Clinical Trial/Experimental Study

Medicine

OPEN

Corneodesmosin as a potential target of oral squamous cell carcinoma

Yong-Bin Di, MD[a], Yang Bao, MD[b], Jie Guo, MD[b], Wei Liu, MD[b], Su-Xin Zhang, MD[b], Guan-Hua Zhang, MD[b], Tian-Ke Li, MD[b,*]

Abstract:

Objective: The relationship between oral squamous cell carcinoma (OSCC) and Corneodesmosin (CDSN) remains unclear. This study aims to explore the correlation between CDSN and the prognosis and survival time of patients with OSCC.

Methods: Bioinformatics were used to identify the hub role of CDSN in the OSCC. A total of 200 patients with OSCC were recruited. Clinical and follow-up data were recorded, and the expression level of CDSN was detected. Pearson chi-square test and Spearman correlation coefficient were used to analyze the relationship between prognosis and related parameters in patients with OSCC. Univariate and multivariate Logistic regression and Cox proportional risk regression were applied for further analysis, and receiver operating characteristic curve and survival curve of subjects were plotted.

Results: CDSN was identified as the most significant hub gene of the OSCC by the cytoHubba. By the comparative toxicogenomics database (CTD) analysis, there was strong relationship between the CDSN and mouth neoplasms, head and neck neoplasms, squamous cell carcinoma of head and neck. The OSCC patients with low expression level of CDSN have poor overall survival compared with the high expression level of CDSN (HR = 0.75, 95% confidence interval [95%CI]: 0.57-0.98, $P = .036$). Spearman correlation coefficient analysis showed that CDSN expression level was significantly correlated with prognosis ($\rho = -0.528$, $P < .001$). Multivariate Logistic regression analysis showed that poor prognosis (odds ratio [OR] = 0.096, 95%CI: 0.049-0.189, $P < .001$) was significantly associated with low expression of CDSN. Cox regression analysis showed that the survival time of OSCC patients was shorter when CDSN expression was low (HR = 0.588, 95%CI: 0.420-0.823, $P = .002$). Strong predictive value of CDSN for the OSCC survival time was obtained by the biological process (BP)-neural network and support vector machine (SVM).

Conclusion: CDSN was significantly correlated with OSCC, and the shorter the survival time of patients with OSCC was, the worse the prognosis was.

Abbreviations: 95% CI = 95% confidence interval, BP = biological process, CDSN = corneodesmosin, CTD = comparative toxicogenomics database, DAVID = database for annotation, visualization and integrated discovery, DEGs = differently expressed genes, GEO = gene expression omnibus, OR = odds ratio, OSCC = oral squamous cell carcinoma, PCA = principal component analysis, PPI = protein-protein interaction, STRING = search tool for the retrieval of interacting genes, SVM = support vector machine.

Keywords: CDSN, oral squamous cell carcinoma, prognostic targets, survival time

1. Introduction

The most common cell source in oral cancer is squamous cells.[1] Oral squamous cell carcinoma (OSCC) is an aggressive tumor with different degrees of differentiation, which tends to early and extensive lymph node metastasis. It has become an increasingly serious worldwide problem, mainly occurring in smokers and alcoholics aged 40 to 70.[2,3] OSCC refers to the malignant transformation of oral squamous epithelium and the formation of malignant tumor. It is a common and harmful malignant tumor in oral cavity, which needs to be paid attention to.[4] Symptoms of OSCC may include a localized mass, lesions of the white mucous membrane, and localized ulcers. It can cause the loss of oral function, such as long influence the language

This work was supported by the Key science and technology research project of Hebei Provincial Health Commission [grant numbers202110801].

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

This study was approved by the Ethics Committee of the Fourth Hospital of Hebei Medical University. The research conformed to the Declaration of Helsinki.

Supplemental Digital Content is available for this article.

[a] Department of Stomatology, The First Hospital of Hebei Medical University, Shijiazhuang, P.R. China, [b] Department of Stomatology, The Fourth Hospital of Hebei Medical University, Shijiazhuang, P.R. China.

**Correspondence: Tian-Ke Li, Department of Stomatology, The Fourth Hospital of Hebei Medical University, Shijiazhuang 050000, P.R. China (e-mail: litianb112@163.com).*

How to cite this article: Di Y-B, Bao Y, Guo J, Liu W, Zhang S-X, Zhang G-H, Li T-K. Corneodesmosin as a potential target of oral squamous cell carcinoma. Medicine 2022;101:39(e28397).

Received: 30 May 2022 / Received in final form: 25 August 2022 / Accepted: 30 August 2022

http://dx.doi.org/10.1097/MD.0000000000030851

on the tongue, swallowing function, the function of the oral cavity of the affected, when further development may cause the stench or distant metastasis, and lead to local fester, affect the quality of life, in the end may also harm to the patient's life, Once detected, OSCC must be treated early.[5,6] Early OSCC after treatment, the cure rate is high. If metastasis occurs, a certain degree of cure can be achieved, usually in combination with radiotherapy and chemotherapy.[7]

CDSN (Corneodesmosin) encodes keratindesmosin, an adhesion glycoprotein secreted by keratinocytes granulosa cells and a key protein for human epidermal barrier function. Combined with the existing desmosomes, it exists in the connective structure between cuticle - cuticle desmosomes, also plays a key role in the adhesion of cuticle, is a part of the desmosomes plaque modified at the top of the epidermis.[8] CDSN is a protein-coding gene located in the major histocompatibility complex Class I region of chromosome 6.[9] The CDSN gene encodes a protein found in intercuticles, which locate in the human epidermis and other keratinized squamous epithelium. The coding protein undergoes a series of cleavages during keratinocyte maturation. This gene is highly polymorphic in the population, with variations associated with skin diseases such as psoriasis, oligotrichosis, and skin peeling syndrome through pathways including keratosis and developmental biology.[10] CDSN -associated genebody annotations include protein homologous dimer activity. Related studies have shown that CDSN is associated with ankylosing spondylitis.[11] It has also been suggested that dermatitis is a widespread inflammatory form of dermatitis syndrome caused by autosomal recessive nonsense mutations of CDSN.[12] Thus, CDSN may be associated with various types of inflammation. However, the relationship between CDSN and OSCC remains unclear.

Bioinformatics studies biological problems using the methods of applied mathematics, informatics, statistics, and computer science. Bioinformatics research materials and results are a variety of biological data, its research tools are computers, research methods include biological data search (collection and screening), processing (editing, sorting, management, and display) and utilization (calculation, simulation). The research interests include sequence alignment, gene recognition, gene recombination, protein structure prediction, gene expression, protein response prediction, and development of evolutionary models.

We hypothesized that CDSN expression plays an important role in the development of OSCC. In this study, bioinformatics analysis was used to verify the potential role of CDSN in OSCC, and 200 OSCC patients were recruited to study the impact of CDSN expression mutation on the prognosis and survival time of OSCC patients, so as to provide new ideas for its molecular mechanism.

2. Methods

2.1. The description and verification of the dataset

At the April 20, 2022, the keywords were set as "oral squamous cell carcinoma," "human," "Expression profiling by array," and "gene symbol," and the datasites were screened. After reviewing the abstract and design of per dataset, GSE41613 was downloaded from the gene expression omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds/?term=). In the GSE41613 dataset, there were 97 OSCC samples, including 41 low grade cases and 56 high grade OSCC tissues. Principal component analysis (PCA) is a statistical method. A set of potentially correlated variables is transformed into a set of linearly uncorrelated variables through orthogonal transformation, and the transformed set of variables is called principal components. In order to verify the dataset GSE41613, the PCA was performed.

2.2. Screening of differently expressed genes (DEGs) related with the OSCC

GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) is a software for differential analysis of expression profile chips based on the GEO database. Using this tool we can compare 2 sets of samples (low grade OSCC and high grade OSCC) in the GSE41613 dataset and obtain DEGs. After the analysis by GEO2R, the DEGs were identified by the criteria of P value ≤.05 and Log_2 (fold change) ≥1 or ≤ -1.

2.3. Protein-protein interaction (PPI) analysis and the identification of hub genes

The Search Tool for the Retrieval of Interacting Genes (STRING) database (https://string-db.org/) is an online database for searching known protein interactions. Currently, a total of 2031 species and 9643763 protein interactions are stored. The above DEGs related with the OSCC were added into the STRING database, and the PPI was constructed.

Cytoscape (version 3.5.1) is a software focused on open source network visualization and analysis. Its core is to provide basic functional layout and query network, and to combine basic data into a visual network. Cytoscape is originated from systems biology and is used to integrate biomolecular interaction networks with high-throughput gene expression data and other molecular state information. Its most powerful function is for large-scale PPIs, and genetic interactions. And the software was used for visualization for the PPI network. Furthermore, the cytoHubba, one plug-in of the Cytoscape, was implied for the identification of the hub gens related with the OSCC. In order to explore the most significant hub genes, the EcCentricity, EPC, and MNC were used, and the VENN diagram was performed. After that, the BottleNeck was used for the verification of the most significant hub genes.

