

Pathogenesis and prognosis of primary oral squamous cell carcinoma based on microRNAs target genes: a systems biology approach

Amir Taherkhani¹, Shahab Shahmoradi Dehto², Shokoofeh Jamshidi³, Setareh Shojaei^{2*}

¹Research Center for Molecular Medicine, Hamadan University of Medical Sciences, Hamadan 6517838678, Iran

²Department of Oral and Maxillofacial Pathology, Faculty of Dentistry, Hamadan University of Medical Sciences, Hamadan 6517838678, Iran

³Dental Research Center, Department of Oral and Maxillofacial Pathology, School of Dentistry, Hamadan University of Medical Sciences, Hamadan 6517838678, Iran

Oral squamous cell carcinoma (OSCC) is the most prevalent head and neck malignancy, with frequent cervical lymph-node metastasis, leading to a poor prognosis in OSCC patients. The present study aimed to identify potential markers, including microRNAs (miRNAs) and genes, significantly involved in the etiology of early-stage OSCC. Additionally, the main OSCC's dysregulated Gene Ontology annotations and significant signaling pathways were identified. The dataset GSE45238 underwent multivariate statistical analysis in order to distinguish primary OSCC tissues from healthy oral epithelium. Differentially expressed miRNAs (DEMs) with the criteria of p -value < 0.001 and $|\text{Log}_2 \text{fold change}| > 1.585$ were identified in the two groups, and subsequently, validated targets of DEMs were identified. A protein interaction map was constructed, hub genes were identified, significant modules within the network were illustrated, and significant pathways and biological processes associated with the clusters were demonstrated. Using the GEPI2 database, the hub genes' predictive function was assessed. Compared to the healthy controls, main OSCC had a total of 23 DEMs. In patients with head and neck squamous cell carcinoma (HNSCC), upregulation of *CALM1*, *CYCS*, *THBS1*, *MYC*, *GATA6*, and *SPRED3* was strongly associated with a poor prognosis. In HNSCC patients, overexpression of *PIK3R3*, *GIGYF1*, and *BCL2L11* was substantially correlated with a good prognosis. Besides, "proteoglycans in cancer" was the most significant pathway enriched in the primary OSCC. The present study results revealed more possible mechanisms mediating primary OSCC and may be useful in the prognosis of the patients with early-stage OSCC.

Keywords: biomarkers, cancer, head and neck, oral, pathogenesis, prognosis

Introduction

Oral cancer (OC) is a global public health problem, accounting for the fourth highest frequency of malignancy [1]. Oral squamous cell carcinoma (OSCC) is the most common histological type of OC [2-4]. The progression of OSCC is step-by-step and is characterized by the transformation of normal oral epithelium to precancerous lesions and cancerous squamous epithelial tissues [5]. Human papillomavirus infections, smoking, con-

suming alcohol, and having poor oral hygiene are a few risk factors that have been linked to OSCC [6]. The major treatment approach for individuals with late stages of OSCC is surgery, which may be combined with chemotherapy and radiation. This is because many OSCC patients are often diagnosed in advanced stages. OSCC has remained one of the mortal malignancies worldwide, with a 5-year survival rate of approximately 50%, indicating that the prognosis of OSCC patients is poor [7-12]. Thus, illustrating the exact mechanisms and markers underlying the development of OSCC may elevate patient survival [8,10].

Protein-coding genes include approximately 2% of the entire human genome, and therefore, the regulatory role of non-coding RNAs (ncRNAs) in signaling pathways, biological processes (BPs), and networks is inevitable [13,14]. The crucial involvement of ncRNAs (practically microRNAs [miRNAs]) as tumor suppressors or tumor promoter genes in many kinds of malignancies, including OC, has been supported by a sizable number of research published in the last 10 years [15,16]. A group of ncRNAs with a length of 18–22 nucleotides are known as miRNAs. These small RNAs either encourage mRNA degradation or restrict mRNA translation to adversely affect the expression of their target gene(s) [15]. It was reported that miRNAs control up to 50% of the human gene expression [17]. The miRNAs involve in many critical biological procedures, including cell development, metabolism, proliferation, and signal transduction [18].

The present study hypothesized that the dysregulation of miRNAs in the early-stage OSCC tumor specimens (stages I and II) compared to their adjacent normal epithelium leads to the aberrant expression of their target genes. Therefore, differentially expressed miRNAs (DEMs) in early-stage OSCC tumor tissues compared with the adjacent healthy controls were identified using the multivariate statistical analysis (p -value < 0.001 ; $|\text{Log}_2$ fold change (FC)| > 1.585), and the differential power of DEMs was evaluated using the support vector machine (SVM) algorithm. Additionally, it was hypothesized that the DEMs-targets, which are crucial in the pathogenesis of early-stage OSCC, might serve as prognostic indicators for the condition. In order to study the prognostic significance of the hub genes inside the network, we created a protein interaction map (PIM) based on the DEMs-targets. Furthermore, based on DEMs-targets and important clusters in the PIM, Gene Ontology (GO) annotation and pathway enrichment studies were carried out. The aims of the present study were followed by re-analyzing the dataset GSE45238, which was established by Shiah et al. [19] to monitor miRNA expression profile in OSCC tissue specimens and their adjacent non-cancerous epithelium. All tissue samples were collected from patients who under-

went curative surgery from 1999 to 2010 at the National Cheng Kung University Hospital (Tainan, Taiwan). Liquid nitrogen was used to store the freshly frozen samples until usage. The Institutional Human Experiment and Ethics Committee accepted the trial, and tumor staging was carried out using the American Joint Committee on Cancer staging method (HR-97-100). The present results may be beneficial to reach a more pleasant prognosis in the patients with early-stage OSCC, leading to appropriate treatments on behalf of patients. The present study was confirmed by the Ethics Committee of Hamadan University of Medical Sciences, Hamadan, Iran (ethics no. IR.UMSHA.REC.1400.722).

Methods

MiRNA expression dataset recovery

The miRNA expression dataset GSE45238 [19] was attained from the National Center for Biotechnology Information Gene Expression Omnibus, available from: <http://www.ncbi.nlm.nih.gov/geo> [20]. The GSE45238 contained 80 samples including stage I OSCC ($n = 1$), stage II OSCC ($n = 14$), stage III OSCC ($n = 10$), and stage IV OSCC ($n = 15$), and their adjacent non-tumor epithelium, based on the GPL8179 platform (Illumina Human v2 MicroRNA expression bead chip, Illumina, San Diego, CA, USA). To illustrate markers involved in primary OSCC, a new dataset was selected from the GSE45238 for further analyses, including 15 early-stage OSCC tissue samples (stage I [$n = 1$] and stage II [$n = 14$]) and their corresponding healthy control specimens.

Statistical analysis

The normalized dataset after preprocessing had 30 observations (15 early-stage OSCC and 15 nearby healthy controls) and 858 characteristics (miRNA IDs). Using the "ropls," "limma," and "genefilter" packages from the R language (version 4.2.0) [21], the advanced supervised orthogonal partial least squares discriminant analysis (OPLS-DA) was used to find DEMs in early-stage OSCC samples as opposed to healthy controls. DEMs with the criteria of p -value < 0.001 and $1/3 > \text{FC} > 3$ (absolute Log_2 FC [$|\text{Log}_2$ FC]| > 1.585) were considered significant.

SVM classifier

To evaluate the differential power of DEMs in early-stage OSCC tissues compared to the healthy controls, the SVM algorithm with the linear kernel was constructed using the e1071 package (version 1.6-8) within the R environment (<https://cran.rproject.org/web/packages/e1071>). The model's sensitivity, specificity, and accuracy were calculated using the leave-one-out (LOO) approach, and

cross-validation was carried out by repeating LOO 100 times [22]. Furthermore, the receiver operating characteristic (ROC) curve was drawn, and the area under the curve (AUC) value was then calculated to appraise the model's predictive power.

