human (20,210,312) proteome reference databases.

### Detection of salivary metabolites

A total of 100  $\mu$ L of samples were added to 300  $\mu$ L of pure methanol (containing 1 ppm of 2-chlorophenylalanine) and incubated at -20 °C freezer for 0.5 h. The supernatant was collected in an injection vial and used for liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. The column was a Waters ACQUITY UPLC HSS T3 C18 1.8  $\mu$ m, 2.1 mm  $\times$  100 mm.

### Collection of public data

We downloaded the expression data and clinical data of HNSCC dataset from The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov>) database. Which included 522 tumor and 44 normal samples. GSE30784 and GSE41613 datasets were collected from GEO (Gene Expression Omnibus) database (<https://www.ncbi.nlm.nih.gov/geo/>). GSE30784 included gene expressed profiling of 167 OSCC and 45 normal oral tissues based on GPL570 platform. GSE41613 included gene expression profiling of 97 OSCC samples with OSCC-specific survival information based on GPL570 platform. The human protein atlas (HPA) database (<https://www.proteinatlas.org/>) was used to detected the immunohistochemical results of proteins in this study.

### Difference analysis

The difference analysis for proteins and metabolites between OSCC and controls was performed. Fold changes  $\geq 1.2$  and  $P < 0.05$  ( $t$ -test) were considered statistically significant and then obtained differentially expressed proteins (DEPs) and differential metabolites. The correlation between DEPs and differential metabolites was calculated using Spearman correlation coefficient analysis. Correlation coefficients greater than 0.8 were considered as strong correlations.

The differentially expressed genes (DEGs) between HNSCC and controls in TCGA were identified using EdgeR in R software with the standard  $|\log_2(\text{fold change [FC]})| > 1.0$  and  $P < 0.05$ . The DEGs between OSCC and controls in GSE30784 were identified using limma in R software with the standard  $|\log_2(\text{fold change [FC]})| > 1.0$  and  $P < 0.05$ .

### Survival analysis

The relationship between DEGs expression and patient prognosis in TCGA or GSE41613 was evaluated using Kaplan-Meier method.  $P < 0.05$  was considered significant. A multivariate Cox regression analysis was used to construct risk score model based on the median risk score, then the samples of cancer were divided into high-risk and low-risk group. DEGs between high-risk and low-risk group were also identified using EdgeR in R software for TCGA and limma in R software for GSE41613. Receiver operating characteristics (ROC) curves were used to assess the predicting ability of patients' survival status. Nomograms were established by expression of genes using Rms in R software.

### Enrichment analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of for DEGs between OSCC and controls or between high-risk and low-risk group was performed using clusterProfiler in R software. Gene set enrichment analysis (GSEA) was performed using 3.0 version (<http://www.broadinstitute.org/gsea/>) to identify high-risk group related gene sets.

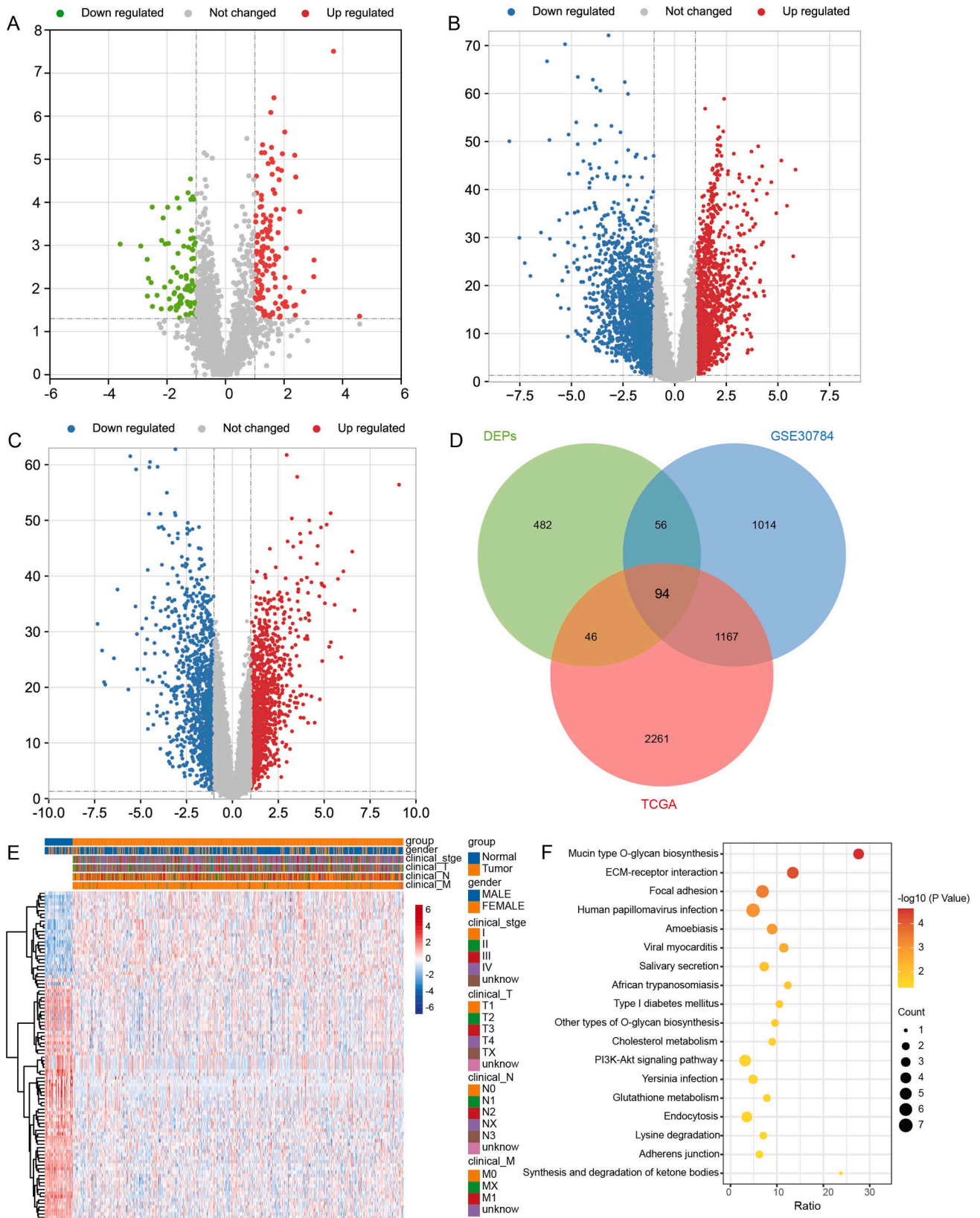
### Infiltration of immune cells

Single-sample gene set enrichment analysis (ssGSEA) was used to calculate the infiltration levels of immune cells. The differential infiltration of immune cells between high-risk and low-risk group was analyzed using limma in R software. Correlation between infiltration of immune cells and expression of proteins were evaluated using Spearman correlation coefficient analysis.