2.4. Functional annotation of the DEGs related with the OSCC by DAVID

Database for Annotation, Visualization and Integrated Discovery (DAVID, (https://david.ncifcrf.gov) is not only a biological database, but also an online analysis software. It can be used for differential gene analysis and pathway enrichment. DAVID could associate genes in the input list with biological annotations. The tool implements functional annotation of genes, presenting in tabular form the annotation of each gene in the input list in the selected database. Enrichment analysis of the DEGs related with the OSCC was made by the DAVID (version 6.7).

2.5. Analysis of the metascape

Metascape integrates multiple authoritative data resources such as genebody, Kyoto Encyclopedia of Genes and Genomes, UniProt and DrugBank, so that it cannot only complete pathway enrichment and biological process (BP) annotation, but also perform gene-related protein network analysis and related drug analysis. In terms of data processing scale, Metascape can process lists containing thousands of genes at one time, and supports simultaneous upload of multiple different gene lists. Therefore, the DEGs related with the OSCC were input in the Metascape to explore the enrichment terms.

2.6. The analysis for the relationship between OSCC and hub genes

The comparative toxicogenomics database (CTD) The CTD database integrates a large number of interaction data between chemical substances, genes, functional phenotypes and diseases,

providing great convenience for the study of disease-related environmental exposure factors and the potential mechanism of action of drugs. The inference score and references of CDSN for the OSCC were calculated by the CTD. In addition, the expression level of hub genes was analyzed by the heatmap. And the relationships among the hub genes were presented. The overall survival analysis of CDSN for the OSCC was made. Finally, the role of hub genes on the pathological stage of OSCC was explored.

2.7. RT-PCR assay

Primer Premier 3.0 was used to design specific primers for CDSN. Primers and probes were synthesized by Fuzhou Boshang Biotechnology Co., Ltd., diluted with RNase-free water to 10 μmol/L, and stored at -20°C for later use. Total RNA was extracted from tissue samples according to the instructions of the RNA extraction kit (Hunan Aikerui Bioengineering Co., Ltd.), and the concentration and purity of RNA were measured with a nucleic acid protein analyzer, and stored at -80 °C for later use. Perform conventional RT-PCR amplification according to the instructions of the one-step RT-PCR kit. The reaction system: One Step Enzyme Mix 1 μL, 2×One-Step Reaction Solution 12.5 μL, upstream and downstream primers CBPV-F/R (10 μmol/L)) 1 μL each, template RNA 2 μL, supplemented with RNase-free water to 25 μL. The reaction conditions were: reverse transcription at 50°C for 30 minute; pre-denaturation at 95°C for 3 minute; denaturation at 95°C for 30 second, annealing at 56°C for 30 second, extension at 72°C for 30 second, 35 cycles; extension at 72°C for 5 minute. The relative expression levels of genes were calculated based on CT values. Take GAPGH was set as a reference.

Primers of CDSN and sequences for PCR analysis were followed:

CDSN-F: GCCGATGACTGAGATAAGG
CDSN-R: GCTGCTGCTGAACTGAAA.

2.8. Immunofluorescence assay

Add 0.01 mol/L PBS with pH 7.4 dropwise to the specimen to be tested, and discard it after 10 min to keep the specimen at a certain humidity. Add appropriately diluted fluorescently labeled antibody solution dropwise to completely cover the specimen, place it in an enamel box with a lid, and keep it for a certain period of time (30 minute). Take out the slides, put them on the slide rack, rinse them with 0.01 mol/L PBS, pH 7.4, and then soak them in 3 cylinders of 0.01 mol/L, pH 7.4 PBS in sequence, 3 to 5 minute in each cylinder, Oscillate from time to time. Remove the slide, absorb excess water with filter paper, but not dry the specimen, add a drop of buffered glycerol, and cover with a coverslip. Immediately observe with a fluorescence microscope.

2.9. Patients and ethics

A total of 200 patients diagnosed with OSCC at Ethics Committee of the Fourth Hospital of Hebei Medical University from March 2015 to June 2020 were selected.

Determine the sample size:

$$n = \frac{Z^2}{d^2}\sigma^2$$

Z is the confidence interval, n is the sample size, d is the sampling error range, and σ is the standard deviation, which is generally 0.5.

Inclusion criteria: 18 to 80 years old, diagnosed with OSCC; Cardiopulmonary function was normal; Normal coagulation.

Exclusion criteria: age < 18 or age > 80 requiring emergency surgery; The patients and their families did not agree to participate in the trial.

The study was approved by Ethics Committee of the Fourth Hospital of Hebei Medical University, and all written informed consent was obtained from the patients.

2.10. The parameter

According to clinical data, patients were classified by gender (male/female), age (≤65/>65), tumor size (≤1 cm/>1 cm), family history of OSCC (no/yes), tumor grade (low/high), CDSN (low/high), tumor stage (early/late), and prognosis (good/poor). Patients' survival time was recorded during follow-up. The tumor grade was classified as grade I, II, III, and IV, with grade I and II as low, grade III and IV as high. The level of CDSN expression was defined according to the mean CDSN expression level of the included patients. Lower than the average level is low expression, vice versa is high expression. The general stage of the tumor is stage I, II, III, and IV, with stage I and II as low, stage III and IV as high. The prognostic classification was based on the recurrence rate and 5-year survival rate.

2.11. Statistical method

The data is expressed as a percentage of the total. Pearson chi-square test and Spearman correlation coefficient were used to analyze the clinical parameters and prognosis of OSCC. Univariate and multivariate Logistic regression analysis were used to calculate the OR values of prognostic variables in OSCC patients. Forward conditional selection was used to build the multivariable model. Any test results reaching a liberal statistical threshold of $P < .2$ for each comparison were then entered multivariable regression model to identify independently predictive factors for OSCC. The risk factors were forced to enter into the multivariate regression model to identify independently predictive factors for OSCC. Then, the overall survival time of OSCC was converted into natural logarithmic equivalent value for statistical analysis. Cox proportional risk regression analysis was used to explore the correlation between survival time and related factors in OSCC patients. The receiver operating characteristic curves were obtained using MedCalc software. Receiver operating characteristic analysis was made to explore the sensitivity and specificity of the expression level of CDSN on the overall survival time of OSCC. And the Kaplan-Meier analysis was performed to explored the effect of CDSN on the overall survival. BP neural network is not only a universal model, but also an error correction function. According to the results obtained by training and the expected results, error analysis is carried out each time, then the weight and threshold value is modified, and the model whose output is consistent with the expected results is obtained step by step. The model finally trained by the support vector machine (SVM) algorithm is determined by some support vectors, which can determine the vector of the final model. Therefore, the BP-neural network and the SVM were constructed by the MATLAB (version: 2017a), and in the model, the expression of CDSN was set as the input value, and the survival time of OSCC was set as the output value. Statistical analyses were performed using SPSS software, version 21.0 (IBM, Armonk, NY). $P < .05$ was considered statistically significant.

3. Results:

3.1. Verification of the GSE41613 and DEGs related with the OSCC

After analysis via the PCA, the samples in the GSE41613 were classified in the PC1 and the value is 68.1% (Fig. 1A). Volcano