PIM, module, signaling pathway, and GO annotation analyses

MiRWalk 2.0 database [23] was utilized to illustrate the experimentally validated targets of DEMs. In this regard, only targets included in the miRTarBase [24] with a score of 1 were assigned as validated DEMs-targets. Then, all validated targets were used to illustrate cellular components (CCs) and molecular functions (MFs) affected in the early-stage OSCC. The STRING knowledge base webserver ([HTTP://string-db.org](http://string-db.org)) [25] was used to unravel possible connections among the target gene. Subsequently, the Cytoscape (version 3.9.1) [26] was utilized to build a PIM from the targets. The MCODE plugin was used to show major linked areas within the network after single genes were removed from the PIM [27]. Only modules with the following criteria were assigned important and considered for further pathway enrichment and BP annotation analysis: minimum MCODE score = 3, the minimum number of genes = 10, Max depth = 100, Degree cutoff = 2, k-score = 2, and node score cutoff = 0.2. The Pathway and GO annotation enrichment analyses were executed using the g:Profiler tool (<https://biit.cs.ut.ee/gprofiler/gost>) [28], and the terms with the criteria of false discovery rate (FDR) < 0.05 and number of enriched genes ≥ 10 were considered significant. The Network Analyzer tool was used to determine the centrality parameters of the network's nodes in addition to topological analysis, and hub status was given to the genes whose degree and betweenness centrality values were higher than the mean of the network's vertexes.

Survival and boxplot analyses

The prognostic impact of the hub genes associated with the early-stage OSCC was studied utilizing Kaplan-Meier curves. The survival analysis was conducted using GEPIA2 database (<http://gepia2.cancer-pku.cn/#survival>) [29]. Re-analyzing RNA sequencing resources, such as The Cancer Genome Atlas [30] and Genotype-Tissue Expression [31] databases (which are linked to human cancer diseases), is made possible by GEPIA2. This process yields more accurate results for survival, box plot, and correlation analysis, as well as similar gene identification in cancer patients as compared to healthy controls. Cox proportional hazards regression model calculated the corrected hazard ratios (HR) and 95% confidence intervals of the hub genes. Kaplan-Meier curves with the log-rank test p-value and HR p-value as < 0.05 were as-

signed significant. Moreover, the expression patterns of prognostic markers in head and neck squamous cell carcinoma (HNSCC) tissue samples and healthy controls were studied using the valuable data from the GEPIA2 database.

Results

Identification of DEMs and their targets in early-stage OSCC

OPLS-DA was performed to differentiate early-stage OSCC tissue samples (n = 15) from healthy controls (n = 15). The model strongly classified two datasets with the following parameters: $R^2X = 0.223$, $R^2Y = 0.959$, and $Q^2 = 0.789$. At a p-value of 0.001 and $|\text{Log}_2 \text{FC}| > 1.585$, 23 DEMs, including nine up-regulated and 14 down-regulated miRNAs, were identified in the early-stage OSCC tissues compared with non-cancerous specimens (Table 1). Six hundred seventy-two distinct genes were shown in the MiRWalk 2.0 database as verified targets for DEMs. The dataset GSE45238's volcano plot is shown in Fig. 1A thanks to the Shiny applications

Table 1. A total of 23 miRNAs were identified to be differentially expressed in early OSCC tissues compared to healthy controls

miRNA name	Log2 FC	Log2 FC	p-value
hsa-miR-21-3p	2.33	2.33	1.24E-09
hsa-miR-503-5p	2.07	2.07	9.64E-07
hsa-miR-31-3p	2.08	2.08	1.03E-05
hsa-miR-196a-5p	2.19	2.19	1.95E-04
hsa-miR-34b-5p	2.47	2.47	2.11E-04
hsa-miR-34c-5p	1.98	1.98	9.73E-04
hsa-miR-424-3p	1.96	1.96	1.16E-06
hsa-miR-7-5p	1.84	1.84	8.04E-07
hsa-miR-944	1.73	1.73	5.53E-04
hsa-miR-411-5p	-2.35	2.35	6.61E-07
hsa-miR-30a-3p	-2.17	2.17	1.64E-05
hsa-miR-410-3p	-2.73	2.73	7.88E-05
hsa-miR-375-3p	-4.54	4.54	1.24E-04
hsa-miR-487b-3p	-1.97	1.97	3.16E-06
hsa-miR-136-5p	-1.91	1.91	4.41E-04
hsa-miR-495-3p	-1.89	1.89	1.27E-04
hsa-miR-370-3p	-1.84	1.84	3.98E-05
hsa-miR-30a-5p	-1.76	1.76	7.06E-09
hsa-miR-502-5p	-1.73	1.73	1.10E-04
hsa-miR-218-5p	-1.70	1.70	2.45E-06
hsa-miR-30c-1-3p	-1.68	1.68	1.00E-04
hsa-miR-154-5p	-1.66	1.66	3.38E-07
hsa-let-7c-5p	-1.65	1.65	3.82E-08

OSCC, oral tongue squamous cell carcinoma; FC, fold change.

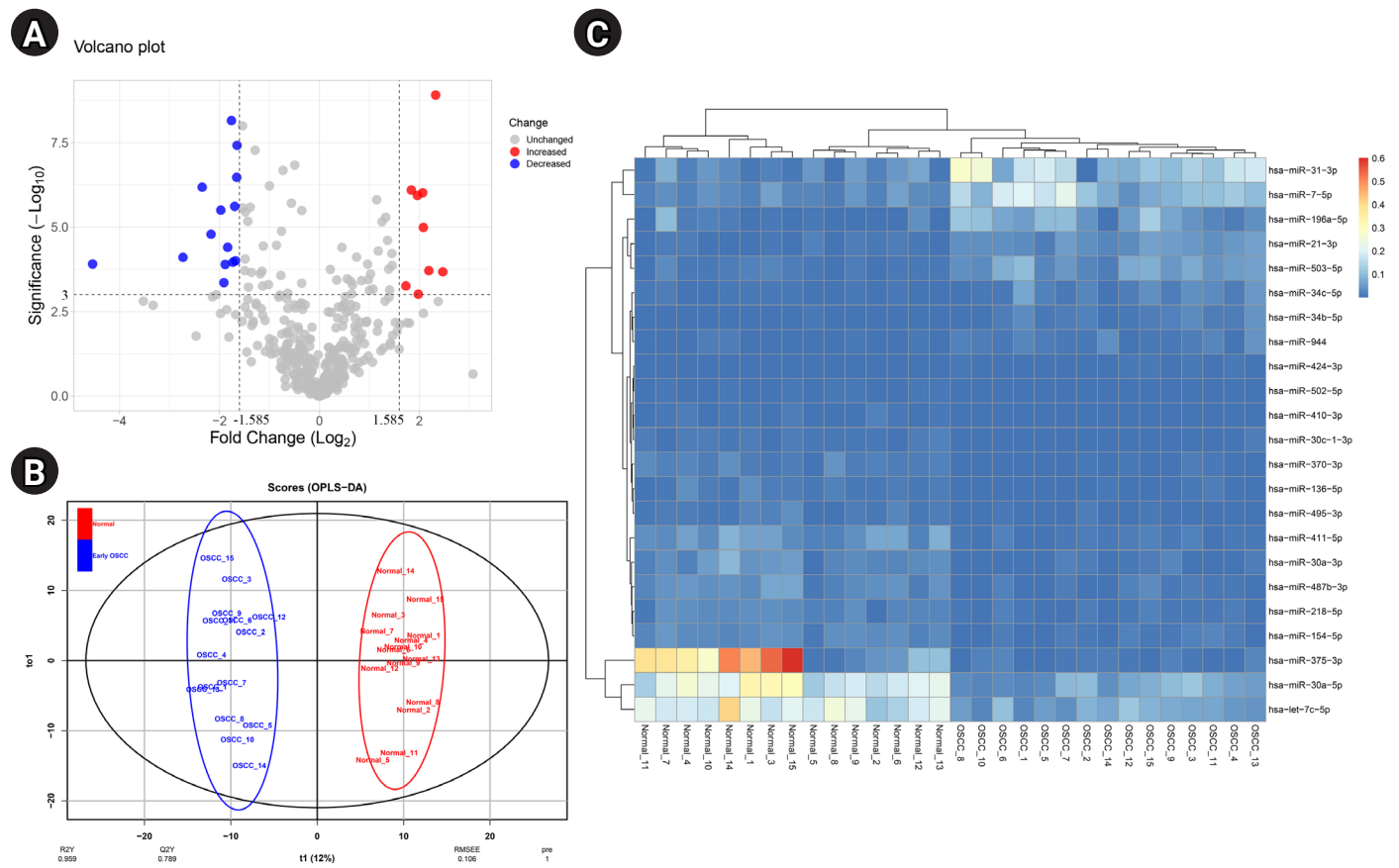


Fig. 1. (A) The volcano plot of miRNAs in the early-stage OSCC tissues compared to the normal epithelium. (B) Score plot in the predictive (x-axis) and orthogonal (y-axis) components of dataset GSE45238 achieved from the OPLS-DA model. (C) Heat-map of DEMs in early-stage OSCC tissues compared with the healthy oral epithelium. OSCC, oral squamous cell carcinoma; OPLS-DA, orthogonal partial least squares discriminant analysis; DEM, differentially expressed miRNA.

tool (available at <https://huygens.science.uva.nl/>) [32]. Fig. 1B displays the OPLS-DA scores plot. Using an unsupervised hierarchical clustering approach in the R environment (version 4.2.0), all early-stage OSCC samples were effectively separated from their nearby healthy controls (Fig. 1C).

