Research in Pharmaceutical Sciences, June 2024; 19(3): 287-302
Received: 22-08-2023
Peer Reviewed: 20-09-2023
Revised: 14-10-2023
Accepted: 13-11-2023
Published: 01-07-2024

School of Pharmacy & Pharmaceutical Sciences
Isfahan University of Medical Sciences

Original Article

A bioinformatics approach of specificity protein transcription factors in head and neck squamous cell carcinoma

Adel Rezvani Sichani[1], Ziba Rezvani Sichani[2], Behnaz Yazdani[3], Mehdi Azizmohammad Looha[4], and Hajar Sirous[5,*]

[1]Department of Food Science and Technology, Shahreza Branch, Islamic Azad University, Shahreza. I.R. Iran.
[2]Department of Biochemistry, Islamic Azad University, Falavarjan Branch, I.R. Iran.
[3]Bioscience Department, Faculty of Science and Technology (FCT), Universitat de Vic—Universitat Central de Catalunya (Uvic-UCC), 08500 Vic, Spain.
[4]Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
[5]Bioinformatics Research Center, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, I.R. Iran.

Abstract

Background and purpose: The seventh most common type of cancer with increasing diagnosis rates around the world is head and neck squamous cell carcinoma (HNSCC). Specificity proteins (SPs) have been known for their role in the regulation of cellular division, growth, and apoptotic pathways in various cancers. In this work, we analyzed the expression levels of SPs in HNSCC to assess their diagnostic and prognostic biomarker potential.
Experimental approach: Differential gene expression and correlation analysis methods were used to determine the top dysregulated genes in HNSCC. Functional enrichment and protein-protein interaction analyses were done with the DAVID database and Cytoscape software to understand their function and biological processes. Receiver operating test, logistic regression, and Cox regression analyses were performed to check SP genes' diagnostic and prognostic potential.
Findings/Results: SP1 (LogFC = -0.27, P = 0.0013) and SP2 (LogFC = -0.20, P = 0.0019) genes were upregulated in HNSCC samples, while SP8 (LogFC = 2.57, $P < 0.001$) and SP9 (LogFC = 2.57, $P < 0.001$) genes were downregulated in cancer samples. A moderate positive correlation was observed among the expression levels of SP1, SP2, and SP3 genes. The SP8 and SP9 genes with AUC values of 0.79 and 0.75 demonstrated diagnostic potential which increased to 0.84 when both genes were assessed by logistic regression test. Also, the SP1 gene held a marginally significant prognostic potential.
Conclusion and implications: Our findings clarify the potential of SP transcription factors as candidate diagnostic and prognostic biomarkers for early screening and treatment of HNSCC.

Keywords: Head and neck squamous cell carcinoma; Specificity protein; SP1; TCGA.

INTRODUCTION

Head and neck cancers are a type of malignancy the cellular origin of most of them lies in the mucosal epithelium layer of the larynx, pharynx, and oral cavity. Around 900,000 new cases are diagnosed each year with this type of cancer, but a great percentage of them face short survival periods (1,2). The high diagnosis rates of head and neck cancers have made this cancer known as the seventh most common form of cancer worldwide (1). The incidence of head and neck cancers has been predicted to elevate up to 30% each year by 2030 (3). There is an increasing need for the identification of biomarkers in the prediction of head and neck squamous cell carcinoma (HNSCC) that could help with better diagnosis and prediction of survival period in patients.

*Corresponding author: H. Sirous
Tel: +98-3137927065, Fax: +98-3136680011
Email: h_sirous@pharm.mui.ac.ir

Access this article online	
	Website: http://rps.mui.ac.ir
	DOI: 10.4103/RPS.RPS_171_23

The genes coding specificity proteins (SP) are transcription factors (TFs) holding a common zinc-binding domain that aids with DNA-binding and the regulation of genes that participate in specific cellular pathways including cell cycle regulation and cellular differentiation (4). SP transcription factors consist of several members and their altered expression level has been associated with the regulator pathways involved in varying biological processes of cancer cells (5). SP TFs have a high binding affinity for GC/GT-rich sequences within the promoter regions of genes. Multiple studies have linked multiple correlations between cellular growth and the metastatic potential of SP TFs in a variety of cancer cells, while their function and expression profile in HNSCC are poorly investigated.

SP factors have been noticed to play a notable role in the regulation of cellular division, growth, and induction of anti-apoptotic signals (6). However, it appears that the tumor suppressor or oncogenic activities of these factors vary across different types of cancer types (7). Most of the previous investigations on SPs have been focused on the molecular functions of SP1 and SP3 genes more than the other members of the SP transcription family. This gap in knowledge about SPs could be due to the higher DNA-binding affinity of these factors with GC-rich promoter regions (8). The regulatory function of SPs in expression levels of pro-apoptotic or anti-apoptotic genes has made these factors an ideal target for the design of new therapeutic opportunities for the treatment of cancers (9). Due to the lack of enough information about the molecular function and expression patterns of SPs in HNSCC, more investigations should be carried out before using these genes as biomarkers in clinical trials.

In the current research, we aimed to investigate the expression patterns of 9 different members of SP TFs in the expression matrix of HNSCC with the use of comprehensive bioinformatic tools. We suggested ideal biomarkers with notable prognostic and diagnostic capability that could serve better approaches in early screening and prediction of survival period with patients with different expression profiles of SPs in HNSCC.

METHODS AND MATERIAL

Data processing and gene expression analysis

The Cancer Genome Atlas (TCGA) program is a highly practical online dataset (Available at the Cancer Genome Atlas Program (TCGA) - NCI) that allows free access usage of RNA-seq count data of more than 13 different types of cancers with normal adjacent tissue samples under the principles organized in the declaration of Helsinki statements. The genome expression matrix of 502 HNSC cancer samples along with 44 normal tissue samples in the format of count data was downloaded and normalized with the help of TCGAbiolinks, Limma, and edgeR packages (Fig. 1).

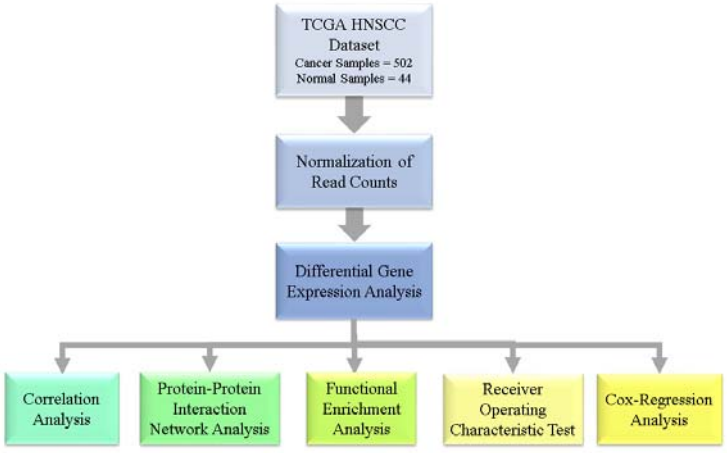

Fig. 1. The flowchart of the study.

The clinical information of the patients has been summarized in Table S1. The count data was converted into a log2 ratio after normalization by the Voom package. The differentially expressed genes (DEGs) were calculated and reordered based on adjusted *P*-values and the top 100 DEGs with the smallest adjusted *P*-values were selected genes in HNSCC samples concerning normal tissue samples for further analysis and *P*-values smaller or equal to 0.01 were considered as statistically significant (10-14).

Correlation analysis

To better understand the interaction and molecular relationships between the SP TFs and the top DEGs in HNSCC, correlation analysis can be used as a useful method to explore this relationship. Accordingly, correlation analysis was performed using the normalized expression data of the SP TFs and top 10 DEGs in HNSCC samples. The metan package in R programming was used for this analysis and the Pearson statistical method was selected for the estimation of *P*-values and correlation coefficient values.

Functional enrichment analysis

To better understand the important biological pathways that are involved in the progression and development of HNSCC, functional enrichment analysis was performed with the help of the DAVID database (version 6.8, available at https://david-d.ncifcrf.gov/), which is an online platform that gives free access opportunities to practical functions and algorithms that can estimate the enrichment of genes in different biological pathways and predict their cellular function and localization. One of the practical analyses that can be done using the DAVID database, is the gene ontology (GO) analysis that was used for the top 200 DEGs in HNSCC to achieve a better perspective of their functions and associated biological pathways (15-17).

Protein-protein interaction analysis

Protein-protein interaction (PPI) network analysis is a practical method that aids with the understanding of the interaction and interplay of a large number of genes that are significantly dysregulated and the function of a majority of them is still unexplored. For this analysis, the STRING online platform (version 10, available at http://www.string.db.org) (18,19) was used from Cytoscape software (version: 3.2.0, available at http://www.cytoscape.org/) (20,21). The PPI network was constructed in Cytoscape using the list of SP TFs with the top 100 DEGs in HNSCC and was analyzed with the CytoNCA tool (available at https://apps.cytoscape.org/apps/cytonca) (22) in Cytoscape software.