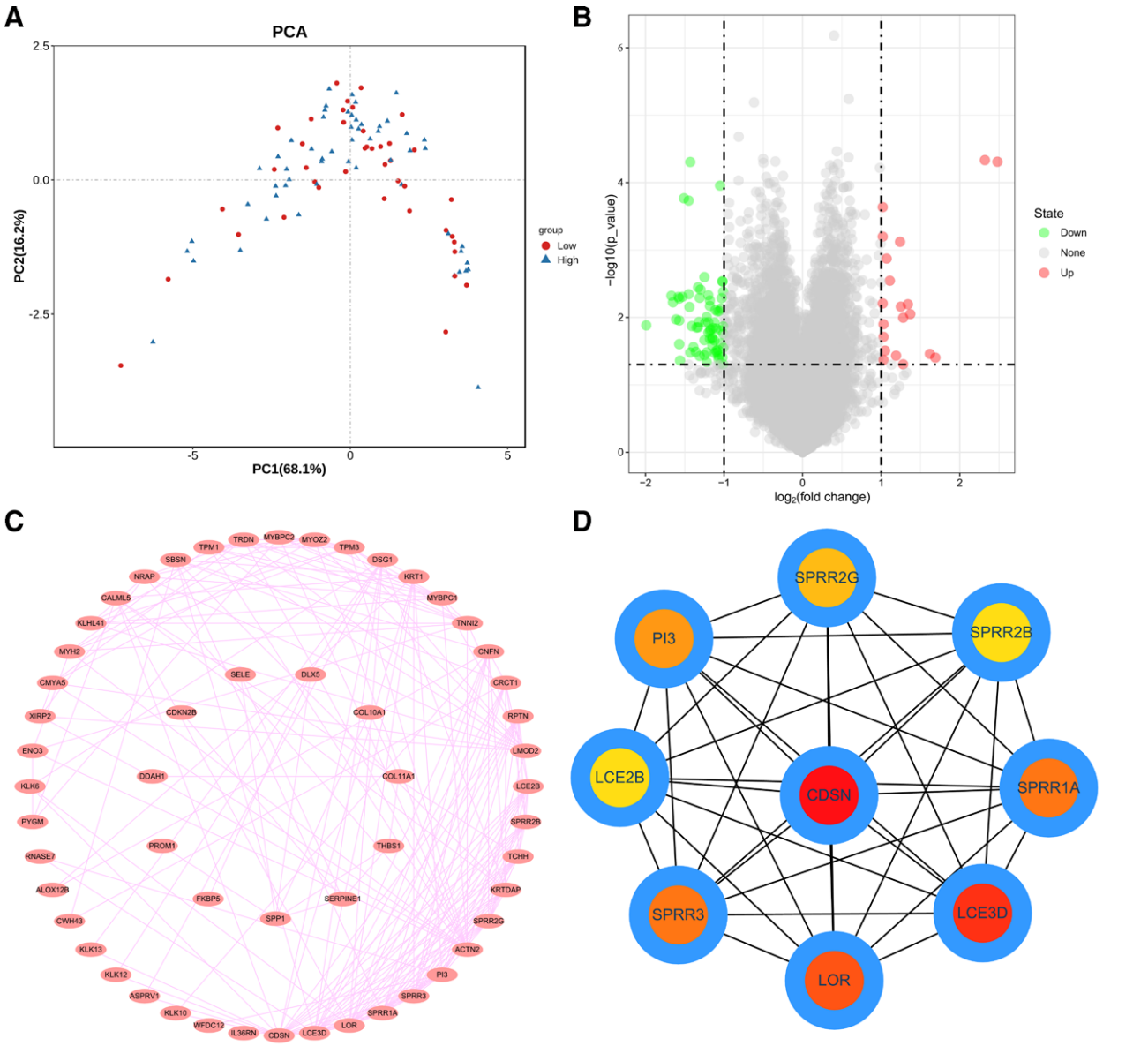

Figure 1. Verification of the GSE41613 and DEGs related with the OSCC. (A) Principal component analysis. (B) Volcano plot manifested that there were plenty of DEGs between the low grade OSCC and high grade OSCC. The green presented the down-regulated DEGs and the red showed the up-regulated DEGs in the high grade OSCC (C) The PPI network of the DEGs. (D) Hub genes by the cytoHubba. DEGs = differently expressed genes, OSCC = oral squamous cell carcinoma, PPI = protein-protein interaction.

plot manifested that there were plenty of DEGs between the low grade OSCC and high grade OSCC, and the green presented the down-regulated DEGs and the red showed the up-regulated DEGs in the high grade OSCC (Fig. 1B). The PPI network was successfully constructed and there was a maze of connections among the DEGs (Fig. 1C). After calculated by the cytoHubba, 9 hub genes were identified, which including SPRR2G, SPRR2B, SPRR1A, LCE3D, LOR, SPRR3, LCE2B, PI3, and CDSN (Fig. 1D).

3.2. CDSN as the most significant hub gene of the OSCC

After calculated by the EcCentricity, 10 hub genes were identified, which including ENO3, PYGM, MYBPC2, CALML5, TCHH, LOR, KRT1, KRTDAP, CDSN, KLK1 (Fig. 2A). Through the EPC analysis, the LCE2B, SPRR2B, LOR, KRTDAP, CRCT1, LCE3D, SPRR1A, SPRR2G, SPRR3, CDSN were identified (Fig. 2B). By the MNC analysis, SPRR3, SPRR2G, SPRR2B, SPRR1A, PI3, LOR, LCE3D, LCE2B, CDSN were screened (Fig. 2C). Through the VENN diagram, the CDSN was identified as the most significant hub gene of the OSCC (Fig. 2D). Finally, after the verification by the BottleNeck, the CDSN was also the most significant hub genes of the OSCC (Fig. 2E).

3.3. Enrichment analysis via the DAVID

In the aspect of BP, the DEGs of the OSCC mainly enriched in the cornification, epidermis development, keratinocyte differentiation, sarcomere organization, cell adhesion, extracellular matrix organization, epithelial cell differentiation (Fig. 3A). In the aspect of cell component, the DEGs of the OSCC mainly enriched in cornified envelope, extracellular region, extracellular space, myofibril, secretory granule, sarcomere, myosin filament, cytosol, sarcoplasmic reticulum, muscle thin filament tropomyosin, platelet alpha granule lumen, stress fiber, extracellular exosome (Fig. 3B). In the aspect of molecular function, the DEGs of

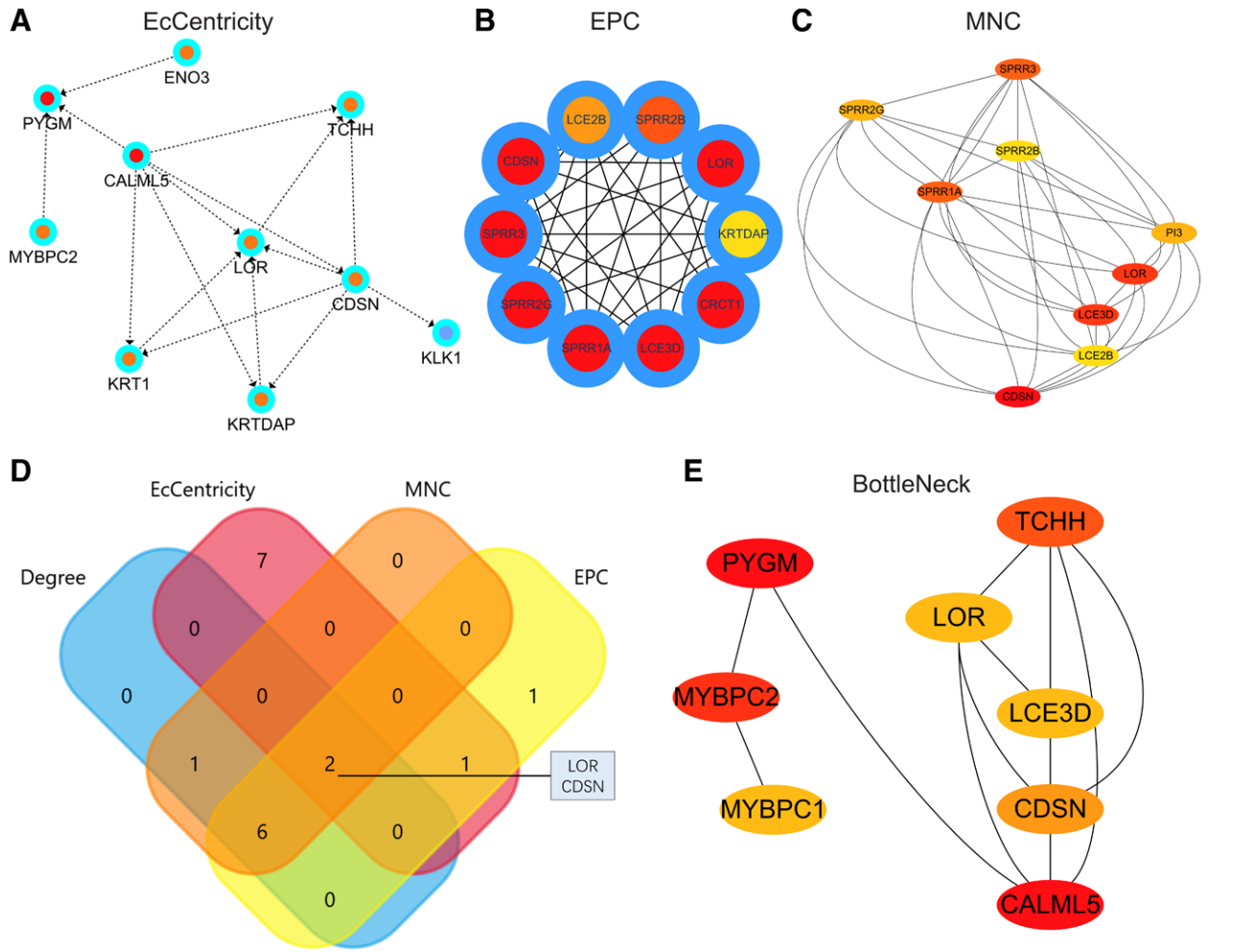

Figure 2. CDSN as the most significant hub gene of the OSCC. (A) Hub genes by the EcCentricity. (B) Hub genes by EPC analysis. (C) Hub genes by MNC analysis. (D) Through the VENN diagram, the CDSN was identified as the most significant hub gene of the OSCC. (E) After the verification by the BottleNeck, the CDSN was also the most significant hub genes of the OSCC. CDSN = corneodesmosin, MNC = major histocompatibility complex, OSCC = oral squamous cell carcinoma.

the OSCC mainly enriched in the actin binding, actin filament binding, structural constituent of muscle, structural constituent of epidermis, extracellular matrix binding, serine-type endopeptidase activity, NADP-retinol dehydrogenase activity, serine-type endopeptidase inhibitor activity, extracellular matrix structural constituent (Fig. 3C). In the aspect of Kyoto encyclopedia of genes and genomes, the DEGs of the OSCC mainly enriched in the apelin signaling pathway, cell adhesion molecules (Fig. 3D).