SVM modeling and cross-validation

A classifier model was constructed using the SVM algorithm via linear kernel to evaluate the differential power of DEMs in early-stage OSCC samples compared to the normal oral epithelium. The dataset included 23 DEMs and 30 observations (early-stage OSCC = 15; normal = 15). The SVM model has the following parameters: cost = 0.01, number of support vectors = 16, and SVM-type = C-classification. The average value of sensitivity, specificity, and accuracy for the SVM was estimated as 100%, 95%, and 97%, respectively, after completing 100 rounds of LOO cross-validations. The ROC curve of model was also achieved using the “Epi” pack-

age in R language (Fig. 2), in which the AUC of the SVM was determined as 0.955. Supplementary Table 1 shows the details of statistics of the model over the 100 times cross-validations.

PIM network, clustering, and functional analyses

A total of 672 unique targets were uploaded into the STRING database, and a primary protein-protein interaction (PPI) network was constructed with a confidence score of ≥ 0.4 . A PIM meeting the criterion of 605 proteins and 2,322 interactions was created after unconnected nodes were removed, and it was then imported into the Cytoscape program for further analysis. The MCODE program found five major modules in all, including clusters Nos. 1, 2, 3, 6, and 7 (Fig. 3). Besides, 85 genes indicated a higher degree and betweenness centralities compared to the average value of the mentioned parameters within the PPI network therefore, were assigned as hub genes associated with the pathogenesis of early-stage OSCC (Supplementary Table 2). The average value for degree

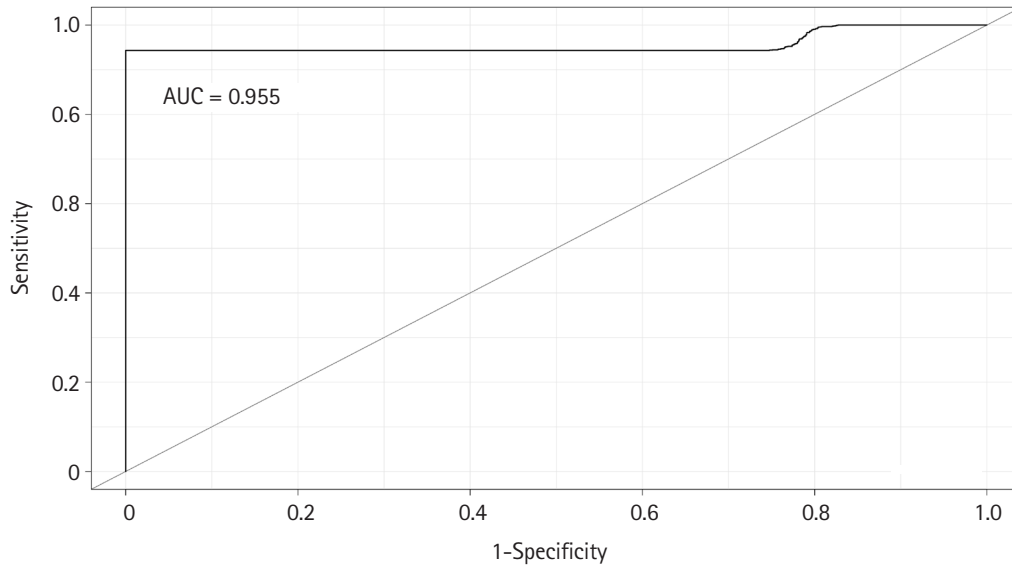


Fig. 2. ROC curve was completed to test the predictive power of the SVM classifier with a linear kernel. The AUC was calculated as 0.955. ROC, receiver operating characteristic; SVM, support vector machine; AUC, area under the curve.

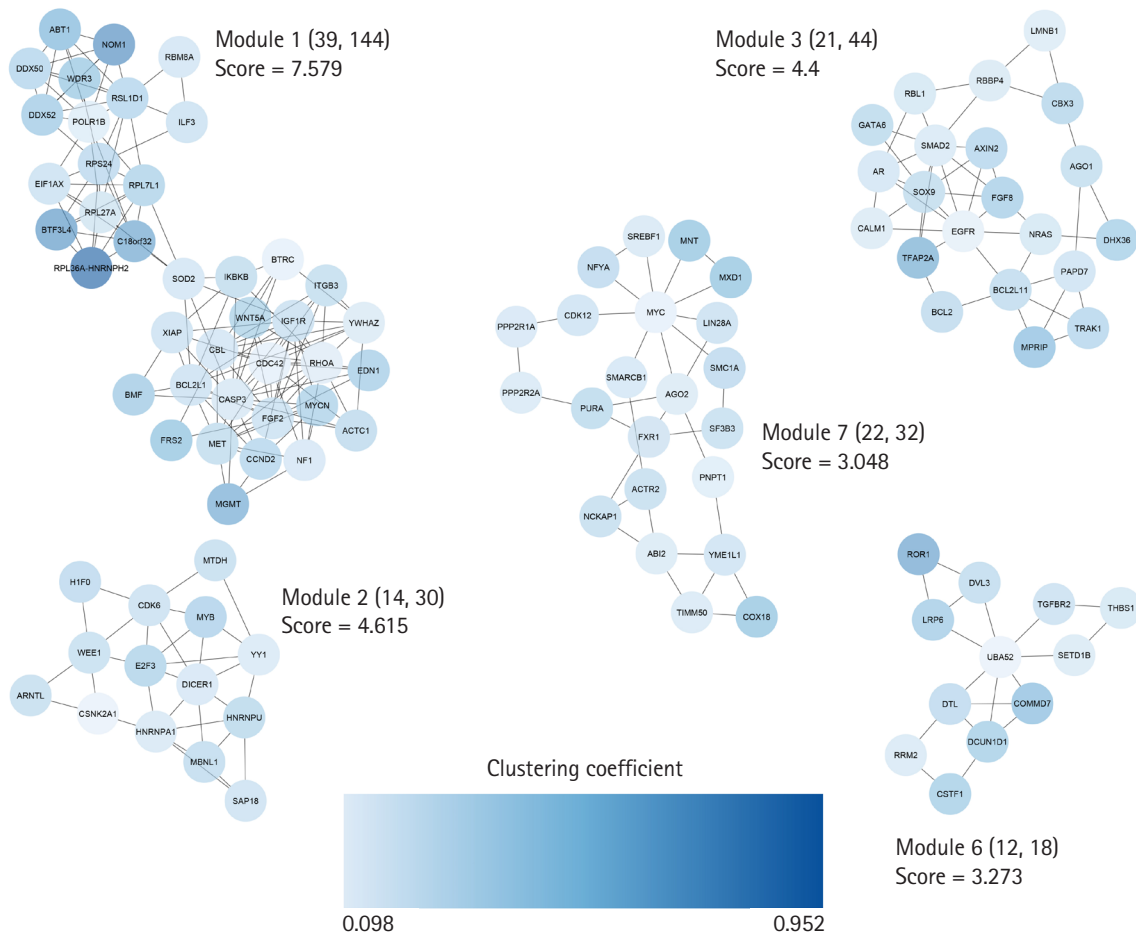


Fig. 3. Clustering analysis was performed with the MCODE plugin. Five salient modules were detected in the PIM linked to the early-stage OSCC. PIM, protein interaction map; OSCC, oral squamous cell carcinoma.

and betweenness centrality was determined as 7.68 and 0.0060, respectively. At an FDR of 0.05 and a number of enriched genes ≥ 10 , a total of 11 MFs, 26 CCs, 174 BPs, and two pathways were found to be significantly dysregulated in early-stage OSCC. Fig. 4 shows top-10 significant MFs, CCs, and BPs, as well as two signaling pathways enriched in the early-stage of OSCC.