Receiver operating characteristic test

A practical statistical method for estimation of the diagnostic potential of genes based on their expression levels in two defined phenotypes of interest is the receiver operating characteristic (ROC) test which can be easily accessed using the GraphPad Prism software (version 9.1.0). Through this analysis, ROC plots are generated based on the sensitivity and specificity of the data, respectively (23). With the help of this method, the diagnostic potential of SP transcription factors in HNSC and normal groups was calculated and ROC plots were generated. To test the potential of the combination of SP genes with moderate AUC values in the diagnosis of HNSCC, the logistic regression statistical test was applied using the R software (version 4.3.1) and *P*-values ≤ 0.05 were considered statistically significant.

Cox-regression analysis

OncoLnc is a highly practical online database (available at http://www.oncolnc.org/) (24) that utilizes specific statistical methods for the estimation of the survival period based on the expression data from the TCGA database. To explore the association between the expression levels of SP genes in HNSCC with the time of survival in patients, the OncoLnc database was used which utilizes the Cox-regression statistical method for survival analysis concerning the clinical data of the patients. It also uses the samples with expression values according to the upper quartile or lower quartile criteria and reports Logrank *P*-values as well, which is a hypothesis test that compares the difference in the survival distributions and Logrank *P*-values smaller than 0.01 would be considered statistically significant (25).

RESULTS

Expression analysis of SP TFs in HNSCC

The gene expression analysis of HNSC cancer samples about normal samples was performed and the top 100 genes with the most statistically significant adjusted *P*-values were identified and reported in Table S2. Differential gene expression analysis of SP TFs in HNSC cancer demonstrated an uneven pattern between the expression levels of these genes, indicating that the expression level of each SP factor contributes differently to the progression of HNSCC. As demonstrated in Fig. 2, among all 9 members of the SP transcription family, SP1 ($P = 0.002$) and SP2 ($P = 0.01$) showed notable decreased expression levels in HNSCC tissue samples, while SP8 ($P < 0.0001$) and SP9 ($P < 0.0001$) genes revealed very significant increased expression levels in HNSCC samples and their high expression ratio might be associated with the biological pathways in HNSCC. The expression levels of SP3 ($P = 0.02$), and SP7 ($P = 0.03$) genes were also notably higher in cancer samples, while the expression levels of SP4 ($P = 0.98$), SP5 ($P = 0.07$), and SP6 ($P = 0.21$) were not statistically significantly different between HNSCC cancer and normal samples.

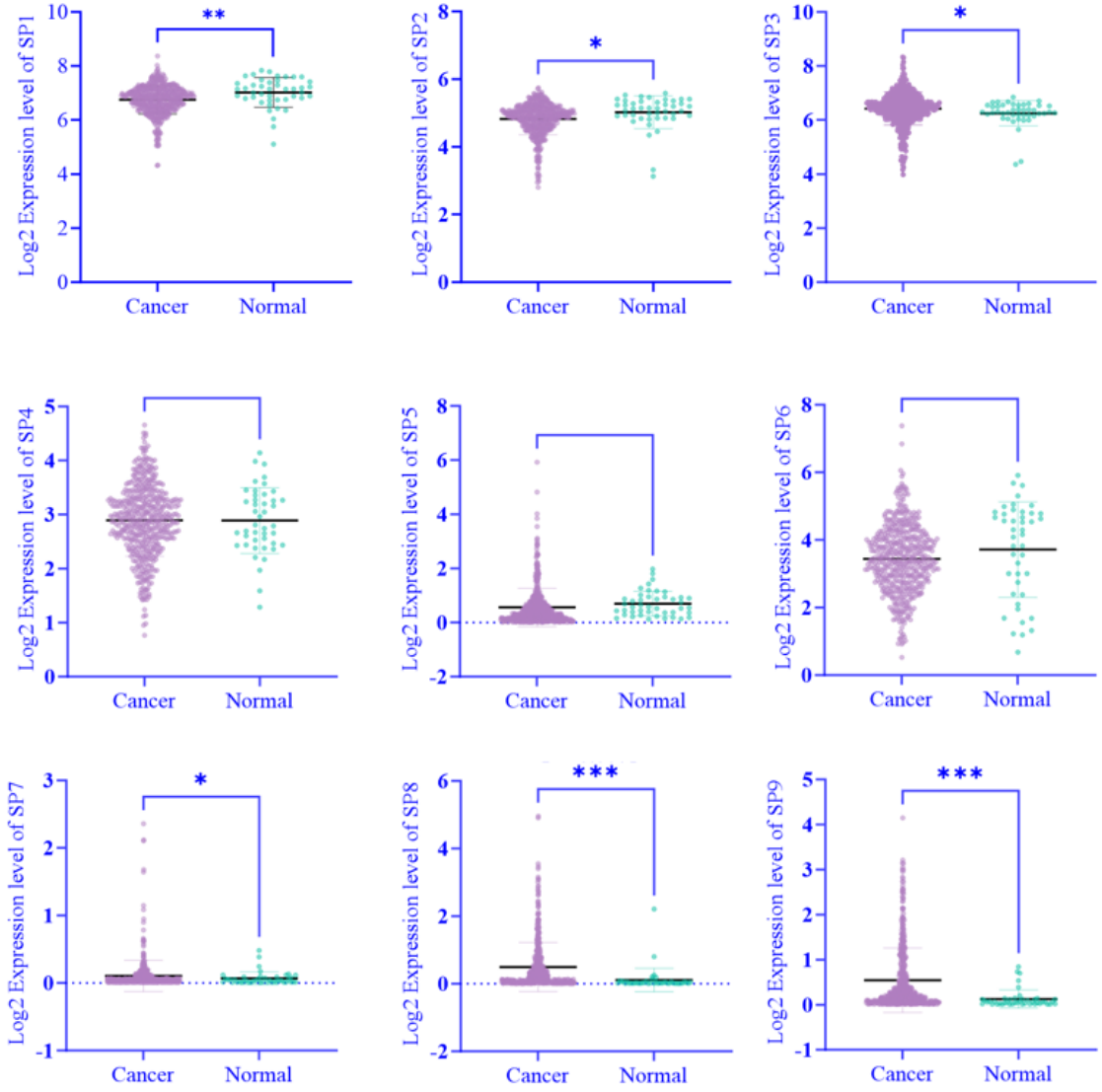

Fig. 2. Differential gene expression analysis of SP transcription factors in HNSCC. The RNA-seq count data of TCGA HNSC cancer and normal tissue samples were normalized and analyzed to identify the top differentially expressed genes. *$P \leq 0.05$, **$P \leq 0.01$, and ***$P \leq 0.001$ indicate statistically significant differences between groups. SP, Specificity protein; HNSCC, head and neck squamous cell carcinoma.

Correlation analysis between top DEGs and expression levels of SP genes

A practical method for the prediction of possible interactions at a molecular level is correlation analysis. This technique was applied using the normalized expression data of SP genes and the top 10 DEGs in HNSCC. As shown in Fig. 3, there was a moderate positive correlation between SP1 with SP3 (correlation coefficient = 0.63, $P \leq 0.001$), and SP1 with SP2 (correlation coefficient = 0.66, $P \leq 0.001$) gene. From the top 10 DEGs in HNSCC, the expression level of the ADIPOQ gene also had a significant positive correlation with the PLIN1 gene (correlation coefficient = 0.94, $P \leq 0.001$).

GO analysis of top DEGs

As depicted in Fig.4, most of the genes were predicted by the GO tool of the DAVID database to be involved in cell division (GO: 00051301) and mitotic cell cycle (GO: 0000278). The molecular function of a majority of the DEGs in HNSCC was associated with protein binding (GO: 0005515), DNA binding (GO: 0003677), and ATP binding (GO: 0005524). The majority of these genes were also estimated to localize in the cytosol (GO: 0005829) and the nucleus (GO: 0005634). The KEGG database also predicted that a large count of the DEGs were involved in the cell cycle (hsa04110) and peroxisome proliferator-activated receptors signaling pathway (hsa03320).