3.4. The analysis by the metascape

Through the metascape, the DEGs were mainly enriched in the formation of the cornified envelope, positive regulation of epithelial cell differentiation, TGF-beta receptor signaling, positive regulation of cation channel activity, cell-cell junction organization, retina homeostasis, regulation of cell adhesion, sensory organ morphogenesis ($P < .05$) (Fig. 3E–G). The significant module in the DEGs related with the OSCC was identified by the Metascape, and the CDSN was also included in the significant module, which verified CDSN as the most significant hub genes (Fig. 4A). Summary of enrichment analysis in COVID showed that the DEGs were mainly enriched in the vanderheiden, lamers intestinal-organoid expansion, proteome stukalov (Fig. 4B). Summary of enrichment analysis in cell type signatures showed that the DEGs were mainly enriched in "BUSSLINGER ESOPHAGEAL LATE SUPRABASAL CELLS," "DESCARTES MAIN FETAL SQUAMOUS EPITHELIAL CELLS" (Fig. 4C). Summary of enrichment analysis in DisGeNET showed that the DEGs were mainly enriched in recurrent tumor (Fig. 4D). Summary of enrichment analysis in PaGenBase showed that the DEGs were mainly enriched in tongue, tonsil (Fig. 5A). Summary of enrichment analysis in Transcription Factor Targets was presented in the Figure 5B and Supplemental Digital Content (Table S1, http://links.lww.com/MD/H438-S6, http://links.lww.com/MD/H438).

3.5. The role of CDSN on the OSCC

By the CTD analysis, there was strong relationship between the CDSN and mouth neoplasms, head and neck neoplasms, squamous cell carcinoma of head and neck (Fig. 6A). Compared with low grade OSCC, the CDSN was down-regulated in the high grade OSCC (Fig. 6B). The heatmap manifested that there were positive correlations among the hub genes ($P < .05$) (Fig. 6C). The OSCC patients with low expression level of CDSN have poor overall survival compared with the high expression level of CDSN (HR = 0.75, 95%CI: 0.57-0.98, $P = .036$) (Fig. 6D). The lower the expression of CDSN was, the higher the clinicopathological stage of OSCC was ($P = .0238$, F value = 3.18) (Fig. 6E).

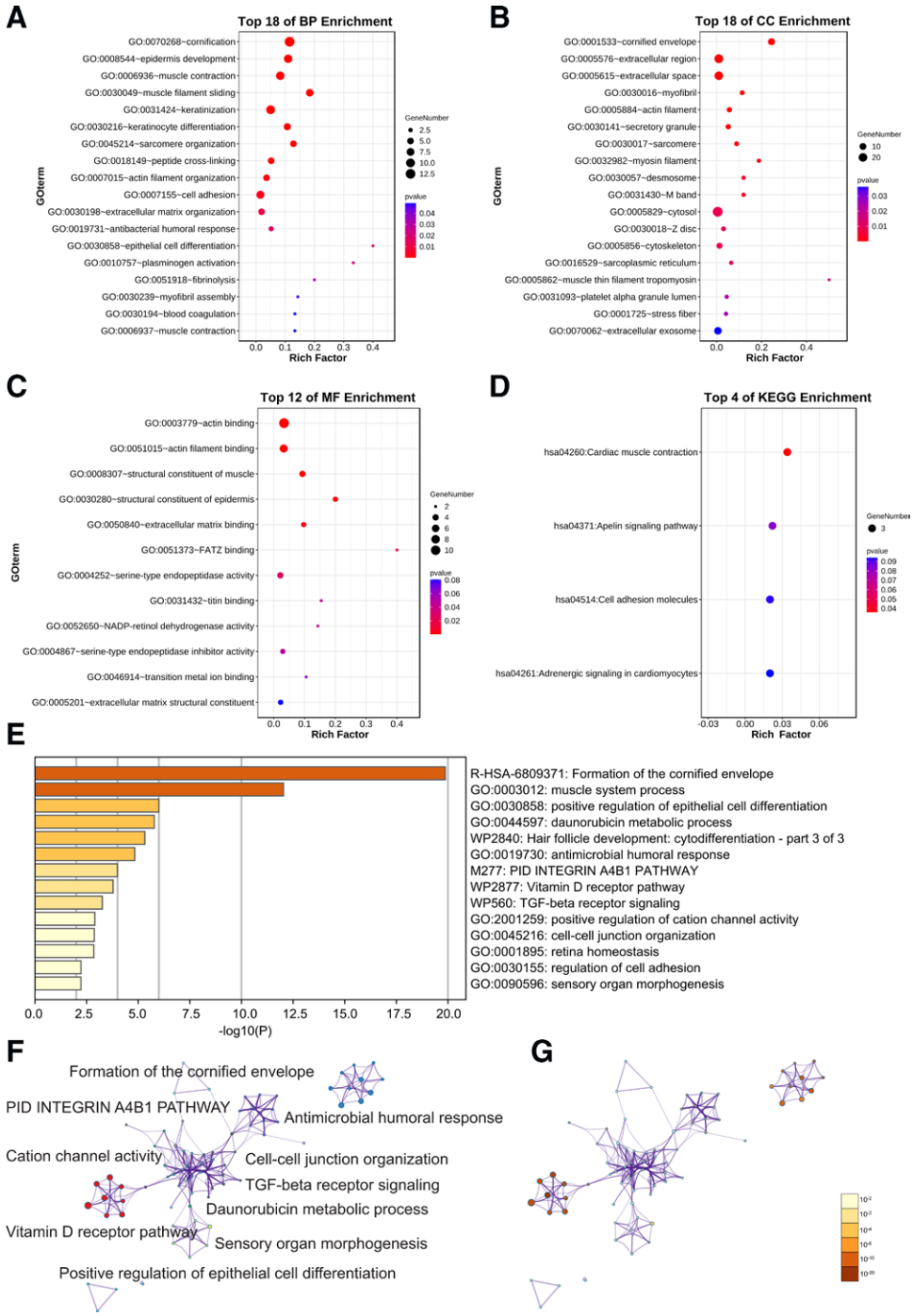

Figure 3. Enrichment analysis via the DAVID, bar graph of enriched terms across input gene lists and network of enriched terms by the Metascape. (A) Biological process (BP). (B) Cell component (CC). (C) Molecular function (MF). (D) KEGG. (E) Bar graph of enriched terms across input gene lists, colored by *P* values. (F) Network of enriched terms: colored by cluster ID, where nodes that share the same cluster ID are typically close to each other. (G) Network of enriched terms: colored by *P* value, where terms containing more genes tend to have a more significant *P* value. DAVID = database for annotation, visualization and integrated discovery, KEGG = Kyoto encyclopedia of genes and genomes.

3.6. Verification of the effect of CDSN on the OSCC by the clinical samples

The expression of CDSN in the low grade OSCC was higher than the high grade OSCC via the RT-PCR ($P < .05$, Fig. 7A). The OSCC patients with low expression level of CDSN have poor overall survival compared with the high expression level of CDSN (HR = 0.096, $P < .001$) (Fig. 7B). Through the immunofluorescence assay, compared with the high grade OSCC, the expression of CDSN was up-regulated in the low grade OSCC ($P < .05$) (Fig. 7C). The expression of CDSN might be molecular target for the diagnosis of the OSCC, and the sensitivity and specificity were high (AUC = 0.764, $P < .05$) (Fig. 7D). The joint effect of all factors (CDSN, age, family history, sex, tumor grade, tumor size, tumor stage) on the diagnosing the OSCC was strong (AUC = 0.814, $P < .05$) (Fig. 7E). The respective role of factors on the diagnosis of OSCC was presented in the Figure 7F.