Prognostic role of the hub genes and expression analysis

The overexpression of six hub genes, including *CALM1*, *CYCS*, *THBS1*, *MYC*, *GATA6*, and *SPRED3*, was strongly related with a poor prognosis in patients with HNSCC, according to the survival analysis performed using the GEPIA2 database. In contrast, upregulation of three of hubs, including *PIK3R3*, *GIGYF1*, and *BCL2L11* was considerably related to a favorable prognosis in HNSCC patients (log-rank test and HR p-values < 0.05). Thus, a prognostic panel including *CALM1* and *CYCS* revealed the most significant result with the HR = 1.8 and $p(\text{HR}) = 0.000042$ (Table 2, Fig. 5).

According to the expression analysis from GEPIA2, all prognostic markers in HNSCC exhibited higher expression in cancerous tissues than the healthy controls (Fig. 6).

Discussion

A dismal prognosis is associated with OC, which continues to be one of the most prevalent carcinomas globally. Additionally, the paucity of reliable markers to forecast OC development contributes to treatment failures [33,34]. The current investigation discovered 23 miRNAs that were differently expressed between healthy control epithelium and primary OSCC tissues. The miR-21-3P indicated the most overexpression in primary OSCC tissues compared to the adjacent normal specimens with the criteria of $\text{Log}_2 \text{FC} = 2.33$ and p-value of $1.24\text{E-}09$. Previous studies have assigned miR-21 as a predictive, diagnostic, and prognostic marker in several human cancers [35]. MiR-21 overexpression is linked to increased cellular proliferation, aggressive behavior, and metastasis because miR-21 targets multiple tumor suppressor genes [36-39]. According to Amirfallah et al. [40], overexpression of miR-21-3p was substantially linked to advanced-stage breast cancer, lymph-node metastases, a low likelihood of survival, and big tumor sizes. The authors reported that miR-21-3p overexpression was related to the downregulation of tumor suppressor genes, suggesting miR-

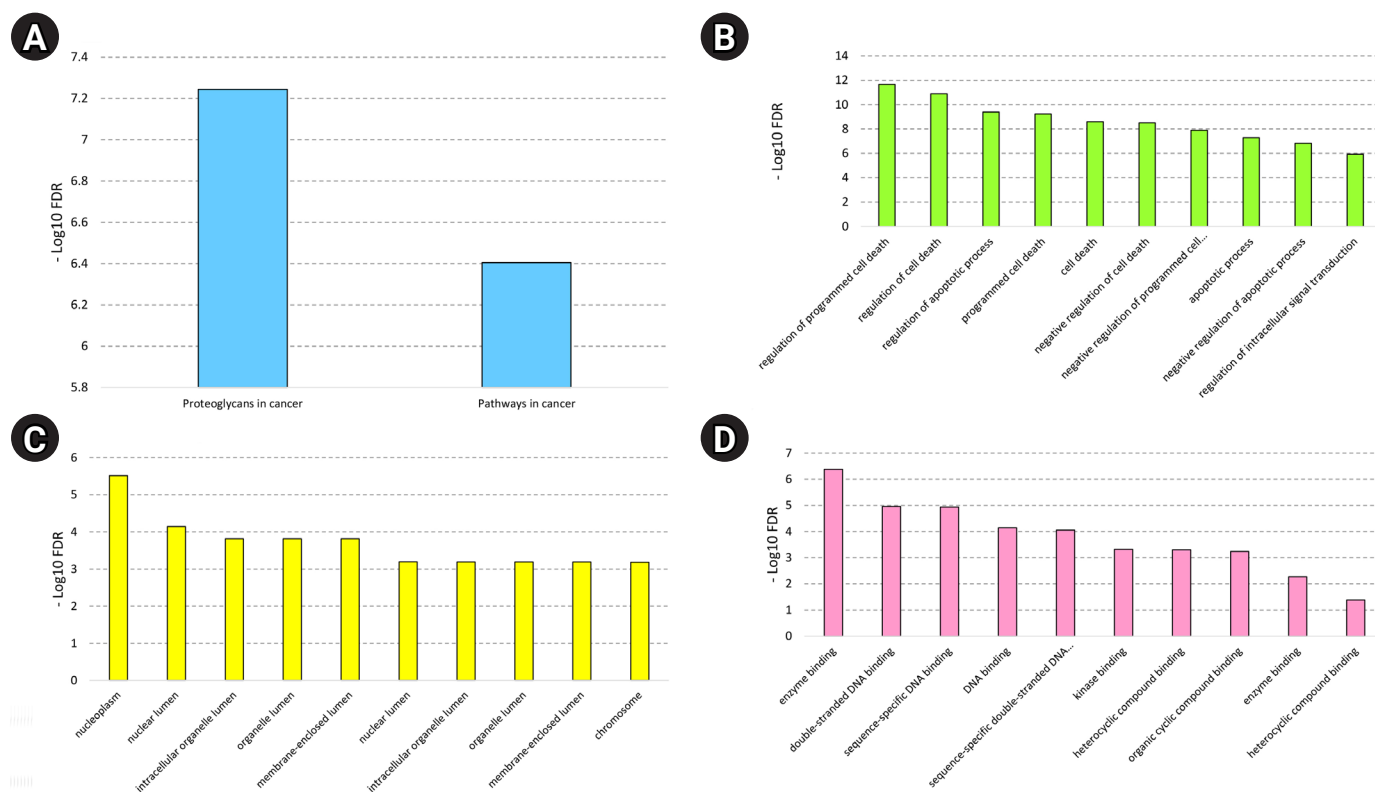


Fig. 4. Most significant signaling pathways (A), biological processes (B), cellular components (C), and molecular functions (D) are affected in the early-stage OSCC. The X-axis shows the name of the term. Y-axis demonstrates the minus value of the Log₁₀ FDR. FDR, false discovery rate; OSCC, oral squamous cell carcinoma.

Table 2. Nine of the hub genes in early OSCC significantly demonstrated a prognostic role in HNSCC

Gene symbol (gene label)	HR (high)	p (log-rank test)	p (HR)
Single-gene			
<i>CALM1</i> (A)	1.6	0.001	0.0011
<i>CYCS</i> (B)	1.4	0.009	0.0096
<i>THBS1</i> (C)	1.4	0.0099	0.01
<i>MYC</i> (D)	1.3	0.028	0.029
<i>GATA6</i> (E)	1.3	0.047	0.049
<i>SPRED3</i> (F)	1.3	0.049	0.049
<i>PIK3R3</i> (G)	0.63	0.00073	0.00081
<i>GIGYF1</i> (H)	0.68	0.0041	0.0042
<i>BCL2L11</i> (I)	0.76	0.041	0.042
Prognostic panel			
Combination of genes			
A + B	1.8	0.000033	0.000042
A + B + C	1.5	0.0015	0.0017
A + B + C + D	1.5	0.002	0.0022
A + B + C + D + E	1.4	0.0069	0.0072
A + B + C + D + E + F	1.6	0.0011	0.0012
G + H	0.73	0.021	0.022
G + H + I	0.75	0.033	0.033

OSCC, oral squamous cell carcinoma; HNSCC, head and neck squamous cell carcinoma; HR, hazard ratio.

21-3p as a potential biomarker in breast tumors. In addition, Gao et al. [41] showed that miR-21-3p upregulation was significantly associated with stemness maintenance of cancer stem cells and a high risk of esophageal squamous cell carcinoma.