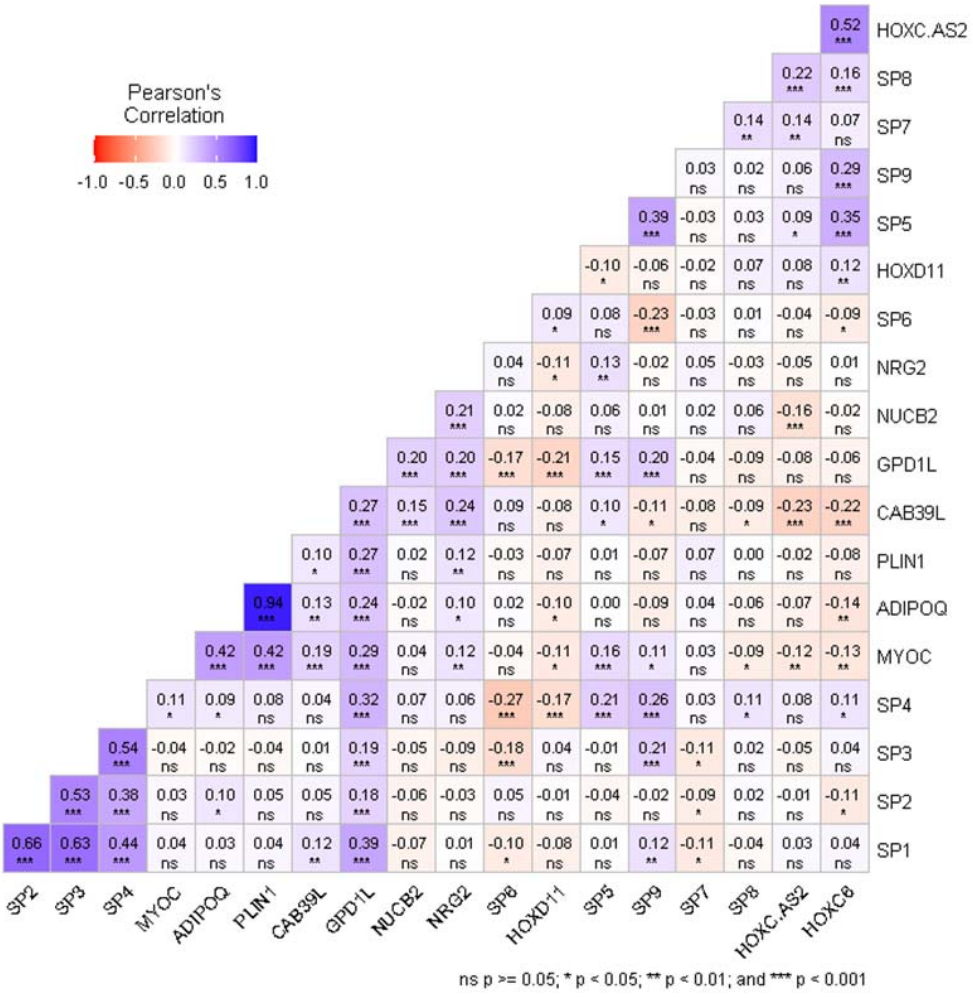

Fig. 3. Correlation analysis between SP transcription factors and top DEGs in HNSCC. The correlation analysis was performed using the Pearson method between the normalized expression levels of SP genes and the top 10 DEGs in the HNSCC. SP, Specificity protein; HNSCC, head and neck squamous cell carcinoma; DEGs, differentially expressed genes.

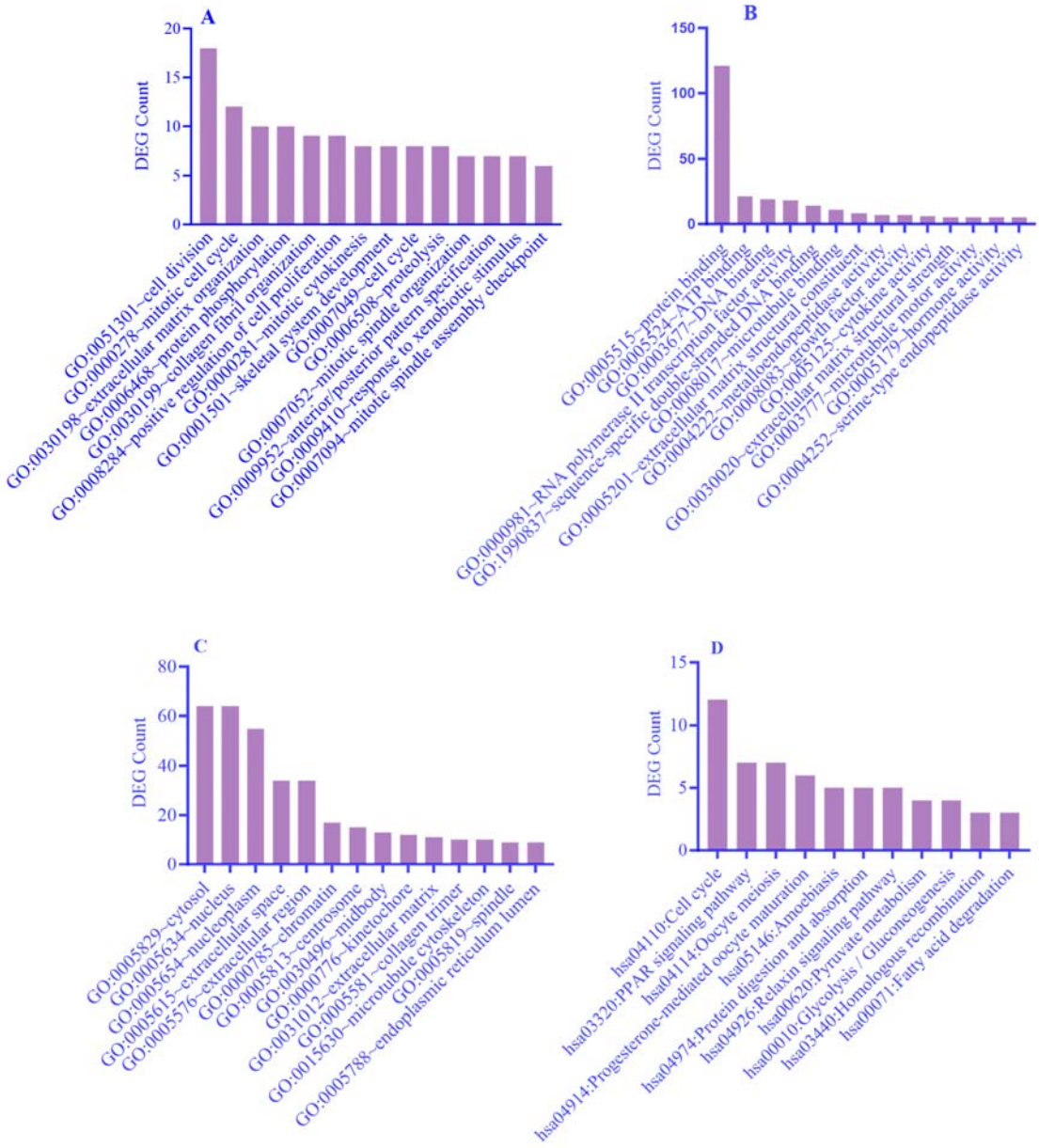

Fig. 4. GO and KEGG pathway analysis of top DEGs in HNSCC. (A) Functional enrichment analysis of top DEGs in the HNSC cancer showed that most of the dysregulated genes were involved in biological pathways such as cell division; (B) molecular functions of most DEGs were associated with protein binding and ATP binding; (C) most of the DEGs were predicted to be localized mostly in the cytosol or nucleus regions; (D) the KEGG pathway also predicted that most of the genes were involved in the regulation of cellular division. GO, Gene Ontology; DEG, differentially expressed gene; HNSCC, head and neck squamous cell carcinoma.

PPI network analysis

HNSCC is a poorly investigated cancer in which the function of the majority of the high DEGs is still under investigation. Therefore, to gain a better insight into the biological processes and molecular interplay between the genes in HNSCC, a PPI network was constructed in Cytoscape and the hub genes in the network were employed with the CytoNCA tool that allows easier measurement of cluster coefficient and node degree in a set of gene list converted into an interactive network. The threshold for the PPI score was set to 0.4, which resulted in the calculation of the

medium-confidence network generated in Cytoscape software. The CytoNCA tool can help with the estimation of the topological parameters of the PPI network and its calculations were performed by excluding the weight, so it can better predict and identify the highly interacting proteins in the network.

The interconnectivity between the DEGs in HNSCC has been shown in Fig. 5, in which each gene in the network has been grouped with other genes that are involved in similar biological processes that were previously predicted using the DAVID database. As it can be understood from the network, intense connectivity exists between the gene sets from the mitotic cell cycle, cell division, centrosome complex assembly, microtubule-based movement, and regulation of cell cycle pathways. The SP1, SP2, and SP7 genes were predicted to interact with genes from the cell division pathway and the SP1 gene also interacted with genes in the regulation of the cell cycle pathway. The top genes with degree scores above 46 from the network topology analysis results by the CytoNCA tool have been summarized in Table 1, in which between the top 19 genes with highest degree, subgraph, betweenness, and closeness scores, the Forkhead box M1, aurora A kinase, RAD51 recombinase, MYB proto-oncogene like 2, baculoviral IAP repeat containing 5, and kinesin family member 14 genes had the highest node degree scores within the PPI network. The Forkhead box M1, aurora A kinase, RAD51 recombinase, and kinesin family member 14 genes had the highest betweenness and closeness scores, which indicates the quicker reach and higher control of these genes to other nodes in the PPI network.