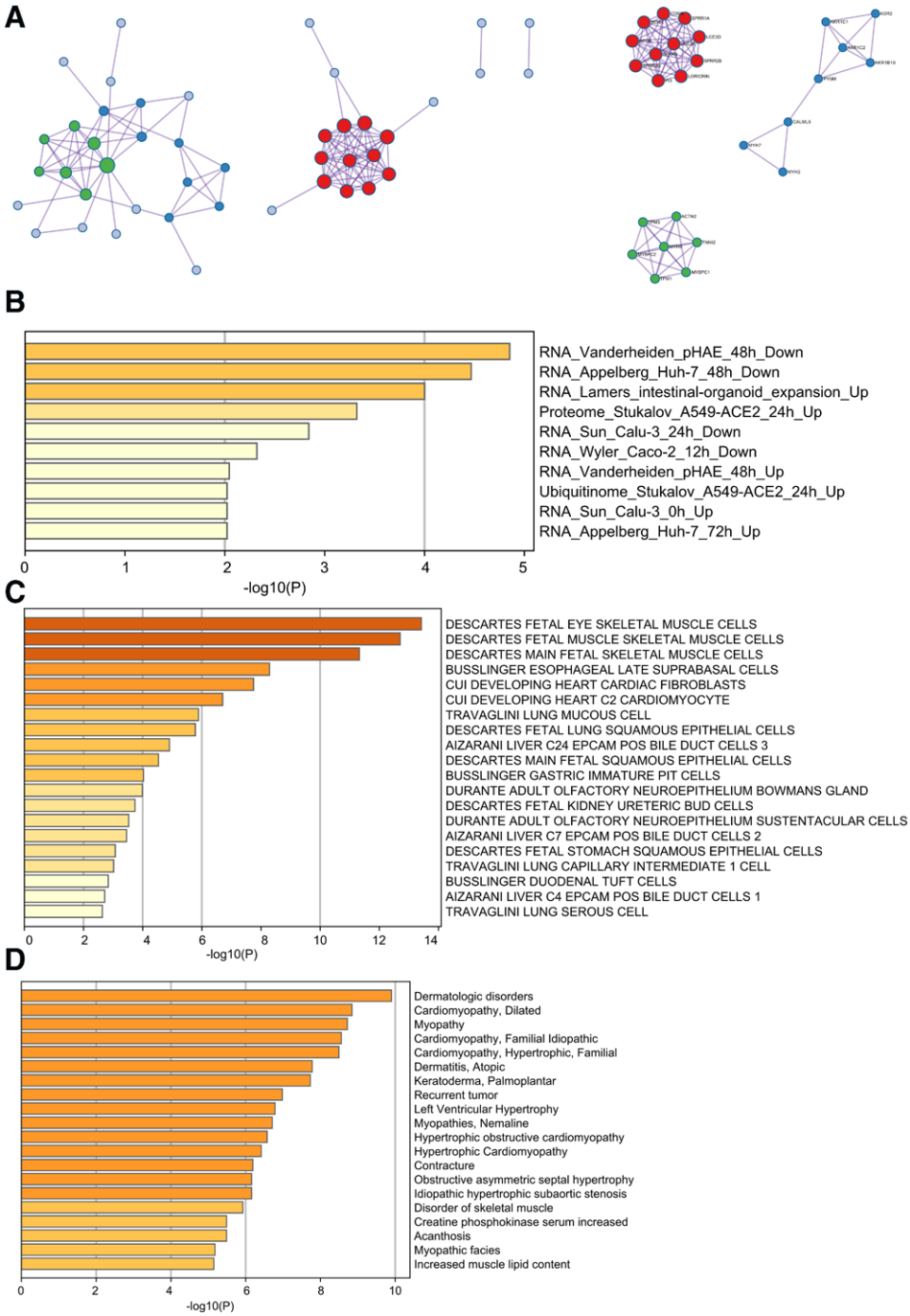

Figure 4. The analysis by the Metascape. (A) Protein-protein interaction network and MCODE components identified in the gene lists. (B) Summary of enrichment analysis in COVID. (C) Summary of enrichment analysis in cell type signatures. (D) Summary of enrichment analysis in DisGeNET.

3.7. Strong predictive value of CDSN for the OSCC survival time

Through the BP-neural network, best training performance is 0.0026893 at epoch 3000, and the training relevance (R) is 0.99413 between CDSN and the OSCC survival time. Furthermore, the percentage errors of the BP-neural network were small (Fig. 8A). In addition, the strong predictive value of CDSN for the OSCC survival time was also verified by the SVM model ($y = 0.6041\times + 20.5610$, $R = 0.8559$) (Fig. 8B).

3.8. Pearson's Chi-square test was used to analyze the correlation between CDSN expression and related factors of OSCC

Pearson's Chi-square test was used to summarize the relationship between CDSN expression level and related factors of OSCC. Prognosis ($P < .001$) was significantly correlated with CDSN expression level. However, Sex ($P = .571$), Age ($P = .121$), Tumor size ($P = .261$), Family History ($P = .325$), Tumor Grade ($P = .199$), and Tumor Stage ($P = .524$) were not significantly correlated with CDSN expression (Table 1).

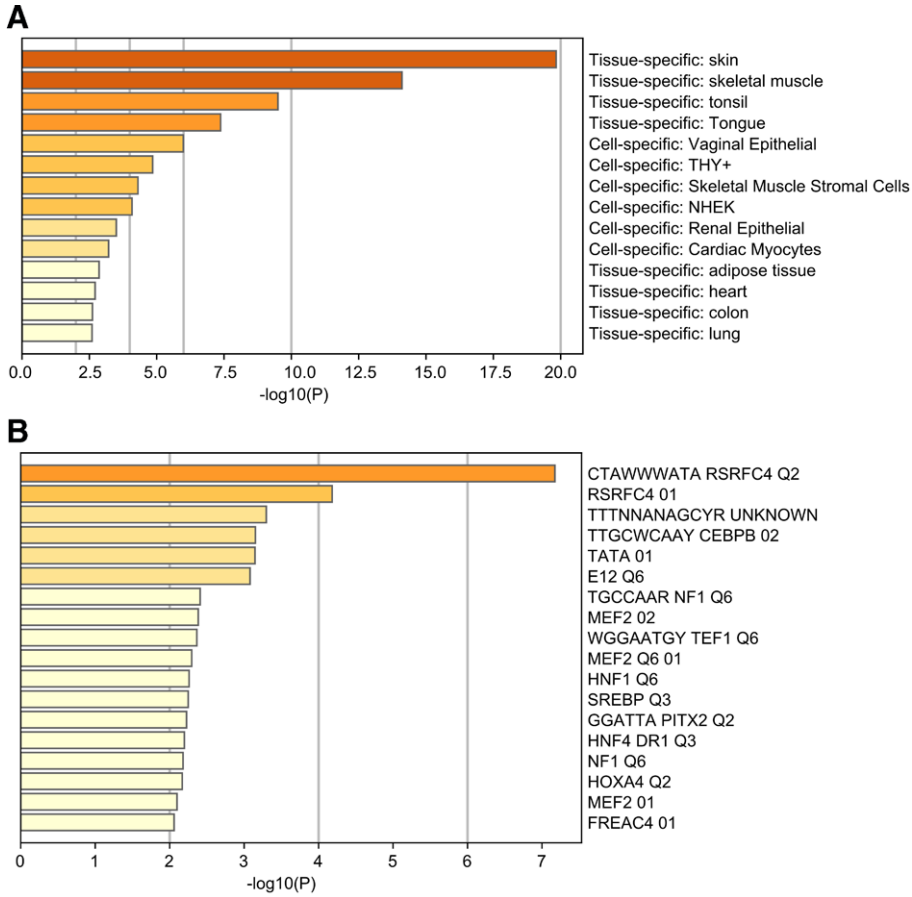

Figure 5. Summary of enrichment analysis in PaGenBase and Transcription Factor Targets. (A) PaGenBase. (B) Transcription factor targets.

3.9. Spearman correlation coefficient was used to analyze the correlation between CDSN expression level and related factors of OSCC

Further analysis of Spearman correlation coefficient showed that CDSN expression level was significantly correlated with prognosis (ρ = -0.528, $P < .001$). However, Gender (rho = 0.040, $P = .573$), Age (rho = 0.110, $P = .122$) and Tumor Size (rho = 0.029, $P = .679$), and Family History (rho = 0.070, $P = .328$), Tumor Grade (ρ = -0.091, $P = .201$), Tumor Stage (ρ = -0.045, $P = .527$) had no significant correlation with CDSN expression (Table 2).

3.10. Univariate logistic regression analysis of prognosis and related factors of OSCC

Binary Logistic regression was used to determine the relationship between OSCC related parameters and CDSN expression level, odds ratio (OR), and 95% confidence interval (95% CI). Table 3 describes the OR and 95%CI of the subjects at the univariate Logistic regression level, and the results show that the prognosis of subjects with low CDSN expression level is significantly worse than that of subjects with high CDSN expression level (OR = 0.095, 95%CI: 0.049-0.183, $P < .001$). However, Gender (OR = 0.852,95%CI: 0.488-1.485, $P = .571$), Age (OR = 0.637,95%CI: 0.359-1.128, $P = .122$), Tumor Size (OR = 0.889,95%CI: 0.510-1.549, $P = .677$), Family History (OR = 0.752,95%CI: 0.427-1.327, $P = .326$), Tumor Grade (OR = 0.694,95%CI: 0.397-1.213, $P = .200$), Tumor Stage (OR = 0.834,95%CI: 0.477-1.458, $P = .524$) had no significant correlation with CDSN expression level (Table 3).

3.11. Multivariate logistic regression analysis of the correlation between OSCC and CDSN expression level

Multivariate Logistic regression was used to describe the OR and 95%CI of the subjects at the multivariate level. The results showed that the prognosis of patients with low CDSN expression was significantly worse than that of patients with high CDSN expression (OR = 0.096, 95%CI: 0.049-0.189, $P < .001$). While Gender (OR = 0.827, 95%CI: 0.244-1.610, $P = .575$), Age (OR = 0.940, 95%CI: 0.472-1.871, $P = .860$), Tumor Size (OR= 1.070, 95%CI: 0.553-2.067, $P = .842$), Family History (OR = 0.794, 95%CI: 0.407-1.551, $P = .500$), Tumor Grade (OR = 0.973, 95%CI: 0.499-1.896, $P = .935$) and Tumor Stage (OR = 0.935, 95%CI: 0.482-1.813, $P = .842$) had no significant correlation with CDSN expression (Table 4).

3.12. Multivariate logistic regression analysis of prognosis and associated characteristics in patients with OSCC

Multivariate Logistic regression was used to describe the OR and 95%CI of the subjects at the multivariate level. The prognosis of patients with low CDSN expression level was significantly worse than that with high CDSN expression level (OR = 0.096, 95%CI: 0.049-0.189, $P < .001$). While Gender (OR = 0.855, 95%CI: 0.431-1.695, $P = .653$), Age (OR = 2.000, 95%CI: 0.996-4.016, $P = .051$), Tumor Size (OR = 1.404, 95%CI: 0.717-2.747, $P = .322$), Family History (OR = 1.120, 95%CI: 0.562-2.231, $P = .747$), Tumor Grade (OR = 1.782, 95%CI: 0.907 -- 3.502, $P = .094$) and tumor stage (OR = 1.144,95%CI: 0.577-2.267, $P = .700$) had no significant correlation with prognosis (Table 5).