Moreover, the present study indicated that has-miR-375-3p was the most under-expressed miRNA in primary OSCC than normal adjacent tissues (Log2 FC = -4.54 and p = 1.24E-04). Crimi et al. [42] performed a study to determine the expression levels of four miRNAs including hsa-miR-375-3p, hsa-miR-133a-3p, hsa-miR-503-5p, and hsa-miR-196a-5p in liquid biopsies obtained from OC patients and healthy individuals. With p-values of 0.0001 and 0.001, respectively, the droplet digital polymerase chain reaction revealed lower plasma levels of hsa-miR-133a-3p and hsa-miR-375-3p in OC compared to healthy controls. Additionally, Crimi et al. [42] demonstrated that hsa-miR-375-3p and hsa-miR-133a-3p were highly sensitive and specific diagnostic indicators for OC, with AUCs of 0.96 and 0.86, respectively. The significant role of hsa-miR-375-3p in the development of colorectal cancer (CRC) was shown. Xu et al. [43] reported a significant association between miR-375-3p downregulation and a dismal survival rate in CRC patients. Besides, upregulation of miR-375-3p elevated the

sensitivity of CRC cells to 5-fluorouracil-based chemotherapy through negative regulation of YAP1 and SP1.

Overexpression of the DEMs target genes *CALM1*, *CYCS*, *THBS1*, *MYC*, *GATA6*, and *SPRED3* was strongly associated with a poor prognosis in patients with HNSCC. Additionally, patients with HNSCC who had upregulation of *PIK3R3*, *GIGYF1*, and *BCL2L11* had better prognoses. Calmodulin (*CALM*) is a major conserved calcium ion receptor protein in eukaryotic cells and is encoded by *CALM1*, *CALM2*, and *CALM3* in humans [44]. The *CALM*/Ca²⁺ binds to the phosphoinositide 3-kinase α (*PI3K α*), leading to the hyperactivation of *PI3K α* /Akt/mammalian target of rapamycin signaling pathway, cell proliferation, differentiation, motility, and development [45-47]. The overexpression of *CALM1* was indicated in many cancers including prostate cancer [48], bladder cancer [49], and nasopharyngeal carcinoma [50].

A significant part in promoting the signaling pathways connected to programmed cell death is played by cytochrome c (*CYCS*) [51]. However, *CYCS* is a poor prognostic indicator in the majority of malignancies, which may be a result of a malignant tissue response [52,53]. Thrombospondin-1 (*THBS1*) is an adhesive glycoprotein. *THBS1* plays a role in wound healing by mediating cell adhesion, proliferation, and migration. Moreover, *THBS1* is involved in the cancer cell and metastasis [54]. The upregulation of *THBS1* was illustrated in different cancer cells such as ovarian cancer [55], mammary cancer [56], and thyroid cancer [57]. Hayashido et al. [54] studied the expression of *THBS1* at the protein level in OSCC tissues using the immunohistochemical assay. In the OSCC cells, *THBS1* induced the production of matrix metalloproteinase-9, which increased proteolytic activity and haptotactic migration, according to the findings of the research by Hayashido et al. [54]. Furthermore, Xiao et al. [9] showed that exosomes carrying the *THBS1* protein that were released from OC tissues might facilitate the conversion of macrophages into tumor-associated macrophages that resemble M1. Myc proto-oncogene protein (*MYC*) is a well-known transforming gene in many cancers such as breast [58] and lung cancer [59], as well as hepatocellular carcinoma [60]. *MYC* is up-regulated in 80% of the OSCC cases [61] and has shown high frequency by Genome-wide profiling of OSCC [62]. Wang et al. [10] reported that the miR-1294 targeted the 3'-untranslated region of c-*MYC* and led to the downregulation of *MYC* mRNA in OSCC cancer cells.

GATA6 (*GATA* binding protein 6) is a transcription factor involved in a *GATA* family regulating several genes participating in tissue morphogenesis and cell fate decision making [63]. *GATA6* has been described as an independent prognostic marker in ovarian cancer and plays a key role in the differentiation and invasion of

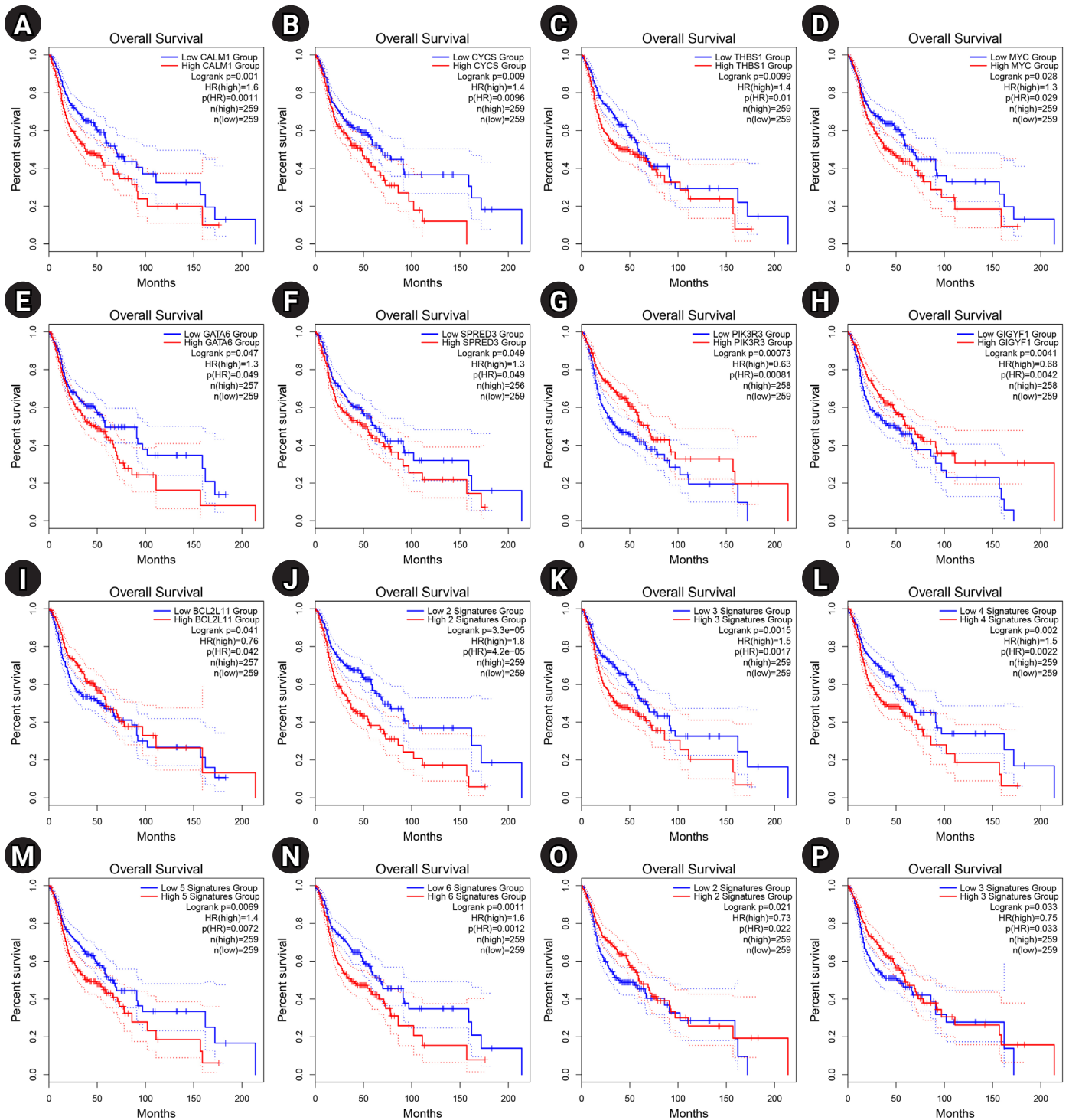


Fig. 5. (A–P) Prognostic impact of *CALM1* (A), *CYCS* (B), *THBS1* (C), *MYC* (D), *GATA6* (E), *SPRED3* (F), *PIK3R3* (G), *GIGYF1* (H), and *BCL2L11* (I) was significant in HNSCC. Prognostic role of the panel of genes is also exhibited. The X-axis and Y-axis represent the survival time of patients with HNSCC and the survival probability, respectively. The dotted lines are 95% confidence intervals. HNSCC, head and neck squamous cell carcinoma; HR, hazard ratio.