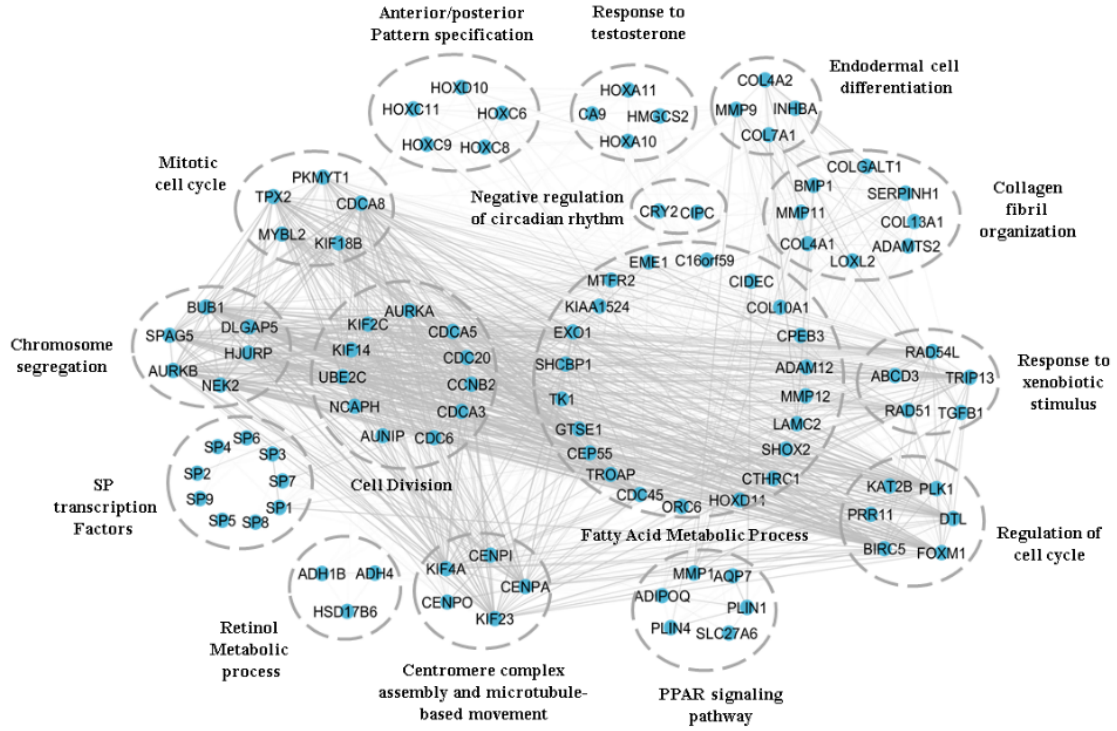

Fig. 5. Protein-protein interaction network of top DEGs in head and neck squamous cell carcinoma. Network analysis of the top DEGs with the help of the STRING database and CytoNCA plugin in the Cytoscape software revealed a notable connection among important cellular pathways that most of the genes were predicted previously by GO analysis to be enriched and involved in. SP factors also showed to interact with other proteins and pathways shown in the network. GO, Gene Ontology; DEG, differentially expressed gene; SP, specificity protein.

Table 1. Top 19 genes from the top 200 differentially expressed genes in head and neck squamous cell carcinoma cancer with the highest degrees in protein-protein interaction network analyzed by CYTOCNA application in Cytoscape.

Ensemble protein ID	Gene ID	Description	Degree	Betweenness	Closeness
ENSP00000342307	FOXM1	Forkhead box M1	53	840.25543	0.2820513
ENSP00000216911	AURKA	Aurora kinase A	51	403.89383	0.27576602
ENSP00000372088	RAD51	RAD51 recombinase	51	420.04904	0.27423823
ENSP00000217026	MYBL2	MYB proto-oncogene like 2	51	270.12604	0.27348065
ENSP00000301633	BIRC5	Baculoviral IAP repeat containing 5	50	362.1902	0.27348065
ENSP00000356319	KIF14	Kinesin family member 14	50	224.12881	0.27272728
ENSP00000300403	TPX2	TPX2 microtubule nucleation factor	49	258.1087	0.27272728
ENSP00000336868	CENPA	Centromere protein A	48	132.00188	0.27123287
ENSP00000405726	CDC45	Cell division cycle 45	48	90.10279	0.26902175
ENSP00000300093	PLK1	Polo-like kinase 1	48	60.806026	0.2682927
ENSP00000361540	CDC20	Cell division cycle 20	47	200.96227	0.26756757
ENSP00000313950	AURKB	Aurora kinase B	47	63.645523	0.26612905
ENSP00000302530	BUB1	BUB1 mitotic checkpoint	47	66.95278	0.26190478
ENSP00000363524	KIF4A	Kinesin family member 4A	47	15.906779	0.26121372
ENSP00000362146	CDCA8	Cell division cycle associated 8	47	15.906779	0.26121372
ENSP00000260363	KIF23	Kinesin family member 23	47	15.906779	0.26121372
ENSP00000275517	CDCA5	Cell division cycle associated 5	47	70.41854	0.26121372
ENSP00000355506	EXO1	Exonuclease 1	47	49.754776	0.26121372
ENSP00000288207	CCNB2	Cyclin B2	47	110.61793	0.26052633

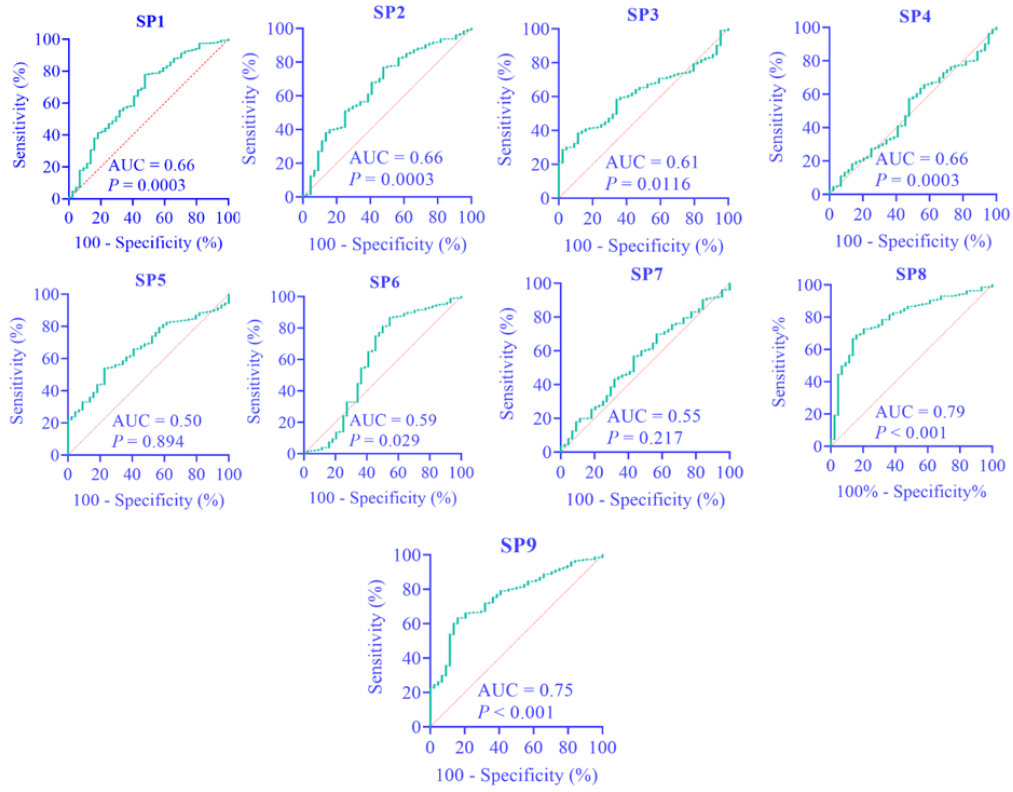

Fig. 6. Receiver operating characteristic test of SP transcription factors in HNSCC. The receiver operating characteristic test was used to assess and clarify the potential of SP transcription factors as diagnostic biomarkers in HNSCC and the results revealed that only SP8 and SP9 genes could hold a diagnostic potential for HNSCC. Logistic regression analysis was performed for the prediction of the diagnostic power of the combination of SP8 and SP9 genes in HNSCC. The AUC values below 0.70 were considered weak diagnostic capability and $P \leq 0.01$ are considered statistically significant. SP, Specificity protein; HNSCC, head and neck squamous cell carcinoma; AUC, the area under the curve.