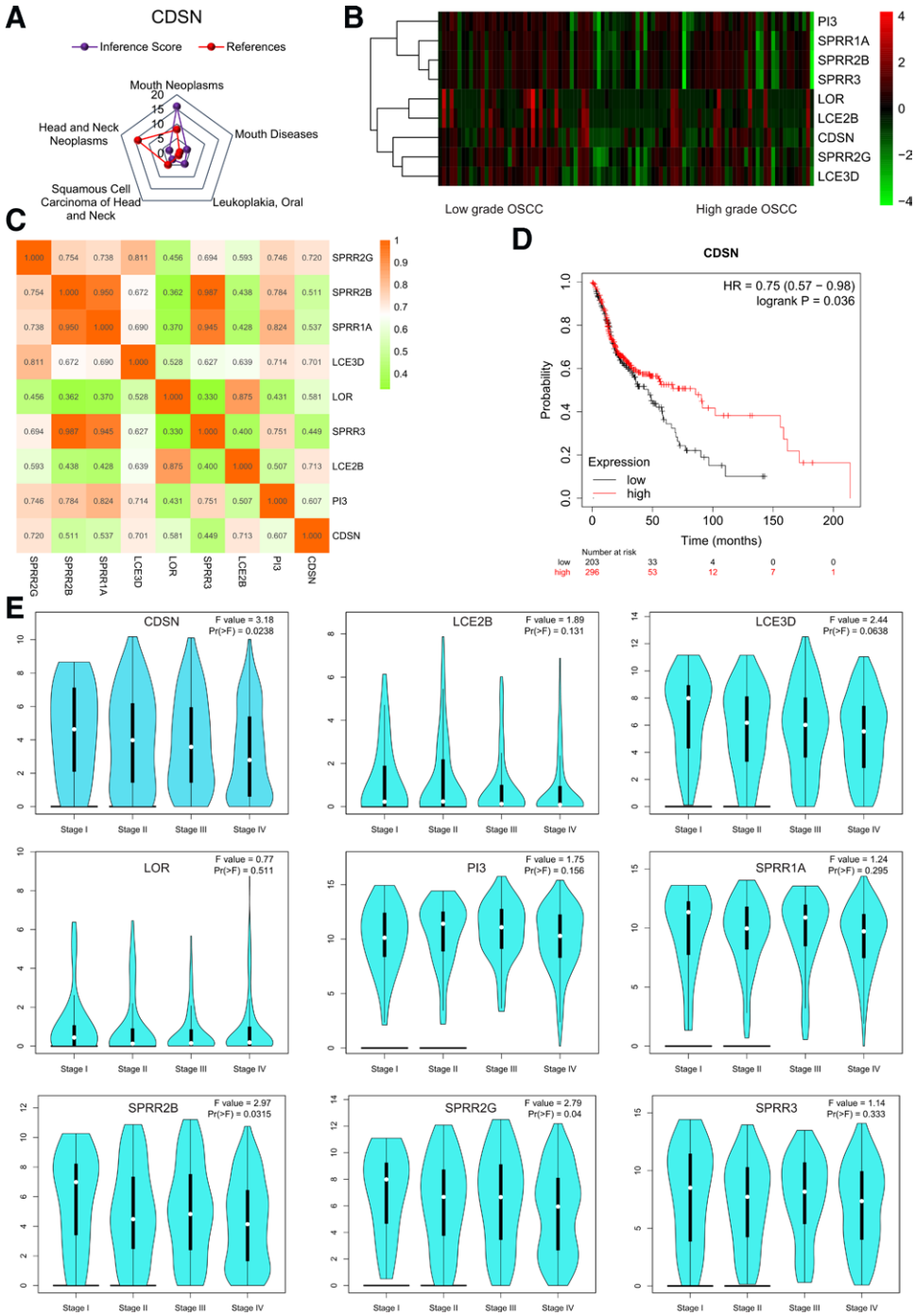

Figure 6. The role of CDSN on the OSCC. (A) By the CTD analysis, there was strong relationship between the CDSN and mouth neoplasms. (B) Compared with low grade OSCC, the CDSN was down-regulated in the high grade OSCC. (C) The heatmap manifested that there were positive correlations among the hub genes ($P < .05$). (D) The OSCC patients with low expression level of CDSN have poor overall survival compared with the high expression level of CDSN (HR = 0.75, 95%CI: 0.57-0.98, $P = .036$). (E) The lower the expression of CDSN was, the higher the clinicopathological stage of OSCC was ($P = .0238$, F value = 3.18). CDSN = corneodesmosin, OSCC = oral squamous cell carcinoma.

3.13. Cox regression analysis of OSCC patients survival time proportional risk

In order to effectively control the influence of confounding factors, all factors were included in the Cox regression model. Cox proportional regression analysis showed that the survival time of OSCC patients with low CDSN expression level was significantly lower than that of patients with high CDSN expression level (HR = 0.588, 95%CI: 0.420-0.823, $P = .002$) (Fig. 1). However, Gender (HR = 1.100, 95%CI: 0.787-1.536, $P = .578$), Age (HR = 1.401, 95%CI: 0.994-1.974, $P = .054$), and Tumor Size (HR = 1.048 95%CI: 0.765-1.435, $P = .770$), Family History (HR = 0.953, 95%CI: 0.690-1.318, $P = .772$), Tumor Grade (HR = 1.502, 95%CI: 1.090-2.070, $P = .013$) and Tumor Stage (HR = 1.265, 95%CI: 0.918-1.742, $P = .151$) were not significantly correlated with survival time (Table 6).

4. Discussion:

Pearson's Chi-square test showed that prognosis of patients with OSCC ($P < .001$) was significantly correlated with CDSN

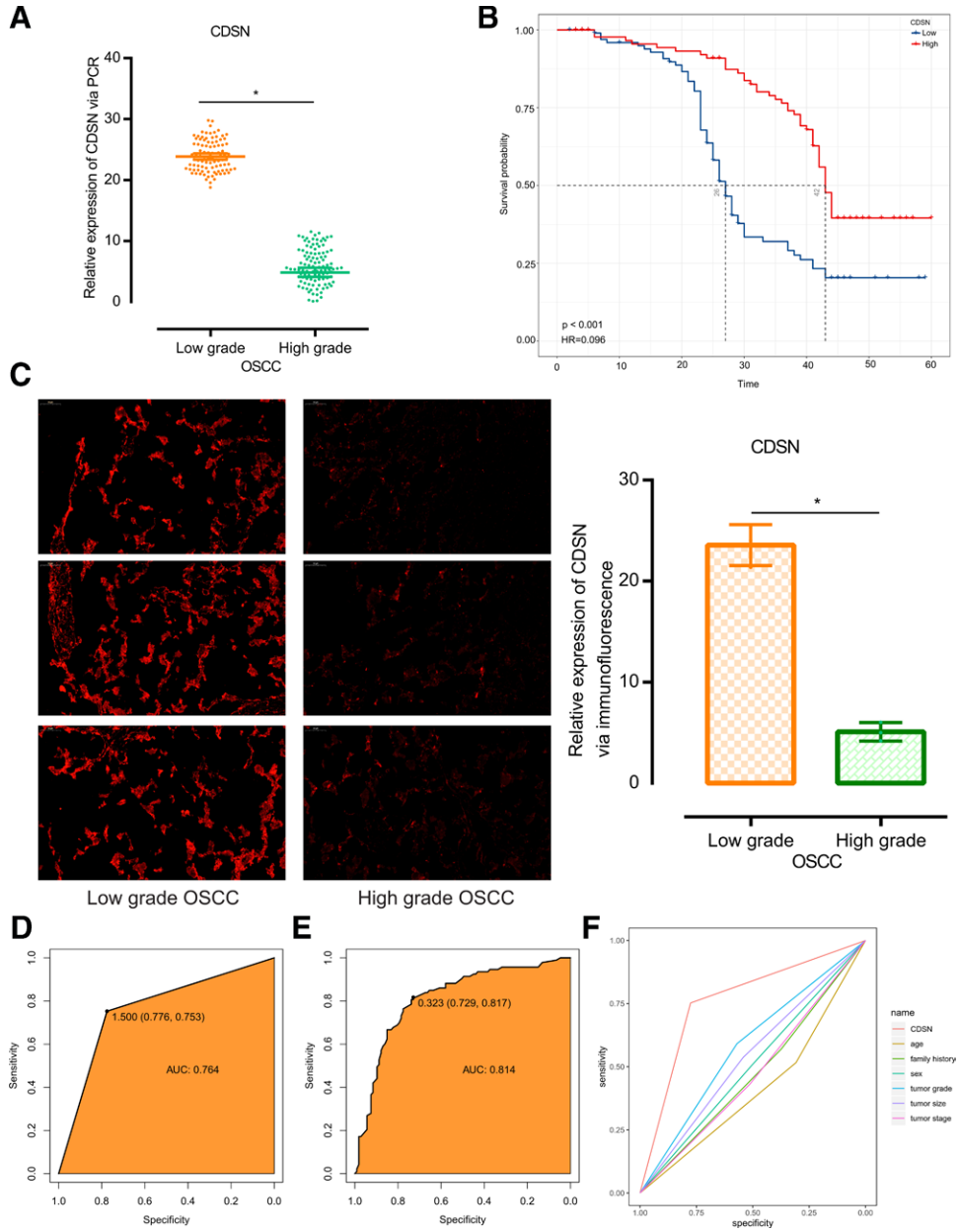

Figure 7. Verification of the effect of CDSN on the OSCC by the clinical samples. (A) The expression of CDSN in the low grade OSCC was higher than the high grade OSCC via the RT-PCR. (B) The OSCC patients with low expression level of CDSN have poor overall survival compared with the high expression level of CDSN (HR = 0.096, $P < .001$). (C) Through the immunofluorescence assay, compared with the high grade OSCC, the expression of CDSN was up-regulated in the low grade OSCC ($P < .05$). (D) The expression of CDSN might be molecular target for the diagnosis of the OSCC, and the sensitivity and specificity were high (AUC = 0.764, $P < .05$). (E) The joint effect of all factors (CDSN, age, family history, sex, tumor grade, tumor size, tumor stage) on the diagnosing the OSCC was strong (AUC = 0.814, $P < .05$). (F) The respective role of factors on the diagnosis of OSCC. CDSN = corneodesmosin, OSCC = oral squamous cell carcinoma.