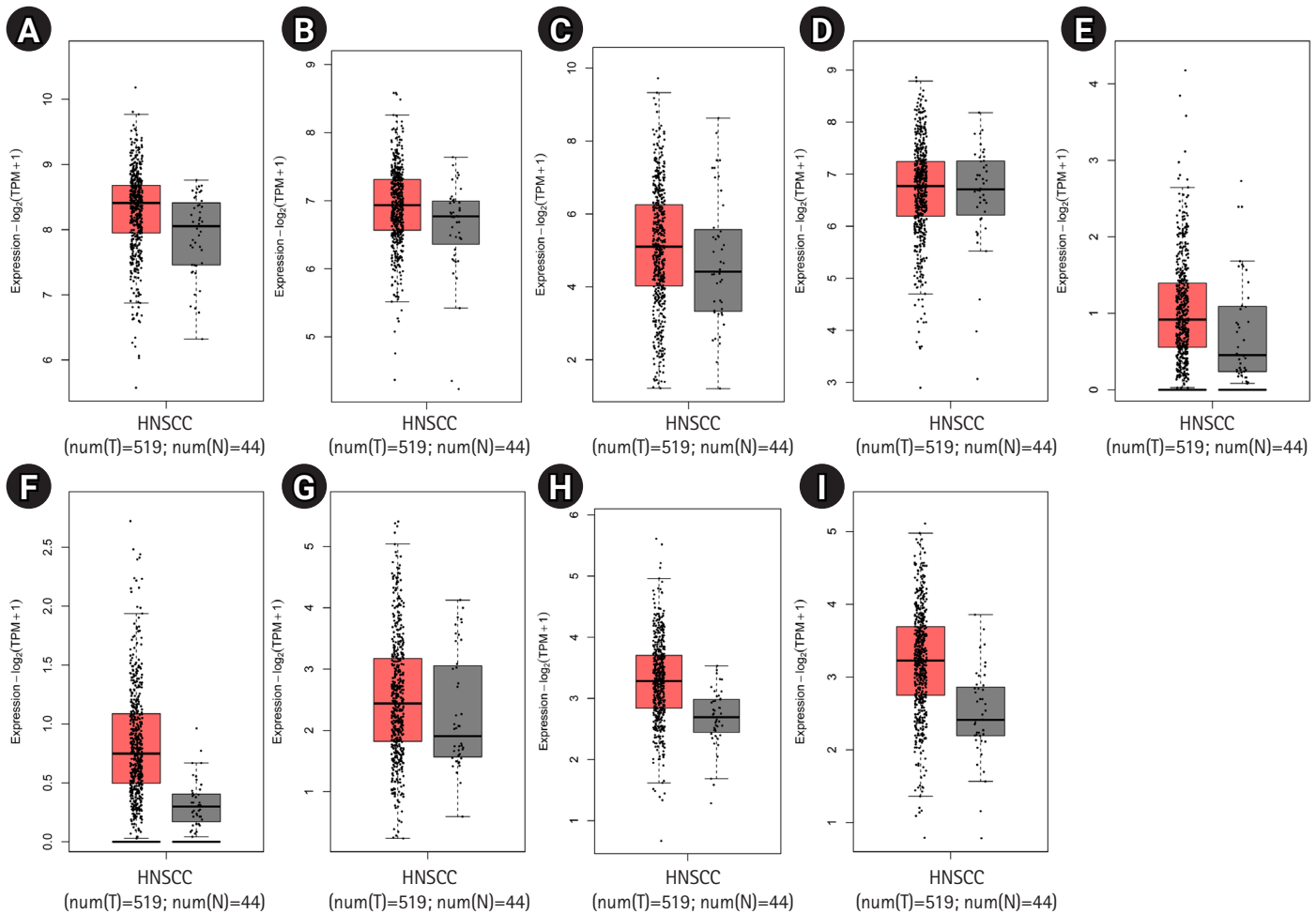


Fig. 6. Gene expression of prognostic markers in HNSCC from the expression analysis in the GEPIA2 database. Box plots are based on 44 normal tissues (gray color) and 519 HNSCC samples (red color). The boxplot analysis show overexpression of *CALM1* (A), *CYCS* (B), *THBS1* (C), *MYC* (D), *GATA6* (E), *SPRED3* (F), *PIK3R3* (G), *GIGYF1* (H), and *BCL2L11* (I) in HNSCC. However, *MYC* shows a mild upregulation in HNSCC. HNSCC, head and neck squamous cell carcinoma.

tumor cells [64]. *GATA6* upregulation has been shown in several tumors in the past, including CRC [65], gastric cancer [66], and cholangiocarcinoma [67]. According to Zhai and Luo [8], OSCC cell lines (HN4, HN6, SCC9, and Cal27) overexpress *GATA6* in comparison to human normal oral mucosal epithelial keratinocytes. Furthermore, *GATA6* knockdown suppressed the colony formation, proliferation, migration, and invasion of cancerous cells. According to a previous study, *GATA6* expression was diminished by miRNA-506 and attenuated the progression of OSCC [68].

Sprouty-related, EVH1 domain-containing protein 3 (*SPRED3*) downregulates the Ras/Raf/mitogen-activated protein kinase (MAPK) signaling pathway during the organogenesis [69]. It was indicated that *SPRED3* is frequently loosed in glioblastoma [70].

Furthermore, He et al. [71] showed that non-small cell lung cancer with repressed *SPRED3* gene expression increased Ras/Raf/MAPK signaling, resulting in resistance to epidermal growth factor receptor-tyrosine kinase inhibitors. According to the present results, the has-miR-370-3p was significantly down-regulated in the early-stage OSCC tissues than the healthy controls ($p = 0.0000398$; $\text{Log}_2 \text{FC} = -1.84$), which may be hypothesized that *SPRED3* is overexpressed in the early-stage OSCC tissues, which may be in line with the cancer response. However, this requires further confirmation.

In order to demonstrate important pathways and GO annotations impacted in main OSCC, gene set enrichment analysis was performed. The genes linked to prominent clusters in the PPI network connected to the etiology of the illness were used for this.

The most important route, then, was "proteoglycans in cancer" (KEGG:05200). Proteoglycans are classified as highly glycosylated proteins involved in various parts of tissues, including the cell membrane, extracellular matrix, and pericellular space. These heavy molecules mediate many important BPs associated with angiogenesis, tissue regeneration, cell growth, and motility. Numerous forms of cancer have been linked to aberrant proteoglycan expression [72]. In solid tumors and hematological malignancies, proteoglycans make cancer cells more aggressive [73]. Additionally, proteoglycans participate in the receptor-ligand interaction between cancer cells and their environment, which may facilitate cell migration and increase the number of circulating tumor cells [74]. Besides, previous studies indicated that heparan sulfate proteoglycans (HSPGs) are involved in cellular differentiation and proliferation, apoptosis, immune evasion, angiogenesis, and matrix remodeling in cancer. Therefore, targeting proteoglycans and HSPGs could be assigned as a practical therapeutic approach in cancer [75,76].

In conclusion, the present study identified 23 DEMs in primary OSCC tissues compared to the normal oral epithelium. MiR-21-3p and miR-375-3p were the most salient up-regulated and down-regulated miRNAs in primary OSCC tissues compared to healthy controls, respectively. It has been hypothesized that DEMs contribute to OSCC's onset or undergo dysregulation in response to a tumor mass. Additionally, there was a strong correlation between a poor outcome and the upregulation of *CALMI*, *CYCS*, *THBS1*, *MYC*, *GATA6*, *SPRED3*, and the under-expression of *PIK3R3*, *GIGYF1*, and *BCL2L1* in HNSCC patients. Furthermore, "proteoglycans in cancer" was demonstrated as the most salient signaling pathway affected in primary OSCC. These results may help the prognosis of the patients with early-stage OSCC, leading to more useful therapeutic approaches.