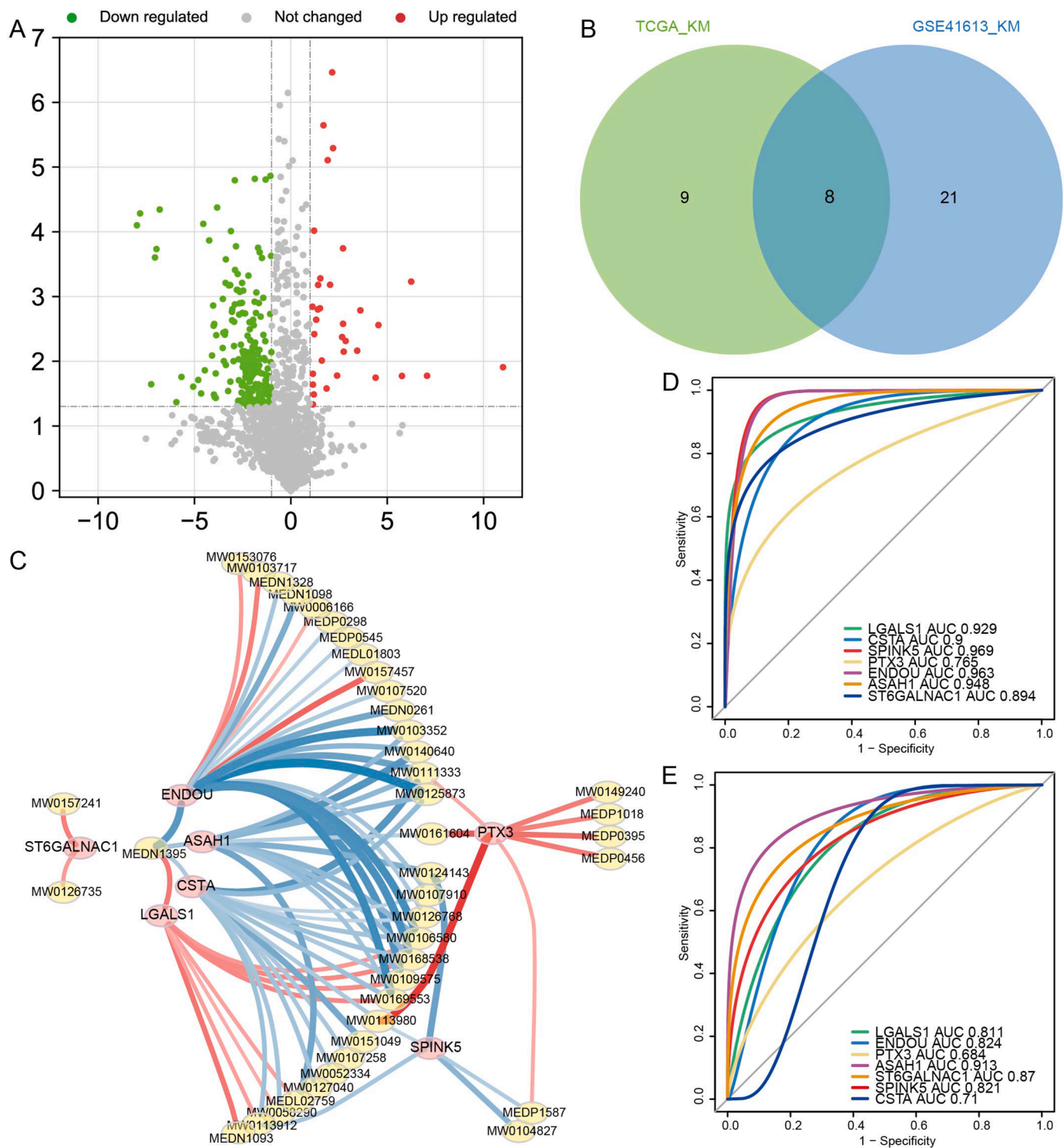
## Results

### Identification of proteins associated with OSCC

To identify the DEPs between OSCC and controls, we first performed differential expressed analysis. In the results, we found 678 DEPs in



**Fig. 1.** Differentially expressed proteins and differentially expressed genes between OSCC and controls. (A) Differentially expressed proteins between OSCC and controls by proteome sequencing in saliva samples. Red is upregulation and green is downregulation. (B) Differentially expressed genes between OSCC and controls in TCGA. Red is upregulation and blue is downregulation. (C) Differentially expressed genes between OSCC and controls in GSE30784 dataset. Red is upregulation and blue is downregulation. (D) Intersection analysis of DEPs and DEGs. DEPs, differentially expressed proteins; TCGA, the Cancer Genome Atlas. (E) Heatmap of 94 intersected DEGs in TCGA. (F) Main KEGG pathways of 94 intersected DEGs enriched.



**Fig. 2.** Recognition of proteins related to metabolism affecting prognosis in OSCC patients. (A) Differential metabolites between OSCC and controls by metabolome sequencing in saliva samples. Red is upregulation and green is downregulation. (B) Intersection of the K-M analysis results that significantly affected OS of OSCC patient in TCGA and GSE41613 dataset. (C) Differential metabolites whose correlation with core protein was greater than 0.8. Red is positive correlation and blue is negative correlation. The darker the color the higher the correlation. ROC curves of core proteins in GSE30784 dataset (D) and TCGA (E).

OSCC according to proteome sequencing (Fig. 1A). We also identified 3568 DEGs in TCGA (Fig. 1B) and 2331 DEGs in GSE30784 dataset (Fig. 1C). The results of intersection analysis showed that there were 94 intersected DEPs in three groups (Fig. 1D-E). Enrichment analysis found that 94 intersected DEPs were mainly involved in mucin type O-glycan biosynthesis, ECM-receptor interaction, and focal adhesion (Fig. 1F).

#### Identification of core proteins associated with metabolism

A total of 255 differential metabolites between OSCC and controls were also identified (Fig. 2A). In addition, there were 17 proteins of intersected DEPs significantly affected OS of patients in TCGA and 29 proteins of intersected DEPs significantly affected OS of patients in GSE41613 dataset. Finally we found 8 proteins affected the OS in both

**Table 1**  
Eight proteins affected the OS of OSCC patients.

Gene name	Protein description	Fold changes of protein levels	logFC in GSE30784	logFC in TCGA
LGALS1	Spermatogenesis-associated protein 1	1.477281	1.70969	1.687104
CSTA	Cystatin-A	0.532297	-1.03433	-1.07421
SPINK5	Serine protease inhibitor Kazal-type 5	0.547915	-4.34491	-3.26448
PTX3	Pentraxin-related protein PTX3	1.644838	1.476131	-1.47328
ENDOU	Uridylate-specific endoribonuclease	0.969514	-5.27523	-3.94611
ASAH1	Acid ceramidase	0.616373	-1.11863	-1.11154
RAB11FIP1	Rab11 family-interacting protein 1	1.477499	-1.83969	-1.49088
ST6GALNAC1	Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 1	0.452882	-3.18651	-2.96107

FC, Fold change.

datasets (Fig. 2B). The expression of 8 proteins in the proteome, TCGA and GSE30784 dataset were shown in Table 1 and Fig. S1. ASAH1, CSTA, ENDOU, SPINK5, and ST6GALNAC1 were all downregulated expression in OSCC than controls, while LGALS1 was upregulated expression. Pearson correlation analysis results between 8 proteins and differential metabolites obtained 7 proteins (ASAH1, CSTA, ENDOU, LGALS1, PTX3, SPINK5, and ST6GALNAC1) with high correlation ( $R^2 > 0.8$ ) and considered as core proteins (Fig. 2C, Table S1). ENDOU associated with 22 metabolites, which were the most. The corresponding metabolites are shown in Table S2. Among them, MW0106580 (Eflornithine), MW0109575 (S-Allyl-L-Cysteine), MW0169553 (NIMUSTINE), and MW0168538 (3-(2,5-dimethoxyphenyl)propanoic acid) were all associated with four proteins, which included ENDOU. The diagnostic predictive abilities of the core proteins for OSCC in GSE30784 dataset were all greater than 0.76 (Fig. 2D), and that in TCGA were all greater than 0.68 (Fig. 2E). The K-M curves of core proteins in GSE41613 were shown in Fig. 3. Importantly, we also checked the expression of core proteins in OSCC and controls in HPA database (Fig. 4).