expression. Spearman correlation coefficient analysis showed that CDSN expression level was significantly correlated with prognosis (ρ = -0.528, $P < .001$). Univariate Logistic regression analysis showed that prognosis (OR = 0.095, 95%CI: 0.049-0.183, $P < .001$) was significantly correlated with CDSN expression. Multivariate Logistic regression analysis showed that prognosis (OR = 0.096, 95%CI: 0.049-0.189, $P < .001$) was significantly correlated with CDSN expression level. Cox regression analysis showed that CDSN (HR = 0.588, 95%CI: 0.420-0.823, $P = .002$) was significantly correlated with survival time.

OSCC is the most common malignant tumor of the head and neck.[13] OSCC is a multifactorial disease, which is mainly related to chronic stimulation of tobacco and alcohol. Chronic inflammation, viral infection (human papillomavirus) and genetic predisposition are also risk factors for oral tumors.[14] OSCC is a malignant tumor originating from epithelial tissue. Generally speaking, the malignant degree is relatively low and the growth rate is relatively slow. Surgical treatment is the main treatment, and postoperative chemotherapy can be assisted.[15] In recent years, oral microbiome has been reported to play an important role in the development of OSCC. Oral microbiome induces inflammatory responses through the production of cytokines and chemokines, thereby enhancing the proliferation and survival of tumor cells.[16] Studies have shown that the histopathological examination of OSCC is often associated with significant inflammatory reactions. Other studies have shown that there is a significant pro-inflammatory immune environment in OSCC.[17] In the tumor microenvironment, the production of a variety of

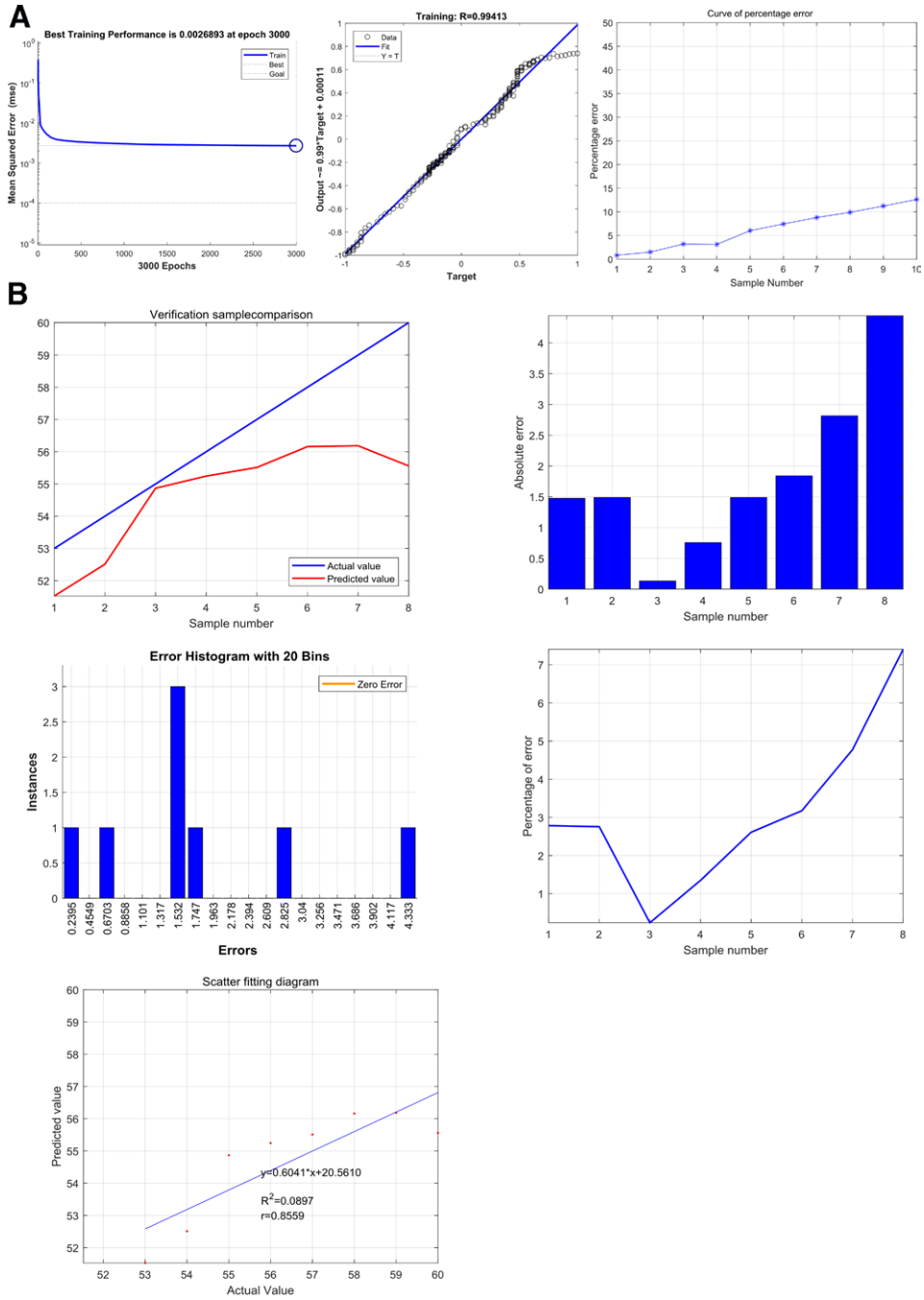

Figure 8. Strong predictive value of CDSN for the OSCC survival time. (A) BP-neural network, best training performance is 0.0026893 at epoch 3000, and the training relevance (R) is 0.99413 between CDSN and the OSCC survival time, the percentage errors of the BP-neural network were small (B) The strong predictive value of CDSN for the OSCC survival time was also verified by the SVM model (y = 0.6041× + 20.5610, *R* = 0.8559). CDSN = corneodesmosin, OSCC = oral squamous cell carcinoma.

pro-inflammatory cytokines can promote the survival and proliferation of tumor cells.[18,19]

CDSN is located in the core of desmosomes in the cuticle.[20] Some experimental evidence supports the role of CDSN in keratinocyte cohesion.[21] Therefore, we speculated that when CDSN is highly expressed, a strong barrier can be constructed for oral mucosa to prevent the invasion of external pathogens, thus reducing the damage of oral mucosa and preventing the occurrence and development of OSCC.

CDSN is also closely related to inflammation.[22] Related studies have shown that CDSN mutations can cause inflammatory peeling syndrome. CDSN is the only PSORS1 gene specifically expressed in the final differentiated keratinocytes,[23] CDSN is also an important component of keratinocyte desmosomes, hair follicle epidermal keratinocytes and inner root sheath desmosomes,[24] and plays an important role in maintaining epidermal integrity.[25] Scalp simple oligotrichosis is associated with nonsense mutations in CDSN encoding keratin.[26] Keratinocytes are hydrolyzed gradually during maturation. This treatment is a prerequisite for peeling. Changes in the skin's proteolytic balance can lead to inflammation, which can lead to clinical symptoms.[27] Thus, low expression of CDSN induces oral inflammation, which is usually related to the occurrence and development of cancer. It is one of the initiation processes for cells to enter the tumor microenvironment through specific cytokines called chemokines.[28] Inflammation involves the

Table 1

Relevant characteristics of patients with oral squamous cell carcinoma.

Characteristics			CDSN Low	CDSN High	*P*
Sex	Male	100	51 (25.5%)	49 (24.5%)	.571
	Female	100	55 (27.5%)	45 (22.5%)	
Age	≤65	78	36 (18.0%)	42 (21.0%)	.121
	>65	122	70 (35.0%)	52 (26.0%)	
Tumor size	≤ 1cm	99	51 (25.5%)	48 (24.0%)	.261
	>1cm	101	55 (27.5%)	46 (23.0%)	
Family history	No	80	39 (19.5%)	41 (20.5%)	.325
	Yes	120	67 (33.5%)	53 (26.5%)	
Tumor grade	Low	101	49 (24.5%)	52 (26.0%)	.199
	High	99	57 (28.5%)	42 (21.0%)	
Tumor stage	Early	108	55 (27.5%)	53 (26.5%)	.524
	Late	92	51(25.5%)	41 (20.5%)	
Prognosis	Good	93	23 (11.5%)	70 (35.0%)	<.001*
	Poor	107	83 (41.5%)	24 (12.0%)	

CDSN = corneodesmosin.
Pearson chi-square test
*$P < .05$.