ORCID

Amir Taherkhani: <https://orcid.org/0000-0002-6546-8785>

Shahab Shahmoradi Dehto: <https://orcid.org/0000-0002-0780-4453>

Shokoofeh Jamshidi: <https://orcid.org/0000-0002-0646-1882>

Setareh Shojaei: <https://orcid.org/0000-0001-8323-1231>

Authors' Contribution

Conceptualization: AT, SS, SJ. Data curation: AT, SS, SMD, SJ.

Formal analysis: AT, SMD, SS. Methodology: AT, SMD, SS, SJ.

Writing - original draft: AT. Writing - review & editing: SS, SJ.

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

The authors would like to appreciate the Dental Research Center, Deputy of Research and Technology, and Research Center for Molecular Medicine, Hamadan University of Medical Sciences, Hamadan - Iran for their supports. This paper was extracted from the thesis of Shahab Shahmoradi Dehto.

Supplementary Materials

Supplementary data can be found with this article online at <http://www.genominfo.org>.

References

1. Chou CH, Chou YE, Chuang CY, Yang SF, Lin CW. Combined effect of genetic polymorphisms of *AURKA* and environmental factors on oral cancer development in Taiwan. *PLoS One* 2017;12:e0171583.
2. Rivera C. Essentials of oral cancer. *Int J Clin Exp Pathol* 2015;8:11884-11894.
3. Reyes-Gibby CC, Anderson KO, Merriman KW, Todd KH, Shete SS, Hanna EY. Survival patterns in squamous cell carcinoma of the head and neck: pain as an independent prognostic factor for survival. *J Pain* 2014;15:1015-1022.
4. Lee WH, Chen HM, Yang SF, Liang C, Peng CY, Lin FM, et al. Bacterial alterations in salivary microbiota and their association in oral cancer. *Sci Rep* 2017;7:16540.
5. Saintigny P, Zhang L, Fan YH, El-Naggar AK, Papadimitrakopoulou VA, Feng L, et al. Gene expression profiling predicts the development of oral cancer. *Cancer Prev Res (Phila)* 2011;4:218-229.
6. Mehrotra R, Singh M, Kumar D, Pandey AN, Gupta RK, Sinha US. Age specific incidence rate and pathological spectrum of oral cancer in Allahabad. *Indian J Med Sci* 2003;57:400-404.
7. Dave K, Ali A, Magalhaes M. Increased expression of PD-1 and PD-L1 in oral lesions progressing to oral squamous cell carcinoma: a pilot study. *Sci Rep* 2020;10:9705.
8. Zhai J, Luo G. *GATA6* induced *FN1* activation promotes the proliferation, invasion and migration of oral squamous cell carcinoma cells. *Mol Med Rep* 2022;25:102.

9. Xiao M, Zhang J, Chen W, Chen W. M1-like tumor-associated macrophages activated by exosome-transferred THBS1 promote malignant migration in oral squamous cell carcinoma. *J Exp Clin Cancer Res* 2018;37:143.
10. Wang Z, Yan J, Zou T, Gao H. MicroRNA-1294 inhibited oral squamous cell carcinoma growth by targeting c-Myc. *Oncol Lett* 2018;16:2243-2250.
11. Ries LA, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, et al. SEER cancer statistics review, 1975-2002. Bethesda: National Cancer Institute, 2002. Accessed 2022 Jun 10. Available from: https://seer.cancer.gov/archive/csr/1975_2002/.
12. Torabi M, Haghani J, Hashemipour MA, Ebrahimi M. Mast cells density in hyperkeratosis, dysplastic oral mucosa and oral squamous cell carcinoma. *Avicenna J Dent Res* 2018;10:67-70.
13. Chawla JP, Iyer N, Soodan KS, Sharma A, Khurana SK, Priyadarshni P. Role of miRNA in cancer diagnosis, prognosis, therapy and regulation of its expression by Epstein-Barr virus and human papillomaviruses: with special reference to oral cancer. *Oral Oncol* 2015;51:731-737.
14. Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet* 2006;15 Spec No 1:R17-R29.
15. Solomon MC, Radhakrishnan RA. MicroRNAs: the vibrant performers in the oral cancer scenario. *Jpn Dent Sci Rev* 2020;56:85-89.
16. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer* 2006;6:857-866.
17. Wienholds E, Plasterk RH. MicroRNA function in animal development. *FEBS Lett* 2005;579:5911-5922.
18. Cheng G, Danquah M, Mahato RI. MicroRNAs as therapeutic targets for cancer. In: *Pharmaceutical Perspectives of Cancer Therapeutics* (Lu Y, Mahato RI, eds.). New York: Springer, 2009. pp. 441-444.
19. Shiah SG, Hsiao JR, Chang WM, Chen YW, Jin YT, Wong TY, et al. Downregulated miR329 and miR410 promote the proliferation and invasion of oral squamous cell carcinoma by targeting Wnt-7b. *Cancer Res* 2014;74:7560-7572.
20. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets: update. *Nucleic Acids Res* 2013;41:D991-D995.
21. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2013.
22. Oskouie AA, Ahmadi MS, Taherkhani A. Identification of prognostic biomarkers in papillary thyroid cancer and developing non-invasive diagnostic models through integrated bioinformatics analysis. *Microna* 2022;11:73-87.
23. Dweep H, Gretz N. miRNAWalk2.0: a comprehensive atlas of mi-croRNA-target interactions. *Nat Methods* 2015;12:697.
24. Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, Chan WL, et al. miRTarBase: a database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res* 2011;39:D163-D169.
25. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 2021;49:D605-D612.
26. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2011;27:431-432.
27. Bayat Z, Ahmadi-Motamayel F, Parsa MS, Taherkhani A. Potential biomarkers and signaling pathways associated with the pathogenesis of primary salivary gland carcinoma: a bioinformatics study. *Genomics Inform* 2021;19:e42.
28. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler: a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res* 2007;35:W193-W200.
29. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017;45:W98-W102.
30. Cancer Genome Atlas Research Network; Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* 2013;45:1113-1120.
31. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580-585.
32. Goedhart J, Luijsterburg MS. VolcanoR is a web app for creating, exploring, labeling and sharing volcano plots. *Sci Rep* 2020;10:20560.
33. Caspa Gokulan R, Devaraj H. Stem cell markers CXCR-4 and CD133 predict aggressive phenotype and their double positivity indicates poor prognosis of oral squamous cell carcinoma. *Cancers (Basel)* 2021;13:5895.
34. Noorbakhsh F, Rezaie F, Abdolmaleki M, Ghasemi N, Abdollahi B, Nojan F, et al. Dental students' awareness about prevention, early diagnosis and referral of patients with oral cancers in Tabriz in 2018. *Avicenna J Dent Res* 2018;10:89-94.
35. Bautista-Sanchez D, Arriaga-Canon C, Pedroza-Torres A, De La Rosa-Velazquez IA, Gonzalez-Barrios R, Contreras-Espinosa L, et al. The promising role of miR-21 as a cancer biomarker and its importance in RNA-based therapeutics. *Mol Ther Nucleic Acids* 2020;20:409-420.
36. Nguyen P, Bar-Sela G, Sun L, Bisht KS, Cui H, Kohn E, et al. BAT3 and SET1A form a complex with CTCFL/BORIS to