#### Functional evaluation of core proteins

A multivariate Cox regression analysis was performed based on the 7 core proteins, and a risk score was calculated. A risk score model was constructed based on the median risk score, and the OSCC samples were divided into high-risk and low-risk groups in GSE41613 dataset (Fig. 5A). The results of the time-dependent ROC curves showed that the AUC values of the median risk score for predicting 1-, 3-, and 5-year OS of OSCC patients were all greater than 0.7 in GSE41613 dataset (Fig. 5B). The prognosis in high-risk group was significantly poorer than that in low-risk group (Fig. 5C). In the TCGA, the risk score model was also constructed using median risk score (Figure S2A). The area under the ROC curve (AUC) values of the median risk score for predicting 1-, 3-, and 5-year OS of OSCC patients were all greater than 0.6 (Fig. S2B). The prognosis was significantly different between patients in the high-risk and low-risk group (Figure S2C).

In the result of PCA, the high-risk and low-risk groups samples had a high degree of outlier (Fig. 5D). This was also found in TCGA (Fig. S2D). Nomograms were used to assess the impact of core proteins on the prognosis of OSCC patients in GSE41613 dataset (Fig. 5E-F) and in TCGA (Fig. S2E-F).

#### Biological function enriched in the high- and low-risk groups

To further identify the differences between the high- and low-risk

groups, we performed differential expression analysis. We identified 558 DEGs in GSE41613 dataset (Fig. 6A) and 489 DEGs in TCGA (Fig. 6B). Among these DEGs, we found 224 intersected genes (Fig. 6C). Enrichment results showed that intersected genes were mainly enriched in Amoebiasis, protein digestion and absorption, and IL-17 signaling pathway (Fig. 6D). In the GSEA results, we identified Notch signaling pathway, epithelial mesenchymal transition (EMT), and angiogenesis were activated in the high-risk group both in GSE41613 dataset (Fig. 6E) and in TCGA (Fig. 6F).

#### Immune cell infiltrated in the high-risk and low-risk groups

To further evaluate the differences in immune cell infiltration in high- and low-risk groups of GSE41613 dataset, we performed ssGSEA. Results found that the infiltrated levels of activated B cells, activated CD8+T cells, effector memory CD8+T cells, eosinophil, immature B cells, and T follicular helper cells were all lower in the high-risk group than in the low-risk group (Fig. 7A). While activated dendritic cells, and regulatory T cells (Treg) were higher in the high-risk group. Correlation between infiltration of immune cells and expression of core protein results showed that correlation between Treg cells and core protein was the highest (Fig. 7B).

In addition, we evaluated the infiltration of immune cells in the TCGA. Infiltrated levels of activated B cells, and immature B cells were also lower in the high-risk group than in the low-risk group, and Treg was higher (Figure S3A). Interestingly, the infiltration of Treg was higher in OSCC than that in controls (Fig. S3B-S3C).

#### Discussion

Little progress has been made in long-term survival of OSCC over the past decades, and it remains a fatal disease. Many studies have found that OSCC is a polygenic disease, and researchers can obtain gene expression signatures of specific types of tumors by analyzing the gene expression profiles of tumor samples [16]. In this study, protein expression profiles, metabolic profiles, and gene expression profiles from patients with OSCC were analyzed in TCGA and GEO, and seven core proteins associated with OS were identified. Utilization of the core protein performed well in the prognostic stratification of OSCC patients and may serve as a promising biomarker for patient prognosis prediction.

A gene set for clinical prognosis prediction is already available. For example, Oncotype DX provides a breast cancer recurrence score based on 21 gene expression [17,18]. ColoPrint is a colon cancer recurrence score provided based on 18 gene expression [19,20]. In addition, using different subsets of predicted genes, the study by Yalniz et al., could predict metastasis of OSCC and obtain multiple accurate prediction maps [21]. The core proteins of this study were selected among the prognostic indicators of differentially expressed proteins related to differential metabolites. Patients with OSCC were grouped by the median risk score calculated based on core proteins, and there was a clear difference in the prognosis of patients between groups. PCA results suggested that patients in high- and low risk groups could be highly differentiated.

The core proteins we identified include ASAH1, CSTA, ENDOU, LGALS1, PTX3, SPINK5, and ST6GALNAC1. ASAH1 with decreased expression in this study is an acid ceramidase that converts ceramide to sphingosine and free fatty acids [22]. Recent evidence suggests that ceramide is a bioactive lipid that regulates various signaling pathways and can induce apoptosis [23]. In fact, ASAH1 is downregulated in both lymph node metastasis and extracapsular sprout of oral tongue cell carcinoma [24]. CSTA is mainly expressed in epithelial and lymphoid tissues, and its reduced expression is associated with the recurrence of HNSCC [25]. Consistent with our findings, CSTA was found to be downregulated in the saliva of OSCC patients [26]. CSTA inhibits EMT by preventing cathepsin induced protein cleavage of extracellular

