Check for updates

OPEN ACCESS

EDITED BY Fabio Grizzi, Humanitas Research Hospital, Italy

REVIEWED BY Mingyi Wang, Shanghai Jiao Tong University, China Lei Wang, Hebei Medical University, China

*CORRESPONDENCE Xi Yu yuxi316@126.com

SPECIALTY SECTION This article was submitted to Otorhinolaryngology - Head and Neck Surgery, a section of the journal Frontiers in Surgery

RECEIVED 25 May 2022 ACCEPTED 21 November 2022 PUBLISHED 04 January 2023

CITATION

Lv X and Yu X (2023) Signatures and prognostic values of related immune targets in tongue cancer.

Front. Surg. 9:952389. doi: 10.3389/fsurg.2022.952389

COPYRIGHT

© 2023 Lv and Yu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Signatures and prognostic values of related immune targets in tongue cancer

Xiaofei Lv¹ and Xi Yu^{2*}

¹Department of Stomatology, The Second Hospital of Tianjin Medical University, Tianjin, China, ²Department of Anesthesiology, The Second Hospital of Tianjin Medical University, Tianjin, China

Background: Tongue cancer, as one of the most malignant oral cancers, is highly invasive and has a high risk of recurrence. At present, tongue cancer is not obvious and easy to miss the opportunity for early diagnosis when in the advanced stage. It is important to find markers that can predict the occurrence and progression of tongue cancer.

Methods: Bioinformatics analysis plays an important role in the acquisition of marker genes. GEO and TCGA data are very important public databases. In addition to expression data, the TCGA database also contains corresponding clinical data. In this study, we screened three GEO data sets that met the standard, which included GSE13601, GSE34105, and GSE34106. These data sets were combined using the SVA package to prepare the data for differential expression analysis, and then the limma package was used to set the standard to p < 0.05 and $|log2 (FC)| \ge 1.5$.

Results: A total of 170 differentially expressed genes (DEGs) were identified. In addition, the DEseq package was used for differential expression analysis using the same criteria for samples in the TCGA database. It ended up with 1,589 DEGs (644 upregulated, 945 downregulated). By merging these two sets of DEGs, 5 common upregulated DEGs (*CCL20, SCG5, SPP1, KRT75, and FOLR3*) and 15 common downregulated DEGs were obtained.

Conclusions: Further functional analysis of the DEGs showed that *CCL20*, *SCG5*, and *SPP1* are closely related to prognosis and may be a therapeutic target of TSCC.

KEYWORDS

tongue cancer, DEGs, CCL20, SCG5, SPP1

Introduction

Head and neck tumors are the fifth most common malignancies in the world, and smoking and alcohol consumption are two of the most common risk factors associated with these lesions (1). Oral cancer accounts for 2.1% of all new cancer cases worldwide, and squamous cell carcinoma is a known malignancy, accounting for more than 90% of all oral cancers (2, 3). Tongue squamous cell carcinoma (TSCC) is one of the most common malignancies diagnosed in the oral cavity and is associated with a poor prognosis due to its high regional recurrence rate and lymphatic metastasis (4). It is reported that there are about 500,000 new cases of tongue cancer every year, the global incidence of tongue cancer is increasing year by year, and the onset of tongue cancer is getting younger and younger (5). Numerous lymphatic

vessels of the tongue are abundant in blood circulation, and the tongue is active frequently. These factors promote the early spread of cancer cells to adjacent tissues and organs, such as lymph nodes, the base of the mouth, throat, and neck. Although tongue cancer treatments are improving as technology advances, the 5-year survival rate is still poor (3, 6). TSCC has a lower survival rate than squamous cell carcinoma elsewhere in the mouth. There is currently a lack of TSCC markers, so finding reliable predictive markers is of great interest (7).