Diagnostic potential of SP genes in HNSCC

The diagnostic capability of SP transcription factors was assessed using GraphPad Prism software based on the normalized expression matrix of HNSCC and normal samples. As demonstrated in Fig. 6, two genes revealed significant diagnostic potential including the SP8 and SP9 genes compared to the rest of the members of the SP TF family, whose AUC values did not meet the satisfactory statistical criteria. While the diagnostic potential of SP genes in HNSCC was not statistically significant, SP8 and SP9 genes demonstrated better diagnostic capability compared to the other SP genes. The analysis of the logistic regression test for estimation of diagnostic capability for the combination of SP8 and SP9 genes predicted an AUC value of 0.84, which indicates better diagnostic potential of these two genes together in the HNSCC.

Analysis of prognostic biomarker capability of SP genes

Based on the Cox-regression analysis results presented in Fig. 7, only SP1 (Logrank $P = 0.05$) and SP5 (Logrank $P = 0.01$) genes demonstrated marginally significant statistical prognostic potential compared to the rest of the SP members, while the prognostic potential of other SP genes in HNSCC was statistically poor and not significant. Also, it can be seen that SP1 and SP5 expression levels correlated with better survival periods in HNSCC patients. It should be mentioned that the OncoLnc database did not provide any Cox-regression results over the SP7 gene as the expression level of SP7 did not meet the expression cutoff demanded for this analysis. Overall, SP1 and SP5 transcription factors revealed better prognostic potential compared to other SP genes in HNSCC.

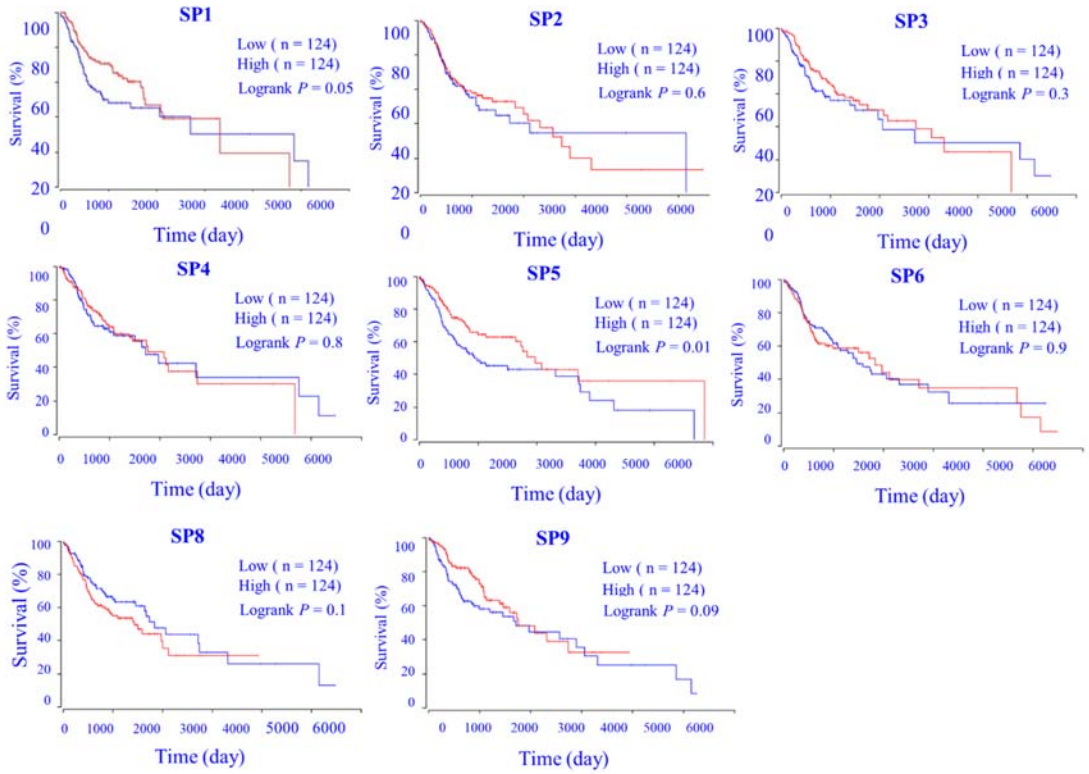

Fig. 7. Cox-regression analysis of SP transcription factors in HNSCC. Survival analysis was performed with the help of the Cox-regression analysis tool provided by the OncoLnc database, which utilized the mRNA expression data from the TCGA database along with the clinical data of the patients. The OncoLnc database calculates the Cox-regression analysis for interest genes using the samples with expression values above the upper quartile ('high' group) and lower quartile ('low' group). This test was done to check the prognostic potential of SP transcription factors in HNSCC, and only SP1 and SP5 genes revealed better prognostic potential compared to other members of the SP family. However, the prognostic values of SP genes were statistically not significant in the HNSCC. The last graph shows the result of the Logistic regression test performed on the combination of SP8 and SP9 genes for the prediction of their diagnostic potential in HNSCC. The Logrank $P \leq 0.01$ was considered statistically significant. SP, Specificity protein; HNSCC, head and neck squamous cell carcinoma; AUC, the area under the curve.

DISCUSSION

One of the common malignancies that affect the head and neck regions of the human body is the HNSCC. The unstable genomic nature and the high metastatic potential of this cancer have increased the need for further investigations on more specific molecular therapies and the identification of better biomarkers for early detection and prevention in HNSCC patients (9,26).

TFs are important and attractive molecular targets for the design of specific therapeutic approaches for the treatment of different types of cancer in humans (27,28). Previous studies that analyzed the expression spectrum of HNSCC identified varying common mutations in the genomic regions of proteins in about 30% of the patients with HNSCC. Specific genes such as tumor protein p63, cyclin-dependent kinase inhibitor A2, and phosphatase and tensin homolog genes were found to be mutated in HNSCC, but their tumor suppressor or oncogenic activities were reported to vary among different types of cancers (29,30).

SPs are a class of TFs with DNA-binding activity and multiple cellular functions that help with the maintenance of cellular homeostasis such as regulation of cellular division, apoptotic pathway, and metastasis that have been reported previously by multiple investigations on different cancers. Multiple members of SP TFs such as SP1, SP3, and SP4 genes have been reported to regulate important pathways that involve cellular growth, division, survival, and inflammatory pathways of different types of cancer cells. The SP1 gene has been reported as a pro-oncogenic factor due to its role in the regulation of survival and metastasis of cancer cells and has been suggested as a molecular target for the design of new anticancer drugs and chemotherapies(31,32).

In this study, we performed genome expression analysis using RNA-seq count data of HNSCC samples from the TCGA database and clarified the expression pattern of 9 members of the SP TF family. The Differential gene expression analysis revealed a notable decrease in the expression levels of SP1 and SP2 genes along with a significant increase in the expression levels of SP8 and SP9 genes in HNSCC samples. This indicates that each SP gene has a unique role and expression pattern in the HNSCC. Previous studies have investigated on SP1 gene far more than the other SP members in multiple types of human cancer cell lines, such as pancreatic ductal adenocarcinoma, colorectal cancer, and glioma cancer (33-36).

Differential gene expression and correlation analysis results of the HNSCC expression matrix revealed that there is a moderate positive correlation among the expression levels of SP1, SP2, and SP3 genes. This indicates that a co-expression regulation mechanism might exist among these SP genes. There was also a notable positive correlation between the ADIPOQ gene with PLIN1 and SP4 genes. While the expression levels and biological roles of the ADIPOQ gene in HNSCC are poorly investigated, a study has reported an association between different genetic variants of this gene and with risk of breast cancer (37). Other studies also found an association between the human papillomavirus (HPV) with the risk of oropharyngeal squamous cancer cells and HNSCC (38-40).

To achieve a better vision of the common biological pathways in HNSCC, functional enrichment analysis was done and most of the DEGs were shown to participate in major biological pathways in cancer, such as cell division and mitotic cell cycle pathways. Also, their molecular functions were associated with DNA-binding, ATP-binding, and protein-binding activities and most of them were found to localize in the cytosol and nucleus regions of the cells. PPI network analysis revealed extensive connectivity between the genes that are predicted to be involved in the regulation of the cell cycle and the assembly of the chromosome complex. Genes such as SP1, SP2, and SP3 were shown to interact with pathways associated with cellular growth and division as well.