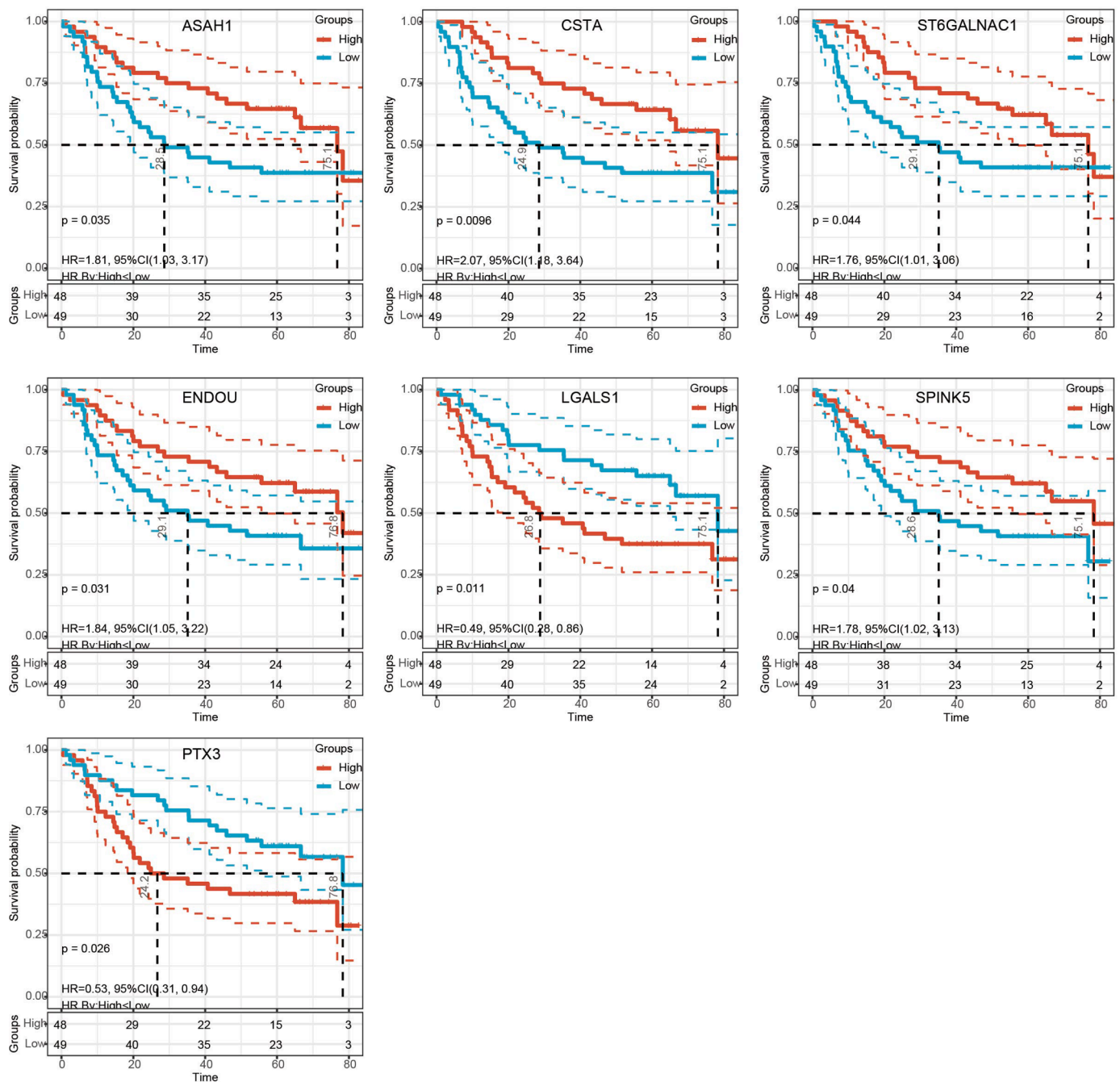


Fig. 3. The K-M curves of core proteins in GSE41613 dataset.

matrix components, adhesion proteins and adherens junctions [27,28]. LGALS1 upregulated in OSCC in our results, which encoded galectin-1, a protein promotes migration and invasion of OSCC [29]. Increased expression of PTX3 is associated with advanced stage of HNSCC and is a promoting factor that mediates EGF induced metastasis of HNSCC [30, 31]. SPINK5 may act as a tumor suppressor in HNSCC and serve as an independent prognostic predictor [32]. Consistent with the results of our analysis, ST6GALNAC1 was protective gene for HNSCC [33].

ENDOU has been shown to be an independent survival marker for HNSCC and negatively correlated with tumor purity and tumor infiltrating macrophages [34]. The results of our analysis showed that ENDOU expression was decreased in both OSCC and high-risk groups. In vitro experiments demonstrated that reduced expression of ENDOU in head and neck cancer has a tumor suppressor role [35]. ENDOU regulated the most metabolites, among which, eflornithine targeted genes

have been a phase II clinical trial (NCT00003076) in the treatment of esophageal cancer [36]. S-Allyl-L-cysteine (SAC) exhibits anticancer, antihepatotoxic, neuroprotective and neurotrophic properties [37]. In addition, apoptosis in glioma cells was induced after NIMUSTINE treatment [38]. Studies have shown that metabolic alterations downstream of proteins are required for their ability to regulate tumours and, in some cases, to achieve productive metastasis [39].

Tumor infiltrating immune cells in the tumor microenvironment play a crucial role in the progression and prognosis prediction of OSCC. ASAH1, CSTA, and ENDOU was negatively associated with Treg, which were higher infiltrated in high-risk group in our results. Compared with normal individuals, the proportion of Tregs in the peripheral blood circulation and cancer microenvironment is significantly higher in HNSCC patients, and the behavior of Tregs can mediate tumor immune evasion [40]. Thus, elevated ASAH1, CSTA, and ENDOU expression in

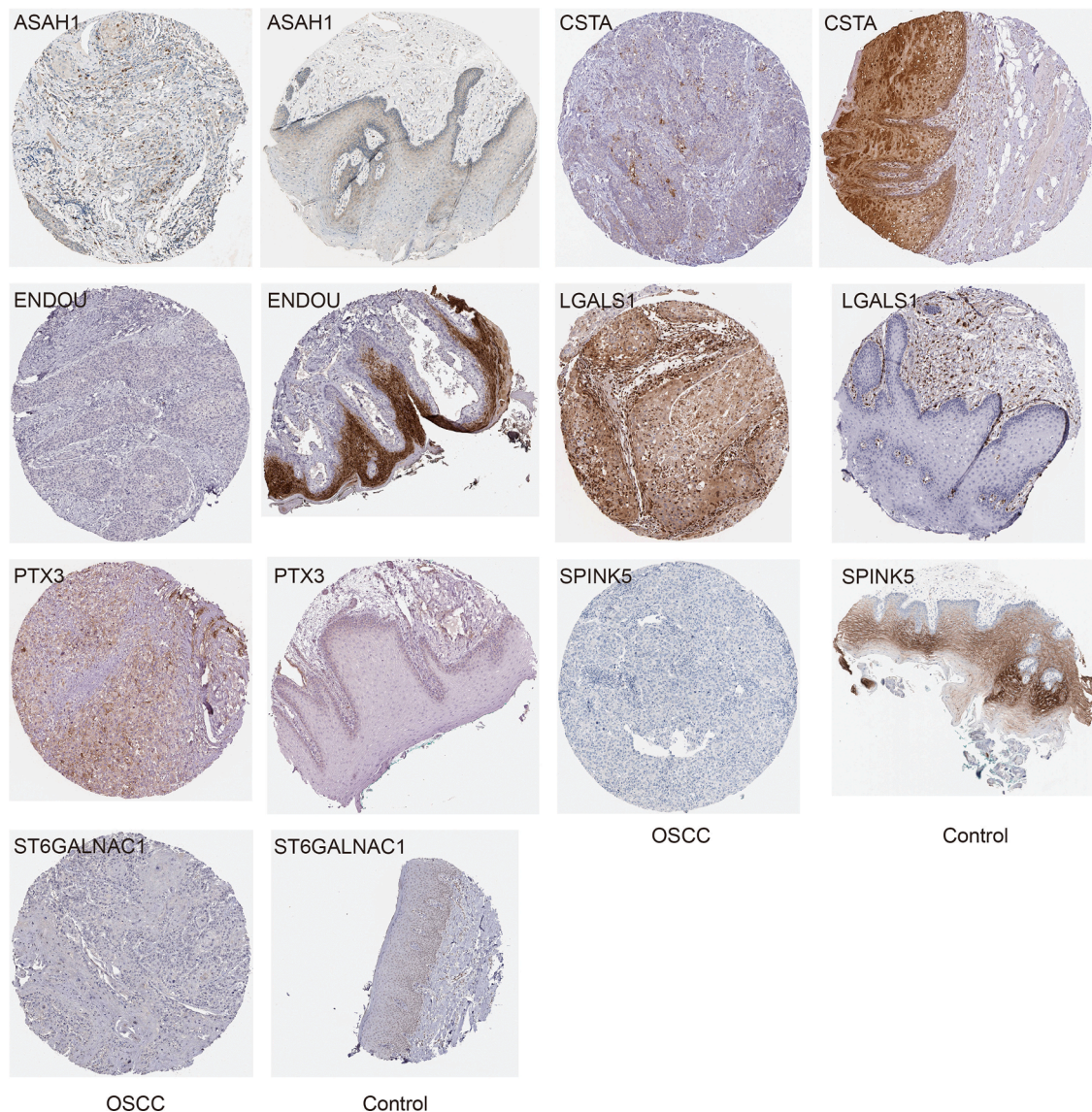


Fig. 4. The immunohistochemical results of core proteins in OSCC and controls in the HPA database.

patients with OSCC, may be a novel target to suppress immune evasion.