Materials and methods

Data preparation

The expression matrix was downloaded from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/). To improve the accuracy of the study, we screened three eligible data sets. These data sets included tongue cancer and normal tissues. GSE13601, GSE34105, and GSE34106 contain 57, 78, and 43 samples, respectively. The

SVA software package in R language was used to combine the samples of these three data sets to form a combined data set with 178 samples. Then, we downloaded the expression data of head and neck tumors from the TCGA database and screened the tongue squamous cell carcinoma samples, a total of 147 samples, including 140 tumor samples and 7 normal samples.

Screening and analysis of differentially expressed genes

The limma package and DESeq package in R language were used for differential expression analysis, and the standard cutoff was set to false discovery rate (FDR) < 0.05 and $|log2 (FC)| \ge 1.5$ for both GEO and TCGA, respectively (8, 9). The volcano package and heatmap package in R language were used to visualize DEGs. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, Gene Ontology (GO) analysis, and Gene Set Enrichment Analysis (GSEA) were performed for DEGs. Metascape website (https://metascape.org/) was used for pathway enrichment analysis of DEGs in TCGA. The *t*-test



FIGURE 1

Differential expression analysis of the GEO merged data set (GSE13601, GSE34105 and GSE34106). (A) PCA analysis was performed to identify differences between tumor tissue and normal tissue in the combined data set. (B) Volcano diagram representing the DEGs obtained by differential expression analysis using the limma package, where red represents upregulated DEGs and blue represents downregulated DEGs. (C) Heatmap showing the differentially expressed genes in three data sets, with the annotations above representing each data set and the annotations in the second row representing different types of samples. The differentially expressed genes are shown on the right, where red represents upregulated genes are shown on the right, where red represents the upregulated genes and blue represents downregulated genes. (D) KEGG pathway enrichment analysis of DEGs, in which red represents the upregulation gene enrichment pathway and gray represents the downregulation gene enrichment pathway. (E) GO enrichment analysis of DEGs.



was used to compare the expression in the two groups, and ANOVA was used to compare that in multiple groups. We further verified the expression of upregulated DEGs using the GEPIA (http://gepia.cancer-pku.cn/).

Protein interaction network

The interaction between proteins corresponding to each gene was analyzed using String website (https://string-db.org/), and then the results were visualized using CytoScope. Genemania website (http://genemania.org/) was used for functional enrichment analysis of DEGs.

Identify the hub gene

The DEGs obtained from combined dataset were crossed with those obtained from TCGA, including 5 up-regulated genes and 15 down-regulated genes.

Survival analysis

Kaplan-Meier Plotter website (https://kmplot.com/) (citation) was used for survival analysis of the screened HUB genes.

Results

Differential expression analysis of the GEO merged data set

We downloaded three data sets GSE13601, GSE34105, and GSE34106 containing tongue squamous cell carcinoma (TSCC) in the GEO database. Each data set contains tumor samples and normal tissue samples. We used the SVA package to merge these data sets and obtain a combined data set. To identify differences between TSCC samples and normal tongue samples, we performed principal component analysis (PCA) and found that the two groups could be



separated in the combined data set (Figure 1A). We used the limma package [$|\log 2 (FC)| \ge 1.5$, p < 0.05] in R to analyze the 178 samples and screened for DEGs. We identified 170 DEGs, including 104 upregulated genes and 66 downregulated genes, which were shown in the volcano diagram, where the red on the right represents the upregulated differentially expressed genes and the blue one represents the downregulated differentially expressed genes (Figure 1B). In the heatmap, groups were made according to data sets and sample characteristics to explore the expression of different genes. The top annotation represents the different data sets. DEGs are marked on the right side of the figure (Figure 1C). To further explore the functions of the screened DEGs, we used R language for further functional enrichment analysis of these genes in the KEGG database and enriched some up-regulated pathways such as cytokine-cytokine receptor interaction, viral protein interaction with cytokine and cytokine receptor, and primary immunodeficiency (Figure 1D). Then, GO analysis was performed on these genes, and the pathways enriched in upregulated genes were mainly cartilage development, connective tissue development, skeletal system development, and chondrocyte differentiation in the biological process (BP) module (Figure 1E). In the cellular component (CC) module, the enrichment pathways were mainly the collagen-containing extracellular matrix, basement membrane, apical part of the cell, and laminin complex. In the molecular function (MF) module, the pathways were mainly enriched in receptor ligand activity, signaling receptor activator activity, and metalloendopeptidase activity. GSEA analysis was also performed on the first six upregulated pathways, and the results were consistent with the previous analysis (**Figure 2**). The KEGG correspondence is in Supplementary Table S2.