To clarify the diagnostic and prognostic capabilities of SP TFs in HNSCC, ROC test and Cox-regression survival analysis were used and we found that most of the SP genes have very weak diagnostic potential in HNSCC except for SP8 and SP9 genes, which demonstrated better sensitivity in the detection of cancer phenotype

from normal tissue samples. Also, the combination of SP8 and SP9 genes revealed a better diagnostic capability in the detection of HNSCC. The Cox-regression analysis also showed that only the SP1 gene showed a marginally significant potential for the prognosis of HNSCC and the estimation of survival period in patients with HNSCC. Previous studies have also noted that the SP1 gene can be a great prognostic biomarker in human gastric cancer (40-43).

Currently, there are no trials that have experimentally investigated the molecular role and expression patterns of all 9 SP genes in HNSCC, and our study has clarified their expression pattern along with their prognostic and diagnostic capabilities for the first time. A recent study reviewed the role of SP1, SP3, and SP4 genes in cancer, while the data on the role of other members of the SP TF family in HNSCC is poorly understood (44). Other studies had also used microarray and RNAseq data of head and neck cancer samples to investigate the expression dysregulation of TFs or microRNA-mRNA interactions of high DEGs in head and neck cancer samples, while in this study we analyzed the expression dysregulation and biomarker capability of 9 different SP TFs in the HNSC samples as their importance in the HNSC was still poorly understood (45-47). Some drugs were reported to be capable of decreasing the expression levels of the SP1 gene in cancer cells indirectly (32,43), such as anti-inflammatory and chemo-preventive agents (32). Further investigations are demanded to clarify the molecular functions and importance of the SP TF expression levels in the progression of HNSCC. These genes can have the potential to be used and investigated as diagnostic and prognostic biomarkers for the development of advanced and specific targeted therapies in HNSC cancer.

CONCLUSION

HNSCC is a type of cancer affecting the regions in the head and neck area and its origin site can be from the epithelium cells located in the oral cavity and pharynx regions. The increasing rates of HNSCC around the world have increased the need for the identification of practical biomarkers for early prediction of HNSCC. In this work, we analyzed the expression matrix of HNSCC tissue samples with adjacent control tissues and we found a significant dysregulation in the expression levels of SP1, SP2, SP8, and SP9 genes from the 9 members of specificity proteins. The correlation analysis results also revealed a positive correlation among the expression levels of SP1, SP2, and SP3 genes, which indicates the presence of a co-expression regulation interaction among these SP genes. A notable correlation was also estimated among SP4, ADIPOQ, and PLIN1 genes in HNSCC samples as well. We suggested that SP8 and SP9 genes can be potential diagnostic biomarkers while the SP1 gene can serve as prognostic biomarkers in faster detection of HNSCC and the survival period in patients with different expression levels of SP genes. Therefore, further experimental investigations are highly needed to examine and validate the real diagnostic and prognostic potential of the SP genes on HNSCC tissue samples.

Acknowledgments

The authors would like to thank the Bioinformatics Research Center at Isfahan University of Medical Sciences, Isfahan, I.R. Iran.

Conflict of interest statement

All authors declared no conflict of interest in this study.

Authors' contribution

The study design was performed by A. Rezvani Sichani, Z. Rezvani Sichani, B. Yazdani, and H. Sirous; data analysis was done by B. Yazdani and M.A. Looha; interpretations of the data and bioinformatics analysis were performed by B. Yazdani, A. Rezvani Sichani, Z. Rezvani Sichani, M.A. Looha, and H. Sirous; the manuscript was written by B. Yazdani, A. Rezvani Sichani, Z. Rezvani Sichani, M.A. Looha, and H. Sirous. The finalized article was read and approved by all authors.

REFERENCES

1. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. Nat Rev Dis Primers. 2020;6(1): 92,1-49. DOI: 10.1038/s41572-020-00224-3.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al*. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-249. DOI: 10.3322/caac.21660.
3. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. Br Dent J. 2022;233(9):780-786. DOI: 10.1038/s41415-022-5166-x.
4. Safe S, Abdelrahim M. SP transcription factor family and its role in cancer. Eur J Cancer. 2005;41(16):2438-2448. DOI: 10.1016/j.ejca.2005.08.006.
5. Bouwman P, Philipsen S. Regulation of the activity of Sp1-related transcription factors. Mol Cell Endocrinol. 2002;195(1-2):27-38. DOI: 10.1016/s0303-7207(02)00221-6.
6. Black AR, Black JD, Azizkhan-Clifford J. Sp1 and krüppel-like factor family of transcription factors in cell growth regulation and cancer. J Cell Physiol. 2001;188(2):143-160. DOI: 10.1002/jcp.1111.
7. Beishline K, Azizkhan-Clifford J. Sp1 and the 'hallmarks of cancer'. FEBS J. 2015;282(2):224-258. DOI: 10.1111/febs.13148.
8. Liu Y, Zhong X, Li W, Brattain MG, Banerji SS. The role of Sp1 in the differential expression of transforming growth factor-β receptor type II in human breast adenocarcinoma MCF-7 cells. J Biol Chem. 2000;275(16):12231-12236. DOI: 10.1074/jbc.275.16.12231.
9. Marur S, Forastiere AA. Head and neck squamous cell carcinoma: update on epidemiology, diagnosis, and treatment. Mayo Clin Proc. 2016;91(3): 386-396. DOI: 10.1016/j.mayocp.2015.12.017.
10. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, *et al*. TCGAbiolinks: an R/bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44(8):e71,1-11. DOI: 10.1093/nar/gkv1507.
11. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-140. DOI: 10.1093/bioinformatics/btp616.
12. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al*. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545-15550. DOI: 10.1073/pnas.0506580102.
13. Yazdani B, Sirous H. Expression analysis of HIF-3α as a potent prognostic biomarker in various types of human cancers: a case of meta-analysis. Res Pharm Sci. 2022;17(5):508-526. DOI: 10.4103/1735-5362.355210.
14. Parsazad E, Esrafili F, Yazdani B, Ghafarzadeh S, Razmavar N, Sirous H. Integrative bioinformatics analysis of ACS enzymes as candidate prognostic and diagnostic biomarkers in colon adenocarcinoma. Res Pharm Sci. 2023;18(4):413-429. DOI: 10.4103/1735-5362.378088.
15. Harris MA, Clark J, Ireland A, Lomax L, Ashburner M, Foulger R, *et al*. The gene ontology (GO) database and informatics resource. Nucleic Acids Res. 2004;32:D258-D261. DOI: 10.1093/nar/gkh036.
16. Touzet H, Perriquet O. CARNAC: folding families of related RNAs. Nucleic Acids Res. 2004;32(suppl_2):W142-W145. DOI: 10.1093/nar/gkh415.
17. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, *et al*. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 2007;35:W169-W175. DOI: 10.1093/nar/gkm415.
18. Mering Cv, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. Nucleic Acids Res. 2003;31(1):258-261. DOI: 10.1093/nar/gkg034.
19. 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(D1):D605-D612. DOI: 10.1093/nar/gkaa1074.
20. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-2504. DOI: 10.1101/gr.1239303.
21. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, *et al*. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc. 2007;2(10):2366-2382. DOI: 10.1038/nprot.2007.324.
22. Tang Y, Li M, Wang J, Pan Y, Wu FX. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks. Biosystems. 2015;127:67-72. DOI: 10.1016/j.biosystems.2014.11.005.
23. Metz CE. Receiver operating characteristic analysis: a tool for the quantitative evaluation of observer performance and imaging systems. J Am Coll Radiol. 2006;3(6):413-422. DOI: 10.1016/j.jacr.2006.02.021.
24. Anaya J. OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. Peer J Comput Sci. 2016;2:e67,1-20. DOI: 10.7717/peerj-cs.67.