To elucidate the in-depth mechanism of core proteins involved in OSCC, we performed GSEA. The results showed that Notch signaling pathway, EMT, and angiogenesis were significantly enriched in the high-risk group. In OSCC, Notch signaling may be supportive through the TME and thereby promote metastasis [41]. EMT is activated during tumor metastasis and is involved in the pathogenesis of OSCC [42,43]. As an important component of the TME, tumor angiogenesis is required for cancer cell spreading, providing nutrient and oxygen support for tumor growth and metastasis [44]. Angiogenesis contributes to cisplatin resistance in OSCC [45].

Our study also had some limitations. First, the number of sequenced samples was small, and further expanded samples are needed for validation. Second, we focused on differential expression at the protein level and evaluated the prognostic value of core proteins using public data, the associates with smoking, alcohol consumption, etc. should be explored in the future. Additionally, the regulatory mechanisms of core proteins and related metabolites have not been explored in depth in cellular and animal experiments. In bioinformatics analysis based both sequencing results and public datasets consistently illustrated the prognostic predictive role of core protein in OSCC, but the clinical availability of core protein still needs to be continuously explored.

## Conclusions

Our study based on transcriptomic, proteomic, and metabolomic approaches, explored 7 core proteins. The core proteins had important implications for the prognosis of OSCC patients and may be used as a survival marker for OSCC. They were involved in regulating metabolites of eflornithine, SAC, NIMUSTINE, and 3-(2,5-dimethoxyphenyl) propanoic acid. GSEA analysis indicated that Notch signaling pathway, EMT, and angiogenesis were significantly activated in the high-risk group.

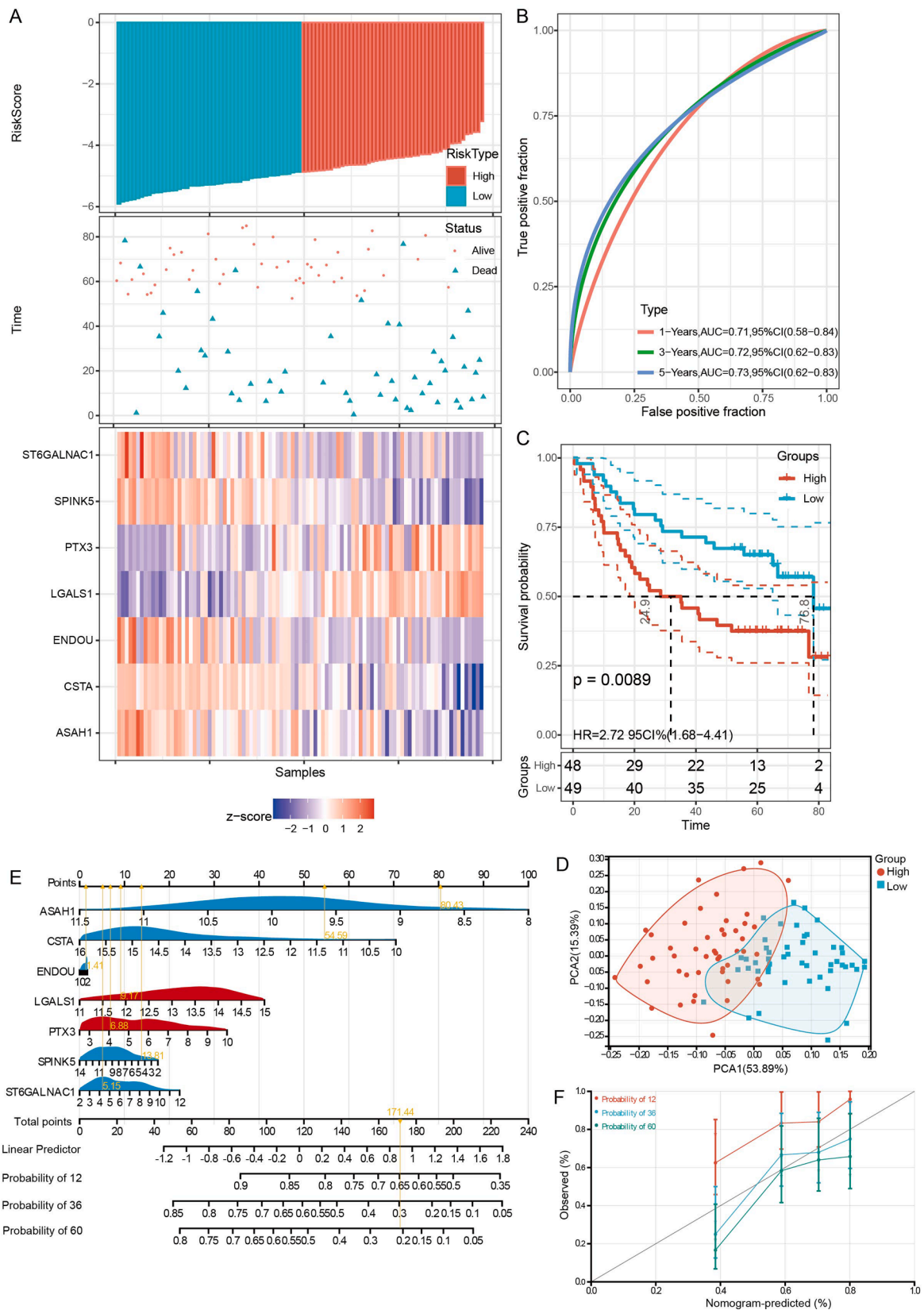
## Data availability

The proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD033080.

The codes used in the manuscript were provided in the supplementary file 1.