Differential expression analysis of TSCC in the TCGA database

DEGs of TSCC vs. normal samples in the TCGA database was performed using the DESeq package in R language [|log2 (FC)| \geq 1.5, p < 0.05]. A total of 644 DEGs upregulated and 945 downregulated DEGs were found, as shown in the volcano diagram (**Figure 3A**). We used Metascape website to analyze these upregulated DEGs and found that the mainly enriched GO pathways were the aminoglycoside antibiotic metabolic process, proximal/distal pattern formation, and negative regulation of blood coagulation (**Figure 3B**). KEGG



enrichment analysis of upregulated DEGs showed that the upregulated pathways were mainly folate biosynthesis, steroid hormone biosynthesis, complement and coagulation cascades, *Staphylococcus aureus* infection, and cytokine–cytokine receptor interaction. The main downregulated pathways were salivary secretion, calcium signaling pathway, and pancreatic secretion (**Figure 3C**).

Acquisition and analysis of hub genes

Differentially expressed genes (DEGs) acquired in GEO were crossed with those in TCGA, and 5 common upregulated genes (*CCL20*, *SCG5*, *SPP1*, *FOLR3*, and *KRT75*) and 15 common downregulated genes (*MAOB*, *HLFSELENBP1*, *CRISP3*, *TMPRSS2*, *TLX1*, *TGFBR3*, *APOD*, *PPP1R3C*, *SLITRK5*, *MFAP4*, *STATH*, *GSTM5*, *RNASE4*, *COX7A1*, *FHL1*, *SLC25A4*, *EFHD1*, *MYH11*, *MGP*, *SCGB2A1*, *LPIN1*, *ELF5*, *ADH1B*, *CKMT2*) were screened (Figure 4A).

The protein–protein interaction (PPI) network analysis of these five upregulated genes indicated that proteins expressed by other genes were all functionally related, except *FOLR3* (**Figure 4B**). Then, we conducted interaction network analysis of these genes individually (**Figure 4C**).

Further expression verification of these hub genes

We further verified the expression of upregulated DEGs using the GSE13601 data set. This data set includes tumor samples and their paired normal samples. After removing the individual samples from the database, we obtained 20 pairs of paired samples and then compared the expression levels of *CCL20*, *SCG5*, *SPP1*, *FOLR3*, and *KRT75* genes in tumor tissues vs. normal tissues. It was shown that the expression of these five genes was higher in TSCC than that in normal tissue both in the whole sample and the most paired samples



(Figure 5A). The same results can be obtained using the TCGA data set (Figure 5B).

To explore the relationship between the expression of these Hub genes and clinical stage, we used the TCGA database to explore and found that the expression level of CCL20 was significantly increased in patients with stages III and IV compared with stage I (p < 0.05) and the expression of SPP1 and SCG5 in stage IV was significantly higher than that in stage I patients (p < 0.05) (Figure 5C).

Multivariate COX regression of TSCC in the TCGA database for each clinical feature and prognosis

To further explore the clinical data of TCGA, multivariate Cox regression was performed using clinical data from the TCGA database (Figure 6). We find that risk increased when T stage > 2. Similarly, years of smoking also has an impact on the survival of patients; when the smoking time is more than 29 years, the risk is significantly increased. In addition, we found that when the TSCC stage is higher than grade III, age > 59, N stage > 1, and the number of cigarettes

smoked per day > 2.5, the risk is increased; however, the *P* value of multivariate Cox analysis did not meet the statistical standards. Kaplan–Meier analysis was performed on clinical data using SPSS software, and we found that patients with a stage higher than III had a significantly worse prognosis than patients with a stage lower than III. AJCC_ pathologic_T had a poor prognosis when it was higher than T2. Similarly, the AJCC_pathologic_N stage higher than N1 and smoking time more than 29 years remained risk factors (Figure 7).