25. Benítez-Parejo N, del Águila MMR, Pérez-Vicente S. Survival analysis and Cox regression. Allergol Immunopathol. 2011;39(6):362-373. DOI: 10.1016/j.aller.2011.07.007.
26. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, *et al*. The mutational landscape of head and neck squamous cell carcinoma. Science. 2011;333(6046):1157-1160. DOI: 10.1126/science.1208130.
27. Bhagwat AS, Vakoc CR. Targeting transcription factors in cancer. Trends Cancer. 2015;1(1):53-65. DOI: 10.1016/j.trecan.2015.07.001.
28. Lambert M, Jambon S, Depauw S, David-Cordonnier MH. Targeting transcription factors for cancer treatment. Molecules. 2018;23(6):1479,1-51. DOI: 10.3390/molecules23061479.
29. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. Lancet. 2008;371(9625):1695-1709. DOI: 10.1016/S0140-6736(08)60728-X.
30. Lee C, Kim JS, Waldman T. PTEN gene targeting reveals a radiation-induced size checkpoint in human cancer cells. Cancer Res. 2004;64(19):6906-6914. DOI: 10.1158/0008-5472.CAN-04-1767.
31. Suske G, Bruford E, Philipsen S. Mammalian SP/KLF transcription factors: bring in the family. Genomics. 2005;85(5):551-556. DOI: 10.1016/j.ygeno.2005.01.005.
32. Safe S, Abbruzzese J, Abdelrahim M, Hedrick E. Specificity protein transcription factors and cancer: opportunities for drug development. Cancer Prev Res. 2018;11(7):371-382. DOI: 10.1158/1940-6207.CAPR-17-0407.
33. Jiang NY, Woda BA, Banner BF, Whalen GF, Dresser KA, Lu D. Sp1, a new biomarker that identifies a subset of aggressive pancreatic ductal adenocarcinoma. Cancer Epidemiol Biomarkers Prev. 2008;17(7):1648-1652. DOI: 10.1158/1055-9965.EPI-07-2791.
34. Gao Y, Gan K, Liu K, Xu B, Chen M. SP1 expression and the clinicopathological features of tumors: a meta-analysis and bioinformatics analysis. Pathol Oncol Res. 2021;27:581998,1-12. DOI: 10.3389/pore.2021.581998.
35. Dong Q, Cai N, Tao T, Zhang R, Yan W, Li R, *et al*. An axis involving SNAI1, microRNA-128 and SP1 modulates glioma progression. PLoS One. 2014;9(6):e98651,1-11. DOI: 10.1371/journal.pone.0098651.
36. Maurer GD, Leupold JH, Schewe DM, Biller T, Kates RE, Hornung HM, *et al*. Analysis of specific transcriptional regulators as early predictors of independent prognostic relevance in resected colorectal cancer. Clin Cancer Res. 2007;13(4): 1123-1132. DOI: 10.1158/1078-0432.CCR-06-1668.
37. Hu X, Cui C, Sun T, Wang W. Associations between ADIPOQ rs2241766 SNP and breast cancer risk: a systematic review and a meta-analysis. Genes Environ. 2021;43:48,1-9. DOI: 10.1186/s41021-021-00221-2.
38. Morris LGT, Sikora AG, Patel SG, Hayes RB, Ganly I. Second primary cancers after an index head and neck cancer: subsite-specific trends in the era of human papillomavirus-associated oropharyngeal cancer. J Clin Oncol. 2011;29(6):739-746. DOI: 10.1200/JCO.2010.31.8311.
39. Westra WH. The changing face of head and neck cancer in the 21st century: the impact of HPV on the epidemiology and pathology of oral cancer. Head Neck Pathol. 2009;3(1):78-81. DOI: 10.1007/s12105-009-0100-y.
40. Wang L, Wei D, Huang S, Peng Z, Le X, Wu TT, *et al*. Transcription factor Sp1 expression is a significant predictor of survival in human gastric cancer. Clin Cancer Res. 2003;9(17):6371-6380. PMID: 14695137.
41. Yao JC, Wang L, Wei D, Gong W, Hassan M, Wu TT, *et al*. Association between expression of transcription factor Sp1 and increased vascular endothelial growth factor expression, advanced stage, and poor survival in patients with resected gastric cancer. Clin Cancer Res. 2004;10(12): 4109-4117. DOI: 10.1158/1078-0432.CCR-03-0628.
42. Wang LW, Li Q, Hua ZL, Zhou F, Keping X, Daoyan W, *et al*. Expression of transcription factor Sp1 in human gastric cancer tissue and its correlation with prognosis. Zhonghua Zhong Liu Za Zhi. 2007;29(2):107-111. PMID: 17645844.
43. Jutooru I, Guthrie AS, Chadalapaka G, Pathi S, Kim K, Burghardt R, *et al*. Mechanism of action of phenethylisothiocyanate and other reactive oxygen species-inducing anticancer agents. Mol Cell Biol. 2014;34(13):2382-2395. DOI: 10.1128/MCB.01602-13.
44. Safe S. Specificity proteins (Sp) and cancer. Int J Mol Sci. 2023;24(6):5164,1-22. DOI: 10.3390/ijms24065164.
45. Zhang B, Wang H, Guo Z, Zhang X. A panel of transcription factors identified by data mining can predict the prognosis of head and neck squamous cell carcinoma. Cancer Cell Int. 2019;19:297,1-10. DOI: 10.1186/s12935-019-1024-6.
46. Jin Y, Qin X. Comprehensive analysis of transcriptome data for identifying biomarkers and therapeutic targets in head and neck squamous cell carcinoma. Ann Transl Med. 2020;8(6): 282,1-17. DOI: 10.21037/atm.2020.03.30.
47. Yan L, Zhan C, Wu J, Wang S. Expression profile analysis of head and neck squamous cell carcinomas using data from The Cancer Genome Atlas. Mol Med Rep. 2016;13(5):4259-4265. DOI: 10.3892/mmr.2016.5054.

Supplementary Materials

Table S1. Summary statistics and distribution of variables in the study population. Descriptive statistics were elucidated in terms of the median with its corresponding interquartile range for numeric variables, while categorical variables were conveyed by their respective frequencies and associated percentages.

Variables	Levels	Mean ± SD / Frequency (%)
Days to last follow-up	-----	616.15 (171.75, 847.25)
Days to death	-----	747.85 (215.50, 800.25)
Tobacco smoking history	-----	2.46 (2.00, 4.00)
Year of tobacco smoking onset	-----	1967.31 (1959.00, 1975.00)
Stopped smoking year	-----	1997.25 (1989.75, 2009.00)
Number of pack years smoked	-----	45.75 (25.00, 60.00)
Amount of alcohol consumption per day	-----	3.24 (0.00, 5.00)
Anatomic neoplasm subdivision	Alveolar ridge	18 (3.41)
	Base of tongue	27 (5.11)
	Buccal mucosa	23 (4.36)
	Floor of mouth	63 (11.93)
	Hard palate	7 (1.33)
	Hypopharynx	10 (1.89)
	Larynx	117 (22.16)
	Lip	3 (0.57)
	Oral cavity	73 (13.83)
	Oral tongue	133 (25.19)
	Oropharynx	9 (1.70)
	Tonsil	45 (8.52)
Gender	Male	386 (73.11)
	Female	142 (26.89)
Vital status	Alive	358 (67.80)
	Dead	170 (32.20)
Clinical stage	Stage I	21 (3.98)
	Stage II	99 (18.75)
	Stage III	107 (20.27)
	Stage IVA	269 (50.95)
	Stage IVB	11 (2.08)
	Stage IVC	7 (1.33)
	NA	14 (2.65)
	Not evaluated	114 (21.59)
	Unknown	8 (1.52)
HPV status	Negative	74 (14.02)
	Positive	41 (7.77)
	NA	291 (55.11)
Alcohol history documented	No	165 (31.25)
	Yes	352 (66.67)
	NA	11 (2.08)

Table S2. The list of the top 100 differentially expressed genes in head and neck squamous cell carcinoma samples has been shown and ordered according to adjusted *P*-value numbers.