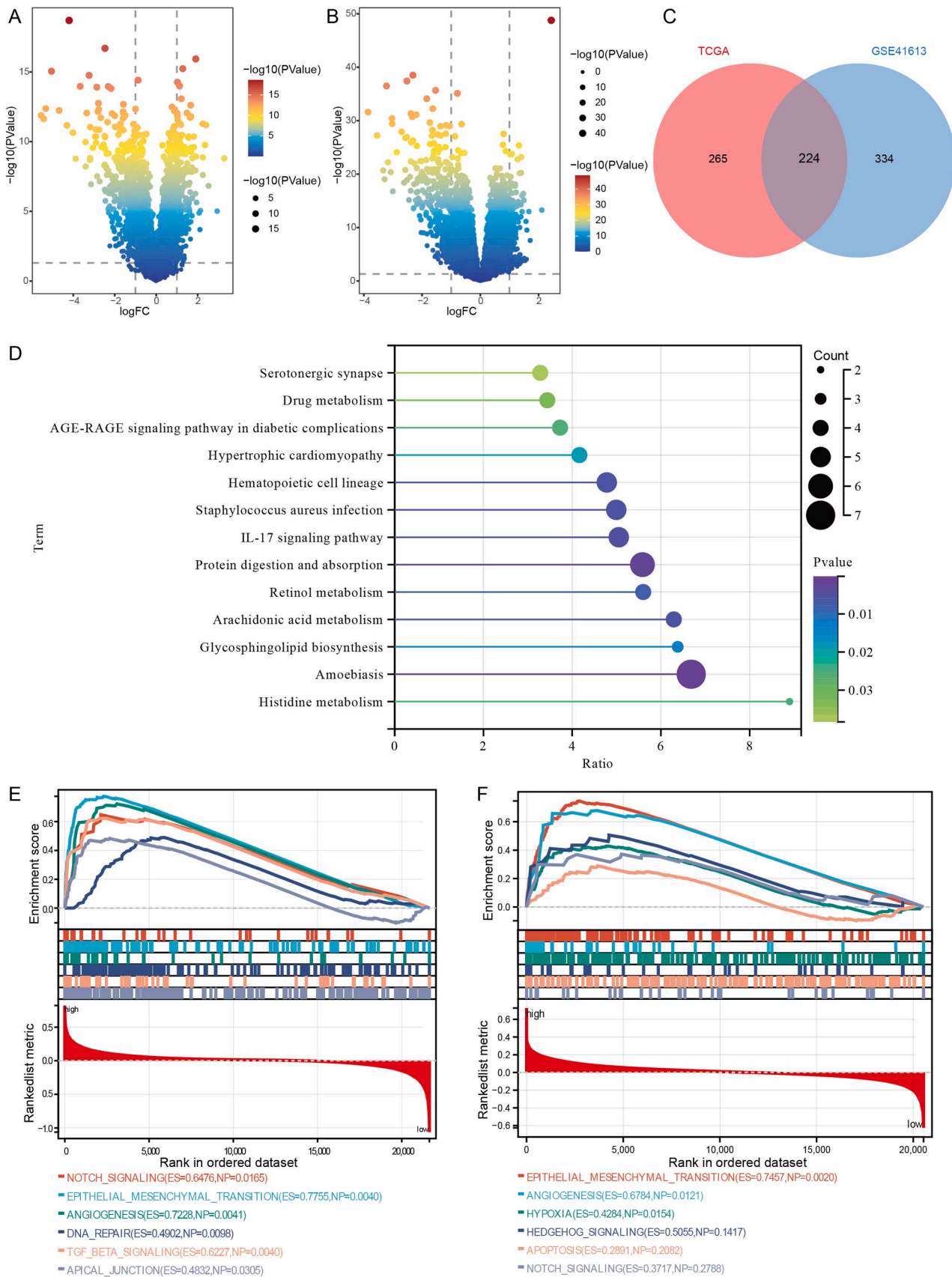
## Funding

There were no fundings.

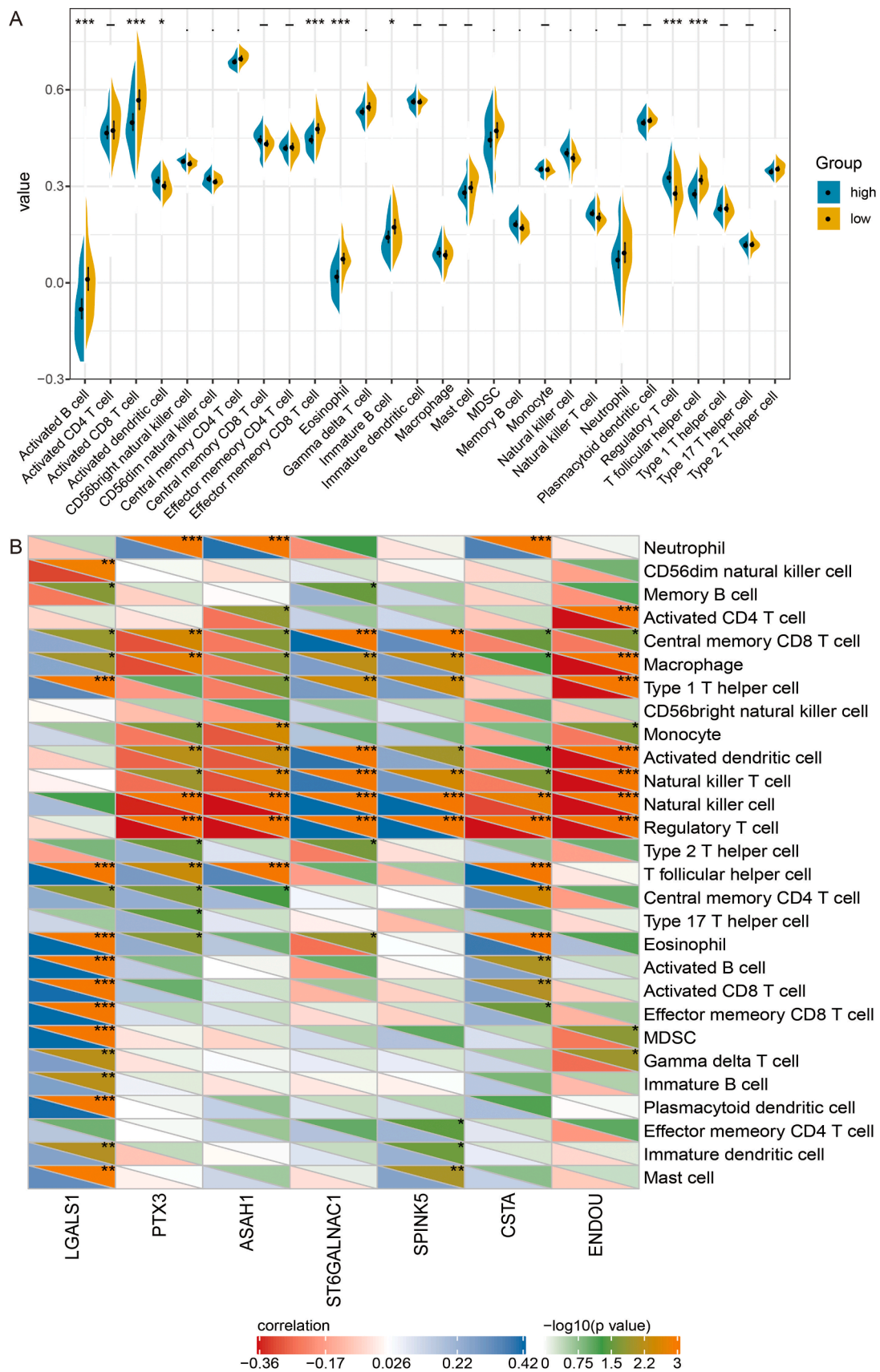


**Fig. 5.** Predictive evaluation of core proteins for the prognosis of OSCC patients in GSE41613 dataset. (A) Risk score model, patient survival status and expression of core proteins in GSE41613 dataset. (B) Time dependent ROC curve of median risk score. (C) K-M curve high-risk and low-risk groups. (D) PCA results of samples in high-risk and low-risk groups. (E) A nomogram was established to visualize the impact of core proteins on patient prognosis. (F) Calibration curves for 1-, 3- and 5-year OS in patients with OSCC from the GSE41613 dataset.





**Fig. 6.** Enrichment and GSEA for DEGs between high- and low-risk groups. (A) Differentially expressed genes between high- and low-risk groups in GSE41613 dataset. (B) Differentially expressed genes between high- and low-risk groups in TCGA. (C) Intersection analysis of two DEGs sets between high- and low-risk groups. TCGA, the Cancer Genome Atlas. (D) KEGG pathways of intersected genes mainly enriched in. GSEA results of intersected genes in GSE41613 dataset (E) and in TCGA (F).