Hub genes analysis in HNSCC

These upregulated hub genes were further verified by GEPIA, and the results showed that the expression of these five upregulated genes in head and neck squamous cell carcinoma (HNSCC) was also higher than that in normal tissues (**Figure 8**). Kaplan–Meier Plotter website was used to analyze the relationship between the expression of these genes and prognosis, and the results showed that high expression of these genes was strongly associated with poor prognosis in HNSCC (p < 0.05) (**Figure 9**).

Characteriistics	Number (%)		HR (95%CI)	P.value
stage				
<= []]	82 (59.0)			
>]]]	57 (41.0)	→	1.65 [0.83, 3.31]	0.155
age				
<=59	58 (41.7)			
>59	81 (58.3)	▶	1.12 [0.61, 2.06]	0.707
ajcc_pathologic_t				
<=T2	65 (46.8)			
>T2	61 (43.9)	⊢ → -	2.83 [1.34, 5.97]	0.006
unkown	13 (9.4)	H	1.14 [0.21, 6.01]	0.881
ajcc_pathologic_n				
<=N1	69 (49.6)			
>N1	51 (36.7)	↓ → ↓	2.10 [0.93, 4.73]	0.075
unkown	19 (13.7)	•	2.38 [0.64, 8.84]	0.197
cigarettes_per_day				
<=2.5	53 (38.1)			
>2.5	24 (17.3)	→	1.29 [0.53, 3.13]	0.578
unkown	62 (44.6)	⊳ I	0.60 [0.27, 1.32]	0.204
years_smoked				
<-29	18 (12.9)			
>29	21 (15.1)	⊢ ◆ −−−	→ 3.56 [1.24, 10.22]	0.018
unkown	100 (71.9)	⊢ ♦−−−{	1.82 [0.69, 4.85]	0.228
treatments_radiation_treatment_or_therapy				
0	41 (29.5)			
1	87 (62.6)	*	0.28 [0.13, 0.59]	0.001
unkown	11 (7.9)	▶	0.72 [0.24, 2.20]	0.567







Conclusion

Key molecular alterations need to be confirmed to identify effective therapeutic targets (10-14). To explore the genes related to the occurrence and progression of tongue cancer, we used GEO and TCGA databases for screening. By increasing the sample size, we used the SVA package in R language for GSE13601, GSE34105, and GSE34106. The three data sets were combined to obtain a data set containing 178 samples (121 TSCC and 57 normal samples). By using the limma package for DEG analysis ($|log_2FC| > 1.5$, p < 0.05), 104 upregulated genes and 66 downregulated genes were screened. Then, the DEGs were obtained by using the DEseq package in R language, and the cut-off value was set the same as before. The DEGs obtained from GEO and TCGA data sets were crossed, obtaining 5 upregulated and 15 downregulated DEGs. The upregulated DEGs are CCL20, SCG5, SPP1, FOLR3, and KET75.

Chemokine (C–C motif) Ligand 20 (CCL20) is a substance involved in tissue validation and homeostasis with a specific receptor C–C chemokine receptor 6 (CCR6) (15). It is expressed in various tissues and immune cells in human body, and the CCL20–CCR6 axis is closely associated with inflammation and infectious diseases. The CCL20–CCR6 axis is closely related to a variety of cancers, which can directly promote the progress of cancer by enhancing the migration and proliferation of cancer cells and can also regulate the tumor microenvironment by immune cells (16).