- modulate H3K4 histone dimethylation and gene expression. *Mol Cell Biol* 2008;28:6720-6729.
37. Zhang X, Gee H, Rose B, Lee CS, Clark J, Elliott M, et al. Regulation of the tumour suppressor PDCD4 by miR-499 and miR-21 in oropharyngeal cancers. *BMC Cancer* 2016;16:86.
 38. Zhu S, Si ML, Wu H, Mo YY. MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). *J Biol Chem* 2007;282:14328-14336.
 39. Wang XS, Prensner JR, Chen G, Cao Q, Han B, Dhanasekaran SM, et al. An integrative approach to reveal driver gene fusions from paired-end sequencing data in cancer. *Nat Biotechnol* 2009;27:1005-1011.
 40. Amirfallah A, Knutsdottir H, Arason A, Hilmarsdottir B, Johannsson OT, Agnarsson BA, et al. Hsa-miR-21-3p associates with breast cancer patient survival and targets genes in tumor suppressive pathways. *PLoS One* 2021;16:e0260327.
 41. Gao Z, Liu H, Shi Y, Yin L, Zhu Y, Liu R. Identification of cancer stem cell molecular markers and effects of hsa-miR-21-3p on stemness in esophageal squamous cell carcinoma. *Cancers (Basel)* 2019;11:518.
 42. Crimi S, Falzone L, Gattuso G, Grillo CM, Candido S, Bianchi A, et al. Droplet digital PCR analysis of liquid biopsy samples unveils the diagnostic role of hsa-miR-133a-3p and hsa-miR-375-3p in oral cancer. *Biology (Basel)* 2020;9:379.
 43. Xu X, Chen X, Xu M, Liu X, Pan B, Qin J, et al. miR-375-3p suppresses tumorigenesis and partially reverses chemoresistance by targeting YAP1 and SP1 in colorectal cancer cells. *Aging (Albany NY)* 2019;11:7357-7385.
 44. Liu T, Han X, Zheng S, Liu Q, Tuerxun A, Zhang Q, et al. CALM1 promotes progression and dampens chemosensitivity to EGFR inhibitor in esophageal squamous cell carcinoma. *Cancer Cell Int* 2021;21:121.
 45. Berchtold MW, Villalobo A. The many faces of calmodulin in cell proliferation, programmed cell death, autophagy, and cancer. *Biochim Biophys Acta* 2014;1843:398-435.
 46. Vergne I, Chua J, Deretic V. Tuberculosis toxin blocking phagosome maturation inhibits a novel Ca²⁺/calmodulin-PI3K hVPS34 cascade. *J Exp Med* 2003;198:653-659.
 47. Chin D, Means AR. Calmodulin: a prototypical calcium sensor. *Trends Cell Biol* 2000;10:322-328.
 48. Adeola HA, Smith M, Kaestner L, Blackburn JM, Zerbini LF. Novel potential serological prostate cancer biomarkers using CT100+ cancer antigen microarray platform in a multi-cultural South African cohort. *Oncotarget* 2016;7:13945-13964.
 49. Zhang L, Feng C, Zhou Y, Zhou Q. Dysregulated genes targeted by microRNAs and metabolic pathways in bladder cancer revealed by bioinformatics methods. *Oncol Lett* 2018;15:9617-9624.
 50. Zamanian Azodi M, Rezaei Tavirani M, Rezaei Tavirani M, Vafae R, Rostami-Nejad M. Nasopharyngeal carcinoma protein interaction mapping analysis via proteomic approaches. *Asian Pac J Cancer Prev* 2018;19:845-851.
 51. Adrain C, Martin SJ. The mitochondrial apoptosome: a killer unleashed by the cytochrome seas. *Trends Biochem Sci* 2001;26:390-397.
 52. Barczyk K, Kreuter M, Pryjma J, Booy EP, Maddika S, Ghavami S, et al. Serum cytochrome c indicates *in vivo* apoptosis and can serve as a prognostic marker during cancer therapy. *Int J Cancer* 2005;116:167-173.
 53. Bayat Z, Farhadi Z, Taherkhani A. Identification of potential biomarkers associated with poor prognosis in oral squamous cell carcinoma through integrated bioinformatics analysis: a pilot study. *Gene Rep* 2021;24:101243.
 54. Hayashido Y, Nakashima M, Urabe K, Yoshioka H, Yoshioka Y, Hamana T, et al. Role of stromal thrombospondin-1 in motility and proteolytic activity of oral squamous cell carcinoma cells. *Int J Mol Med* 2003;12:447-452.
 55. Alvarez AA, Axelrod JR, Whitaker RS, Isner PD, Bentley RC, Dodge RK, et al. Thrombospondin-1 expression in epithelial ovarian carcinoma: association with p53 status, tumor angiogenesis, and survival in platinum-treated patients. *Gynecol Oncol* 2001;82:273-278.
 56. Tuszyński GP, Nicosia RF. Localization of thrombospondin and its cysteine-serine-valine-threonine-cysteine-glycine-specific receptor in human breast carcinoma. *Lab Invest* 1994;70:228-233.
 57. Patey M, Delemer B, Bellon G, Martiny L, Pluot M, Haye B. Immunohistochemical study of thrombospondin and its receptors alpha root of beta 3 and CD36 in normal thyroid and in thyroid tumours. *J Clin Pathol* 1999;52:895-900.
 58. Chen CR, Kang Y, Massague J. Defective repression of c-myc in breast cancer cells: a loss at the core of the transforming growth factor beta growth arrest program. *Proc Natl Acad Sci U S A* 2001;98:992-999.
 59. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 2007;316:1039-1043.
 60. Peng SY, Lai PL, Hsu HC. Amplification of the c-myc gene in human hepatocellular carcinoma: biologic significance. *J Formos Med Assoc* 1993;92:866-870.
 61. Pai R, Pai S, Lalitha R, Kumaraswamy S, Lalitha N, Johnston R, et al. Over-expression of c-Myc oncoprotein in oral squamous cell

- carcinoma in the South Indian population. *Ecancermedical-science* 2009;3:128.
62. Chen YJ, Lin SC, Kao T, Chang CS, Hong PS, Shieh TM, et al. Genome-wide profiling of oral squamous cell carcinoma. *J Pathol* 2004;204:326-332.
63. Tremblay M, Sanchez-Ferras O, Bouchard M. GATA transcription factors in development and disease. *Development* 2018;145:dev164384.
64. Martinelli P, Carrillo-de Santa Pau E, Cox T, Sainz B Jr, Dusetti N, Greenhalf W, et al. GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. *Gut* 2017;66:1665-1676.
65. Tsuji S, Kawasaki Y, Furukawa S, Taniue K, Hayashi T, Okuno M, et al. The miR-363-GATA6-Lgr5 pathway is critical for colorectal tumorigenesis. *Nat Commun* 2014;5:3150.
66. Sulahian R, Casey F, Shen J, Qian ZR, Shin H, Ogino S, et al. An integrative analysis reveals functional targets of GATA6 transcriptional regulation in gastric cancer. *Oncogene* 2014;33:5637-5648.
67. Deng X, Jiang P, Chen J, Li J, Li D, He Y, et al. GATA6 promotes epithelial-mesenchymal transition and metastasis through MUC1/beta-catenin pathway in cholangiocarcinoma. *Cell Death Dis* 2020;11:860.
68. Deng L, Liu H. MicroRNA-506 suppresses growth and metastasis of oral squamous cell carcinoma via targeting GATA6. *Int J Clin Exp Med* 2015;8:1862-1870.
69. King JA, Corcoran NM, D'Abaco GM, Straffon AF, Smith CT, Poon CL, et al. Eve-3: a liver enriched suppressor of Ras/MAPK signaling. *J Hepatol* 2006;44:758-767.
70. Mizoguchi M, Nutt CL, Louis DN. Mutation analysis of CBL-C and SPRED3 on 19q in human glioblastoma. *Neurogenetics* 2004;5:81-82.
71. He Z, Gong F, Liao J, Wang Q, Su Y, Chen C, et al. Spred-3 mutation and Ras/Raf/MAPK activation confer acquired resistance to EGFR tyrosine kinase inhibitor in an EGFR mutated NSCLC cell line. *Transl Cancer Res* 2020;9:2542-2555.
72. Dituri F, Gigante G, Scialpi R, Mancarella S, Fabregat I, Giannelli G. Proteoglycans in cancer: friends or enemies? A special focus on hepatocellular carcinoma. *Cancers (Basel)* 2022;14:1902.
73. Espinoza-Sanchez NA, Gotte M. Role of cell surface proteoglycans in cancer immunotherapy. *Semin Cancer Biol* 2020;62:48-67.
74. Ahrens TD, Bang-Christensen SR, Jorgensen AM, Loppke C, Spliid CB, Sand NT, et al. The role of proteoglycans in cancer metastasis and circulating tumor cell analysis. *Front Cell Dev Biol* 2020;8:749.
75. Nagarajan A, Malvi P, Wajapeyee N. Heparan sulfate and heparan sulfate proteoglycans in cancer initiation and progression. *Front Endocrinol (Lausanne)* 2018;9:483.
76. Garusi E, Rossi S, Perris R. Antithetic roles of proteoglycans in cancer. *Cell Mol Life Sci* 2012;69:553-579.