**Fig. 7.** Immune cell infiltration in high- and low-risk groups of GSE41613 dataset. (A) Differences of immune cell infiltration between high- and low-risk groups. \* $P < 0.05$ , \*\*\* $P < 0.001$ . (B) Correlation results between immune cells and core protein. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

## CRedit authorship contribution statement

**Zhitao Yao:** Supervision, Project administration, Data curation, Writing – original draft. **Wei An:** Formal analysis, Data curation, Investigation, Writing – original draft. **Maimaitituxun Tuerdi:** Investigation, Writing – original draft. **Jin Zhao:** Investigation, Writing – original draft.

## Declaration of Competing Interest

The authors declare that they have no competing interests.

## Acknowledgments

The authors thank all the patients participating in this study.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.tranon.2023.101672](https://doi.org/10.1016/j.tranon.2023.101672).

## References

- J. Ferlay, M. Colombet, I. Soerjomataram, D.M. Parkin, M. Pineros, A. Znaor, F. Bray, Cancer statistics for the year 2020: an overview, *Int. J. Cancer* (2021).
- H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA Cancer J. Clin.* 71 (3) (2021) 209–249.
- Q. Li, Y. Hu, X. Zhou, S. Liu, Q. Han, L. Cheng, Role of oral bacteria in the development of oral squamous cell carcinoma, *Cancers (Basel)* 12 (10) (2020) 2797.
- E.E.W. Cohen, R.B. Bell, C.B. Bifulco, B. Burtness, M.L. Gillison, K.J. Harrington, Q. T. Le, N.Y. Lee, R. Leidner, R.L. Lewis, L. Licitra, H. Mehanna, L.K. Mell, A. Raben, A.G. Sikora, R. Uppaluri, F. Whitworth, D.P. Zandberg, R.L. Ferris, The society for immunotherapy of cancer consensus statement on immunotherapy for the treatment of squamous cell carcinoma of the head and neck (HNSCC), *J. Immunother Cancer* 7 (1) (2019) 184.
- K. Omura, Current status of oral cancer treatment strategies: surgical treatments for oral squamous cell carcinoma, *Int. J. Clin. Oncol.* 19 (3) (2014) 423–430.
- M. Tampa, C. Caruntu, M. Mitran, C. Mitran, I. Sarbu, L.C. Rusu, C. Matei, C. Constantin, M. Neagu, S.R. Georgescu, Markers of oral lichen planus malignant transformation, *Dis. Markers* 2018 (2018), 1959506.
- T. Shibahara, [Oral cancer -diagnosis and therapy-], *Clin. Calcium* 27 (10) (2017) 1427–1433.
- A.A. Altuwajri, T.M. Aldrees, M.A. Alessa, Prevalence of metastasis and involvement of level IV and V in oral squamous cell carcinoma: a systematic review, *Cureus* 13 (12) (2021) e20255.
- H. Jin, H. Miao, Y.W. Nie, Y.Y. Lin, Investigating resistin like beta (RETNLB) as a tumor promoter for oral squamous cell carcinoma, *Head Face Med.* 17 (1) (2021) 20.
- L.L. Leung, N.C.H. Lau, J. Liu, X. Qu, S.K. Tsui, J. Hou, C.T. Law, T.H. Ng, J.W. P. Yam, C. Chow, A.B.W. Chan, J.Y.K. Chan, K. Meehan, The role of arginine metabolism in oral tongue squamous cell carcinoma, *Cancers (Basel)* 13 (23) (2021) 6068.
- S.H. Chen, S.Y. Hsiao, K.Y. Chang, J.Y. Chang, New insights into oral squamous cell carcinoma: from clinical aspects to molecular tumorigenesis, *Int. J. Mol. Sci.* 22 (5) (2021) 2252.
- D. Kumar, J. New, V. Vishwakarma, R. Joshi, J. Enders, F. Lin, S. Dasari, W. R. Gutierrez, G. Leef, S. Ponnurangam, H. Chavan, L. Ganaden, M.M. Thornton, H. Dai, O. Tawfik, J. Straub, Y. Shnyder, K. Kakarala, T.T. Tsue, D.A. Girod, B. Van Houten, S. Anant, P. Krishnamurthy, S.M. Thomas, Cancer-associated fibroblasts drive glycolysis in a targetable signaling loop implicated in head and neck squamous cell carcinoma progression, *Cancer Res.* 78 (14) (2018) 3769–3782.
- M. Cerezo, S. Rocchi, Cancer cell metabolic reprogramming: a keystone for the response to immunotherapy, *Cell Death Dis.* 11 (11) (2020) 964.
- A. Fialova, V. Koucky, M. Hajduskova, K. Hladikova, R. Spisek, Immunological network in head and neck squamous cell carcinoma—a prognostic tool beyond HPV status, *Front. Oncol.* 10 (2020) 1701.
- J. Pillai, T. Chincholkar, R. Dixit, M. Pandey, A systematic review of proteomic biomarkers in oral squamous cell cancer, *World J. Surg. Oncol.* 19 (1) (2021) 315.
- S. Wang, T. Li, H. Liu, W. Wei, Y. Yang, C. Wang, B. Li, Z. Han, Z. Feng, A combined prediction model for lymph node metastasis based on a molecular panel and clinicopathological factors in oral squamous cell carcinoma, *Front. Oncol.* 11 (2021), 660615.
- Z.R. Siow, R.H. De Boer, G.J. Lindeman, G.B. Mann, Spotlight on the utility of the Oncotype DX(R) breast cancer assay, *Int. J. Womens Health* 10 (2018) 89–100.
- S.Y. Wang, W. Dang, I. Richman, S.S. Mougalian, S.B. Evans, C.P. Gross, Cost-Effectiveness analyses of the 21-gene assay in breast cancer: systematic review and critical appraisal, *J. Clin. Oncol.* 36 (16) (2018) 1619–1627.
- S. Kopetz, J. Taberner, R. Rosenberg, Z.Q. Jiang, V. Moreno, T. Bachleitner-Hofmann, G. Lanza, L. Stork-Sloots, D. Maru, I. Simon, G. Capella, R. Salazar, Genomic classifier ColoPrint predicts recurrence in stage II colorectal cancer patients more accurately than clinical factors, *Oncologist* 20 (2) (2015) 127–133.
- M. Maak, I. Simon, U. Nitsche, P. Roepman, M. Snel, A.M. Glas, T. Schuster, G. Keller, E. Zeestraten, I. Goossens, K.P. Janssen, H. Friess, R. Rosenberg, Independent validation of a prognostic genomic signature (ColoPrint) for patients with stage II colon cancer, *Ann. Surg.* 257 (6) (2013) 1053–1058.
- P. Roepman, P. Kemmeren, L.F. Wessels, P.J. Slootweg, F.C. Holstege, Multiple robust signatures for detecting lymph node metastasis in head and neck cancer, *Cancer Res.* 66 (4) (2006) 2361–2366.