CCL20 has been reported to indirectly promote tumor progression by recruiting Treg, Th17, and Th22 cells to maintain the development and immunosuppressive microenvironment (16). Studies have shown that TSCC cells produce CCL20 after interacting with macrophages, and the cells can express CCR6 in the TSCC microenvironment. The CCL20–CCR6 axis may be associated with OSCC progression by inducing CD163 expression in macrophages (17). SPP1 is secreted glycophosphoprotein, which is involved in a variety of functions, including cell adhesion, migration and invasion (18–21). *FOLR3* is highly overexpressed on several tumor cells, including ovarian, nonsmall cell lung, kidney, brain, endometrial, colorectal, breast, pancreatic, gastric, prostate, testicular, and bladder cancer; thus, it is known as therapeutic target for cancer treatment (22–24). It has been reported that the genes are correlated with the progression of a variety of tumors. In addition, according to the results of survival analysis, we found that *CCL20*, *SPP1*, and *SCG5* are closely related to the prognosis of TSCC and HCSN. We have reason to believe that these genes can be developed into prognostic indicators and are expected to be therapeutic targets of TSCC.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

Author contributions

Conceptualization, XY, XL; material preparation, XY; writing—original draft preparation, XY, XL; data analysis and interpretation, XY, XL; and writing—editing and review, all

References

1. Ho AS, Kraus DH, Ganly I, Lee NY, Shah JP, Morris LG. Decision making in the management of recurrent head and neck cancer. *Head Neck*. (2014) 36 (1):144–51. doi: 10.1002/hed.23227

2. Furukawa A, Shinoda M, Kubo A, Honda K, Akasaka R, Yonehara Y, et al. Endothelin signaling contributes to modulation of nociception in early-stage tongue cancer in rats. *Anesthesiology*. (2018) 128(6):1207–19. doi: 10.1097/ALN. 000000000002139

3. Huang L, Luo EL, Xie J, Gan RH, Ding LC, Su BH, et al. FZD2 Regulates cell proliferation and invasion in tongue squamous cell carcinoma. *Int J Biol Sci.* (2019) 15(11):2330–9. doi: 10.7150/ijbs.33881

4. Bagan J, Sarrion G, Jimenez Y. Oral cancer: clinical features. Oral Oncol. (2010) 46(6):414-7. doi: 10.1016/j.oraloncology.2010.03.009

5. Ng JH, Iyer NG, Tan MH, Edgren G. Changing epidemiology of oral squamous cell carcinoma of the tongue: a global study. *Head Neck*. (2017) 39 (2):297-304. doi: 10.1002/hed.24589

6. Zhang Z, Pan J, Li L, Wang Z, Xiao W, Li N. Survey of risk factors contributed to lymphatic metastasis in patients with oral tongue cancer by immunohistochemistry. *J Oral Pathol Med.* (2011) 40(2):127–34. doi: 10.1111/j. 1600-0714.2010.00953.x

7. Heikkinen I, Bello IO, Wahab A, Hagström J, Haglund C, Coletta RD, et al. Assessment of tumor-infiltrating lymphocytes predicts the behavior of early-stage oral tongue cancer. *Am J Surg Pathol.* (2019) 43(10):1392–6. doi: 10.1097/PAS. 00000000001323

8. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43(7):e47. doi: 10.1093/nar/gkv007

9. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. (2010) 11(10):R106. doi: 10.1186/gb-2010-11-10-r106 authors. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the Youth Scientific Research Fund of the Second Hospital of Tianjin Medical University (No. 2020ydey08) and the Research Project of Tianjin Education Commission (No. 2021KJ232).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

10. Hussein AA, Forouzanfar T, Bloemena E, de Visscher J, Brakenhoff RH, Leemans CR, et al. A review of the most promising biomarkers for early diagnosis and prognosis prediction of tongue squamous cell carcinoma. *Br J Cancer.* (2018) 119(6):724–36. doi: 10.1038/s41416-018-0233-4

11. Mithani SK, Mydlarz WK, Grumbine FL, Smith IM, Califano JA. Molecular genetics of premalignant oral lesions. *Oral Dis.* (2007) 13(2):126–33. doi: 10.1111/j.1601-0825.2006.01349.x