Ensemble ID	Gene symbol	Log 2 of fold change	Average expression	*P*-value	Adjusted *P*-value
ENSG00000250133	HOXC-AS2	4.067035229	-1.148967402	2.67E-61	6.77E-57
ENSG00000102547	CAB39L	-2.250606008	2.62578651	1.72E-60	1.45E-56
ENSG00000181092	ADIPOQ	-7.297369413	-4.67972223	1.72E-60	1.45E-56
ENSG00000034971	MYOC	-6.497505799	-4.397185746	6.03E-60	3.82E-56
ENSG00000197757	HOXC6	4.410758803	0.438512566	3.12E-59	1.58E-55
ENSG00000152642	GPD1L	-2.643818057	3.928493656	4.92E-55	1.79E-51
ENSG00000166819	PLIN1	-4.967671556	-2.462249755	4.95E-55	1.79E-51
ENSG00000070081	NUCB2	-1.87934893	5.407735058	1.36E-54	4.29E-51
ENSG00000128713	HOXD11	6.052475249	1.296197492	3.95E-54	1.11E-50
ENSG00000158458	NRG2	-4.18405847	-1.928509504	6.35E-54	1.61E-50
ENSG00000167588	GPD1	-4.703472835	-0.454576838	8.04E-54	1.85E-50
ENSG00000278966	AL031602.1	4.247129426	-0.571218843	1.06E-53	2.22E-50
ENSG00000130309	COLGALT1	1.4236552	7.359154411	5.94E-53	1.16E-49
ENSG00000106351	AGFG2	-1.96785504	4.122153348	8.87E-51	1.60E-47
ENSG00000169258	GPRIN1	2.329497338	4.031573369	1.28E-50	2.16E-47
ENSG00000248554	AC114956.2	3.715658571	-0.008766588	1.48E-50	2.34E-47
ENSG00000150672	DLG2	-3.486651401	-0.589088001	2.16E-50	3.21E-47
ENSG00000237424	FOXD2-AS1	2.30376284	1.254237311	9.66E-50	1.36E-46
ENSG00000146670	CDCA5	2.045166234	5.378767931	1.29E-49	1.72E-46
ENSG00000180806	HOXC9	4.350482873	-0.038974379	2.52E-49	3.19E-46
ENSG00000101057	MYBL2	2.272679251	6.162260408	9.00E-49	1.08E-45
ENSG00000198478	SH3BGRL2	-4.130054038	3.194610672	2.46E-48	2.83E-45
ENSG00000025423	HSD17B6	2.301960774	1.051372234	4.00E-48	4.39E-45
ENSG00000184811	TUSC5	-4.658747805	-4.506695231	4.80E-48	5.06E-45
ENSG00000115163	CENPA	2.128467882	3.25318077	5.85E-48	5.92E-45
ENSG00000214544	GTF2IRD2P1	3.506956725	1.203269492	1.55E-47	1.50E-44
ENSG00000253293	HOXA10	4.176252692	1.161212768	1.88E-47	1.76E-44
ENSG00000139800	ZIC5	5.618903803	0.489738233	3.26E-47	2.95E-44
ENSG00000142945	KIF2C	1.991255599	5.252019293	3.70E-47	3.23E-44
ENSG00000043355	ZIC2	4.629791385	1.849227644	3.85E-47	3.24E-44
ENSG00000258711	AL358334.2	4.171140883	0.311740722	4.64E-47	3.79E-44
ENSG00000198099	ADH4	-4.040719663	-4.611534004	5.47E-47	4.32E-44
ENSG00000187288	CIDEC	-4.903313148	-4.018593361	7.27E-47	5.57E-44
ENSG00000122042	UBL3	-1.63484831	4.967289285	1.09E-46	8.08E-44
ENSG00000171503	ETFDH	-1.469479003	4.073134182	1.67E-46	1.20E-43
ENSG00000248240	AC114956.1	3.378352012	-1.495142579	7.85E-46	5.51E-43
ENSG00000267123	LINC02081	4.565860478	-0.451339804	1.27E-45	8.69E-43
ENSG00000165795	NDRG2	-2.905073821	5.929488402	2.78E-45	1.85E-42
ENSG00000187498	COL4A1	2.800204983	8.231112588	5.42E-45	3.51E-42
ENSG00000171201	SMR3B	-6.332001777	-5.553446752	1.03E-44	6.50E-42
ENSG00000096006	CRISP3	-8.652918707	-1.05688244	1.12E-44	6.94E-42
ENSG00000117122	MFAP2	3.070259914	5.281036959	1.45E-44	8.72E-42
ENSG00000088325	TPX2	2.000500763	6.551291145	2.18E-44	1.28E-41
ENSG00000168309	FAM107A	-3.813698518	1.134875309	2.88E-44	1.66E-41
ENSG00000037965	HOXC8	4.311335319	-0.549416927	4.06E-44	2.28E-41
ENSG00000154920	EME1	1.947067143	2.037572083	4.56E-44	2.51E-41
ENSG00000166851	PLK1	1.956518538	5.633233912	9.19E-44	4.95E-41
ENSG00000179528	LBX2	2.475528163	-0.702628111	1.12E-43	5.90E-41
ENSG00000196616	ADH1B	-6.687233254	-2.394052899	1.31E-43	6.75E-41
ENSG00000234041	AL512326.3	4.596160242	-1.284915057	2.31E-43	1.17E-40
ENSG00000164283	ESM1	3.397047925	1.900907484	4.36E-43	2.16E-40
ENSG00000000005	TNMD	-3.598553987	-5.596754377	4.59E-43	2.23E-40
ENSG00000095752	IL11	3.998145151	1.931565032	4.82E-43	2.30E-40

Table S2. Continued

Ensemble ID	Gene symbol	Log 2 of fold change	Average expression	*P*-value	Adjusted *P*-value
ENSG00000172340	SUCLG2	-1.193031896	5.037065521	7.10E-43	3.33E-40
ENSG00000090889	KIF4A	1.89933935	4.741108111	1.13E-42	5.21E-40
ENSG00000134871	COL4A2	2.534128298	8.525642026	1.62E-42	7.34E-40
ENSG00000138180	CEP55	2.06542102	5.296787324	2.59E-42	1.15E-39
ENSG00000127564	PKMYT1	2.086465524	4.943314123	2.81E-42	1.22E-39
ENSG00000060762	MPC1	-1.477814284	4.013536801	4.14E-42	1.78E-39
ENSG00000134013	LOXL2	3.069342726	5.643654837	5.91E-42	2.49E-39
ENSG00000108381	ASPA	-3.739067205	-2.388581691	8.85E-42	3.67E-39
ENSG00000099953	MMP11	4.564027571	5.426080692	1.40E-41	5.72E-39
ENSG00000134240	HMGCS2	-6.101300197	-3.47146165	1.72E-41	6.82E-39
ENSG00000281386	AP003500.1	-4.308213832	-4.281470987	1.72E-41	6.82E-39
ENSG00000100985	MMP9	4.002109105	6.202632834	1.77E-41	6.87E-39
ENSG00000186185	KIF18B	1.983342292	4.239610824	2.22E-41	8.53E-39
ENSG00000175063	UBE2C	1.998033406	5.452722034	2.91E-41	1.10E-38
ENSG00000094804	CDC6	1.853484263	4.840907474	3.22E-41	1.20E-38
ENSG00000174371	EXO1	1.919035726	3.627558953	3.29E-41	1.21E-38
ENSG00000039537	C6	-4.910536017	-4.57173203	3.92E-41	1.42E-38
ENSG00000124205	EDN3	-5.719837546	-3.797732365	5.27E-41	1.88E-38
ENSG00000171208	NETO2	1.997386805	4.852666648	6.39E-41	2.25E-38
ENSG00000089685	BIRC5	1.921936762	5.913256413	6.82E-41	2.36E-38
ENSG00000008441	NFIX	-1.899657175	6.047103492	1.02E-40	3.49E-38
ENSG00000235097	LINC00330	-4.849489969	-4.192716077	1.30E-40	4.37E-38
ENSG00000127423	AUNIP	1.760474148	2.413044437	1.32E-40	4.41E-38
ENSG00000167676	PLIN4	-4.141935223	1.045125369	1.66E-40	5.47E-38
ENSG00000093009	CDC45	1.956367959	4.269708117	1.82E-40	5.92E-38
ENSG00000123485	HJURP	1.995317333	4.061303218	2.68E-40	8.60E-38
ENSG00000197467	COL13A1	2.398246455	2.115312569	4.20E-40	1.33E-37
ENSG00000272549	LINC02538	-4.380681114	-4.147392374	4.51E-40	1.41E-37
ENSG00000168779	SHOX2	3.210264394	1.719294643	9.26E-40	2.86E-37
ENSG00000163815	CLEC3B	-3.017723679	0.960674494	1.48E-39	4.52E-37
ENSG00000118193	KIF14	2.15364894	3.721370764	1.58E-39	4.76E-37
ENSG00000075218	GTSE1	1.815903525	3.949201309	1.80E-39	5.34E-37
ENSG00000111713	GYS2	-4.635748285	-3.991607189	1.98E-39	5.81E-37
ENSG00000148848	ADAM12	3.712593833	4.722028621	2.22E-39	6.44E-37
ENSG00000107159	CA9	5.470416088	3.343690927	2.83E-39	8.06E-37
ENSG00000165269	AQP7	-4.018186501	-2.308894216	2.84E-39	8.06E-37
ENSG00000133466	C1QTNF6	2.799040555	4.867511201	3.90E-39	1.10E-36
ENSG00000111206	FOXM1	2.055018305	5.935424143	4.42E-39	1.23E-36
ENSG00000111665	CDCA3	1.81237011	4.049239803	4.94E-39	1.36E-36
ENSG00000204889	KRT40	-4.599123136	-4.369657322	5.32E-39	1.45E-36
ENSG00000164932	CTHRC1	3.372348783	4.641354086	7.70E-39	2.07E-36
ENSG00000087586	AURKA	1.745079635	4.74469077	1.09E-38	2.90E-36
ENSG00000113739	STC2	3.088138131	3.966544387	1.53E-38	4.03E-36
ENSG00000181234	TMEM132C	-4.677230557	-3.297734314	1.58E-38	4.11E-36
ENSG00000123388	HOXC11	5.270887041	0.245222717	1.75E-38	4.51E-36
ENSG00000091651	ORC6	1.802906347	3.3609245	3.40E-38	8.68E-36
ENSG00000076382	SPAG5	1.668663738	5.013816972	3.92E-38	9.92E-36