- A. Gebai, A. Gorelik, Z. Li, K. Illes, B. Nagar, Structural basis for the activation of acid ceramidase, *Nat. Commun.* 9 (1) (2018) 1621.
- P. Malvi, R. Janostiak, A. Nagarajan, X. Zhang, N. Wajapeyee, N-acylsphingosine amidohydrolase 1 promotes melanoma growth and metastasis by suppressing peroxisome biogenesis-induced ROS production, *Mol. Metab.* 48 (2021), 101217.
- X. Zhou, S. Temam, M. Oh, N. Pungpravat, B.L. Huang, L. Mao, D.T. Wong, Global expression-based classification of lymph node metastasis and extracapsular spread of oral tongue squamous cell carcinoma, *Neoplasia* 8 (11) (2006) 925–932.
- P. Strojman, M. Budihna, L. Smid, B. Svetic, I. Vrhovec, J. Kos, J. Skrk, Prognostic significance of cysteine proteinases cathepsins B and L and their endogenous inhibitors stefins A and B in patients with squamous cell carcinoma of the head and neck, *Clin. Cancer Res.* 6 (3) (2000) 1052–1062.
- J.S. Yu, Y.T. Chen, W.F. Chiang, Y.C. Hsiao, L.J. Chu, L.C. See, C.S. Wu, H.T. Tu, H. W. Chen, C.C. Chen, W.C. Liao, Y.T. Chang, C.C. Wu, C.Y. Lin, S.Y. Liu, S.T. Chiou, S.L. Chia, K.P. Chang, C.Y. Chien, S.W. Chang, C.J. Chang, J.D. Young, C.C. Pao, Y. S. Chang, L.H. Hartwell, Saliva protein biomarkers to detect oral squamous cell carcinoma in a high-risk population in Taiwan, *Proc. Natl. Acad. Sci. U. S. A.* 113 (41) (2016) 11549–11554.
- V. Gocheva, J.A. Joyce, Cysteine cathepsins and the cutting edge of cancer invasion, *Cell Cycle* 6 (1) (2007) 60–64.
- M. Fonovic, B. Turk, Cysteine cathepsins and extracellular matrix degradation, *Biochim. Biophys. Acta* 1840 (8) (2014) 2560–2570.
- V. Salunkhe, A. Mahajan, N. Prakash, G.L. Pradeep, R. Patil, S.K. Gajdhar, Galectin-1 expression in oral squamous cell carcinoma: an immunohistochemical study, *J. Oral Maxillofac. Pathol.* 24 (1) (2020) 186.
- Y. Guo, W.K. Pan, Z.W. Wang, W.H. Su, K. Xu, H. Jia, J. Chen, Identification of novel biomarkers for predicting prognosis and immunotherapy response in head and neck squamous cell carcinoma based on cerna network and immune infiltration analysis, *Biomed. Res. Int.* 2021 (2021), 4532438.
- W.C. Chang, S.L. Wu, W.C. Huang, J.Y. Hsu, S.H. Chan, J.M. Wang, J.P. Tsai, B. K. Chen, PTX3 gene activation in EGF-induced head and neck cancer cell metastasis, *Oncotarget* 6 (10) (2015) 7741–7757.
- Z. Lv, K. Wu, X. Qin, J. Yuan, M. Yan, J. Zhang, L. Wang, T. Ji, W. Cao, W. Chen, A novel tumor suppressor SPINK5 serves as an independent prognostic predictor for patients with head and neck squamous cell carcinoma, *Cancer Manag. Res.* 12 (2020) 4855–4869.
- L. Cui, H. Chen, X. Zhao, The prognostic significance of immune-related metabolic enzyme MTHFD2 in head and neck squamous cell carcinoma, *Diagnostics* 10 (9) (2020), 689.
- C. Xu, Y. Zhang, Y. Shen, Y. Shi, M. Zhang, L. Zhou, Integrated analysis reveals ENDOU as a biomarker in head and neck squamous cell carcinoma progression, *Front. Oncol.* 10 (2020), 522332.
- J.R. Becker, D. Gallo, W. Leung, T. Croissant, Y.M. Thu, H.D. Nguyen, T.K. Starr, G. W. Brown, A.K. Bielinsky, Flap endonuclease overexpression drives genome instability and DNA damage hypersensitivity in a PCNA-dependent manner, *Nucleic. Acids. Res.* 46 (11) (2018) 5634–5650.
- W. Yan, J.H. Shih, J. Rodriguez-Canales, M.A. Tangrea, K. Ylaya, J. Hipp, A. Player, N. Hu, A.M. Goldstein, P.R. Taylor, M.R. Emmert-Buck, H.S. Erickson, Identification of unique expression signatures and therapeutic targets in esophageal squamous cell carcinoma, *BMC Res. Notes* 5 (2012) 73.
- P. Bronowicka-Adamska, A. Bentke, M. Lasota, M. Wrobel, Effect of S-Allyl-L-cysteine on MCF-7 cell line 3-mercaptopyruvate sulfurtransferase/sulfane sulfur system, viability and apoptosis, *Int. J. Mol. Sci.* 21 (3) (2020).
- M.T. Tomicic, R. Meise, D. Aasland, N. Berte, R. Kitzinger, O.H. Kramer, B. Kaina, M. Christmann, Apoptosis induced by temozolomide and nimustine in glioblastoma cells is supported by JNK/c-Jun-mediated induction of the BH3-only protein BIM, *Oncotarget* 6 (32) (2015) 33755–33768.
- A. Nagarajan, P. Malvi, N. Wajapeyee, Oncogene-directed alterations in cancer cell metabolism, *Trends Cancer* 2 (7) (2016) 365–377.
- Z. Sun, X. Sun, Z. Chen, J. Du, Y. Wu, Head and neck squamous cell carcinoma: risk factors, molecular alterations, immunology and peptide vaccines, *Int. J. Pept. Res. Ther.* 28 (1) (2022) 19.
- J. Kalafut, A. Czerwonka, A. Anameric, A. Przybyszewska-Podstawa, J. O. Misirok, A. Rivero-Muller, M. Nees, Shooting at moving and hidden targets—tumor cell plasticity and the notch signalling pathway in head and neck squamous cell carcinomas, *Cancers (Basel)* 13 (24) (2021) 6219.
- P. Debnath, R.S. Huirem, P. Dutta, S. Palchaudhuri, Epithelial-mesenchymal transition and its transcription factors, *Biosci. Rep.* 42 (1) (2022). BSR20211754.
- S.S. Shetty, M. Sharma, F.P. Fonseca, P. Jayaram, A.S. Tanwar, S.P. Kabekkodu, S. Kapaettu, R. Radhakrishnan, Signaling pathways promoting epithelial

- mesenchymal transition in oral submucous fibrosis and oral squamous cell carcinoma, *Jpn. Dent. Sci. Rev.* 56 (1) (2020) 97–108.
- [44] S. Zhang, G. Li, C. Liu, S. Lu, Q. Jing, X. Chen, H. Zheng, H. Ma, D. Zhang, S. Ren, Z. Shen, Y. Wang, Z. Lu, D. Huang, P. Tan, J. Chen, X. Zhang, Y. Qiu, Y. Liu, miR-30e-5p represses angiogenesis and metastasis by directly targeting AEG-1 in squamous cell carcinoma of the head and neck, *Cancer Sci.* 111 (2) (2020) 356–368.
- [45] Y. Cheng, S. Li, L. Gao, K. Zhi, W. Ren, The molecular basis and therapeutic aspects of cisplatin resistance in oral squamous cell carcinoma, *Front. Oncol.* 11 (2021), 761379.