12. Vigneswaran N, Williams MD. Epidemiologic trends in head and neck cancer and aids in diagnosis. Oral Maxillofac Surg Clin North Am. (2014) 26 (2):123-41. doi: 10.1016/j.coms.2014.01.001

13. Ziegler A, Thorpe E. Oral tongue cancer in a patient with hereditary nonpolyposis colorectal cancer: a case report and review of the literature. *Oral Oncol.* (2019) 92:92–3. doi: 10.1016/j.oraloncology.2019.03.008

14. Lee DY, Kang Y, Im NR, Kim B, Kwon TK, Jung KY, et al. Actin-associated gene expression is associated with early regional metastasis of tongue cancer. *Laryngoscope*. (2021) 131(4):813–9. doi: 10.1002/lary.29025

15. Walch-Rückheim B, Mavrova R, Henning M, Vicinus B, Kim YJ, Bohle RM, et al. Stromal fibroblasts induce CCL20 through IL6/C/EBPβ to support the recruitment of Th17 cells during cervical cancer progression. *Cancer Res.* (2015) 75(24):5248–59. doi: 10.1158/0008-5472.CAN-15-0732

16. Wu N, Wang Y, Wang K, Zhong B, Liao Y, Liang J, et al. Cathepsin K regulates the tumor growth and metastasis by IL-17/CTSK/EMT axis and mediates M2 macrophage polarization in castration-resistant prostate cancer. *Cell Death Dis.* (2022) 13(9):813. doi: 10.1038/s41419-022-05215-8

17. Packer L, Pavey S, Parker A, Stark M, Johansson P, Clarke B, et al. Osteopontin is a downstream effector of the PI3-kinase pathway in melanomas that is inversely correlated with functional PTEN. *Carcinogenesis.* (2006) 27 (9):1778–86. doi: 10.1093/carcin/bgl016

18. Zhou Y, Dai DL, Martinka M, Su M, Zhang Y, Campos EI, et al. Osteopontin expression correlates with melanoma invasion. *J Invest Dermatol.* (2005) 124 (5):1044–52. doi: 10.1111/j.0022-202X.2005.23680.x

19. Deng G, Zeng F, Su J, Zhao S, Hu R, Zhu W, et al. BET Inhibitor suppresses melanoma progression via the noncanonical NF- κ B/SPP1 pathway. *Theranostics*. (2020) 10(25):11428–43. doi: 10.7150/thno.47432

20. Elias EG, Hasskamp JH, Sharma BK. Cytokines and growth factors expressed by human cutaneous melanoma. *Cancers (Basel).* (2010) 2 (2):794–808. doi: 10.3390/cancers2020794

21. Wang Y, Wu N, Jiang N. Autophagy provides a conceptual therapeutic framework for bone metastasis from prostate cancer. Cell Death Dis. (2021) 12 (10):909. doi: 10.1038/s41419-021-04181-x

22. Bin-Alee F, Arayataweegool A, Buranapraditkun S, Mahattanasakul P, Tangjaturonrasme N, Hirankarn N, et al. Transcriptomic analysis of peripheral blood mononuclear cells in head and neck squamous cell carcinoma patients. *Oral Dis.* (2021) 27(6):1394–402. doi: 10.1111/odi.13639

23. Kageyama S, Nagao Y, Ma J, Asakawa M, Yoshida R, Takeshita T, et al. Compositional shift of oral Microbiota following surgical resection of tongue cancer. *Front Cell Infect Microbiol.* (2020) 10:600884. doi: 10.3389/fcimb.2020. 600884

24. Tagliabue M, Belloni P, De Berardinis R, Gandini S, Chu F, Zorzi S, et al. A systematic review and meta-analysis of the prognostic role of age in oral tongue cancer. *Cancer Med.* (2021) 10(8):2566-78. doi: 10.1002/cam4. 3795