Table 2

Relationship between characteristics of oral squamous cell carcinoma and CDSN expression level.

Characteristics	CDSN ρ	*P*
Sex	-0.040	.573
Age	-0.110	.122
Tumor size	-0.029	.679
Family history	-0.070	.328
Tumor grade	-0.091	.201
Tumor stage	-0.045	.527
Prognosis	-0.528	<.001*

CDSN = corneodesmosin.
Spearman correlation analysis
*$P < .05$

Table 3

Association of parameters associated with oral squamous cell carcinoma and CDSN expression level based on univariate Logistic regression analysis.

Characteristics			CDSN OR	CDSN 95% CI	*P*
Sex	Male	100	1		.571
	Female	100	0.852	0.488-1.485	
Age	≤65	78	1		.122
	>65	122	0.637	0.359-1.128	
Tumor size	≤ 1cm	99	1		.677
	>1cm	101	0.889	0.510-1.549	
Family history	No	80	1		.326
	Yes	120	0.752	0.427-1.327	
Tumor grade	Low	101	1		.200
	High	99	0.694	0.397-1.213	
Tumor stage	Early	108	1		.524
	Late	92	0.834	0.477-1.458	
Prognosis	Good	93	1		<.001*
	Poor	107	0.095	0.049-0.183	

95% CI = 95% confidence interval, CDSN = corneodesmosin, OR = odds ratio.
*$P < .05$.

Table 4.

Correlation between CDSN expression level and related parameters of oral squamous cell carcinoma by multivariate Logistic regression analysis.

Characteristics	CDSN OR	CDSN 95%CI	*P*
Sex	0.827	0.424-1.610	.575
Age	0.940	0.472-1.871	.860
Tumor size	1.070	0.553-2.067	.842
Family history	0.794	0.407-1.551	.500
Tumor grade	0.973	0.499-1.896	.935
Tumor stage	0.935	0.482-1.813	.842
Prognosis	0.096	0.049-0.189	<.001*

95% CI = 95% confidence interval, CDSN = corneodesmosin, OR = odds ratio.
*$P < .05$.

Table 5

Multivariate Logistic regression analysis of the prognostic characteristics and influence of oral squamous cell carcinoma.

Characteristics	Prognosis OR	Prognosis 95%CI	*P*
Sex	0.855	0.431-1.695	.653
Age	2.000	0.996-4.016	.051
Tumor size	1.404	0.717-2.747	.322
Family history	1.120	0.562-2.231	.747
Tumor grade	1.782	0.907-3.502	.094
CDSN	0.096	0.049-0.189	<.001*
Tumor stage	1.144	0.577-2.267	.700

95% CI = 95% confidence interval, OR = odds ratio.
*$P < .05$.

Table 6

Influence of correlation characteristics on survival time of patients based on Cox regression analysis.

Characteristics	Survival time HR	Survival time 95% CI	*P*
Sex	1.100	0.787-1.536	.578
Age	1.401	0.994-1.974	.054
Tumor size	1.048	0.765-1.435	.770
Family history	0.953	0.690-1.318	.772
Tumor grade	1.502	1.090-2.070	.013
CDSN	0.588	0.420-0.823	.002*
Tumor stage	1.265	0.918-1.742	.151

95% CI = 95% confidence interval, HR = hazard ratio.
*$P < .05$.

interaction of various immune cells, inflammatory cells, chemokines, cytokines and pro-inflammatory mediators. It also plays a decisive role in the initiation, promotion, malignant transformation, invasion, metastasis and other stages of tumor development.[29] Many cancers are caused by the site of infection, chronic irritation and inflammation. Cell proliferation alone does not cause cancer, but continuous cell proliferation in an environment rich in inflammatory cells, growth factors, activating matrices, and DNA damage promoters certainly enhances and/or promotes the risk of cancer. During injury-related tissue damage, cell proliferation is enhanced, while tissue regeneration occurs, and proliferation and inflammation subside after the aggressor is removed or repair is completed. Instead, proliferating cells undergo DNA damage or mutagenic attack and

continue to prolific in a microenvironment rich in inflammatory cells and growth/survival factors that support their growth. In a sense, tumors are like wounds that do not heal. It is clear that inflammatory cells have a powerful influence on tumor development.[30] These cells are powerful tumor promoters early in the tumor process, creating an attractive environment for tumor growth, promoting genomic instability and promoting angiogenesis. In later tumorigenesis, tumor cells also metastasize the inflammatory mechanism, and although the inflammatory response is also supposed to be anti-tumor, the inflammatory response in cancer patients is often defective. Inflammatory cells and the chemokines and cytokines they produce affect the entire tumor organ and regulate the growth, migration, and differentiation of all cell types in the tumor microenvironment, including tumor cells, fibroblasts, and endothelial cells.[31,32] Therefore, the expression level of CDSN may play a role in the occurrence and development of OSCC.

4.1. Limitations

There are some shortcomings in this study. Although clinical data have been examined and analyzed, the molecular mechanism of CDSN expression on the prognosis and survival time of patients with OSCC has not been verified in animal models. Therefore, future studies should focus on animal experiments to explore the molecular pathway and mechanism of CDSN in OSCC. The sample size of this study is relatively limited. In order to further verify the results, clinical data need to be collected, summarized and analyzed continuously in the process of clinical work.

5. Conclusion

The expression level of CDSN was significantly correlated with the prognosis and survival time of OSCC. Low expression of CDSN is associated with poor prognosis and short survival time in OSCC. CDSN as a potential prognostic target of OSCC provides a new direction for the molecular mechanism of its occurrence and development.

Author contributions

Conceptualization: Yong-Bin Di.
Data curation: Yang Bao.
Formal analysis: Jie Guo.
Investigation: Wei Liu.
Software: Su-Xin Zhang.
Supervision: Guan-Hua Zhang.
Writing – original draft: Tian-Ke Li.
Writing – review & editing: Tian-Ke Li.

References

[1] Reshmi SC, Gollin SM. Chromosomal instability in oral cancer cells. J Dent Res. 2005;84:107–17.
[2] Di Pardo BJ, Bronson NW, Diggs BS, et al. The global burden of esophageal cancer: a disability-adjusted life-year approach. World J Surg. 2016;40:395–401.
[3] Chi AC, Day TA, Neville BW. Oral cavity and oropharyngeal squamous cell carcinoma--an update. CA Cancer J Clin. 2015;65:401–21.
[4] Wang Y, Fan H, Zheng L. Biological information analysis of differentially expressed genes in oral squamous cell carcinoma tissues in GEO database. J BUON. 2018;23:1662–70.
[5] Zhao X, Sun S, Zeng X, et al. Expression profiles analysis identifies a novel three-mRNA signature to predict overall survival in oral squamous cell carcinoma. Am J Cancer Res. 2018;8:450–61.
[6] Rivera C, Oliveira AK, Costa R, et al. Prognostic biomarkers in oral squamous cell carcinoma: a systematic review. Oral Oncol. 2017;72:38–47.
[7] Vitório JG, Duarte-Andrade FF, Dos Santos Fontes Pereira T, et al. Metabolic landscape of oral squamous cell carcinoma. Metabolomics. 2020;16:105.
[8] Telem DF, Israeli S, Sarig O, et al. Inflammatory peeling skin syndrome caused a novel mutation in CDSN. Arch Dermatol Res. 2012;304:251–5.
[9] Wu Y, Wang B, Liu JL, et al. Association of -619C/T polymorphism in CDSN gene and psoriasis risk: a meta-analysis. Genet Mol Res. 2011;10:3632–40.
[10] van der Velden J, van Geel M, Engelhart JJ, et al. Mutations in the CDSN gene cause peeling skin disease and hypotrichosis simplex of the scalp. J Dermatol. 2020;47:3–7.
[11] Sun H, Xu B, Meng Q, et al. PSORS1C1/CDSN is associated with ankylosing spondylitis. Joint Bone Spine. 2014;81:268–72.
[12] Mallet A, Kypriotou M, George K, et al. Identification of the first nonsense CDSN mutation with expression of a truncated protein causing peeling skin syndrome type B. Br J Dermatol. 2013;169:1322–5.
[13] Sun LP, Xu K, Cui J, et al. Cancer-associated fibroblast-derived exosomal miR-382-5p promotes the migration and invasion of oral squamous cell carcinoma. Oncol Rep. 2019;42:1319–28.
[14] St John MA. Inflammatory mediators drive metastasis and drug resistance in head and neck squamous cell carcinoma. Laryngoscope. 2015;125(Suppl 3):S1–11.
[15] Choi S, Myers JN. Molecular pathogenesis of oral squamous cell carcinoma: implications for therapy. J Dent Res. 2008;87:14–32.
[16] Rai AK, Panda M, Das AK, et al. Dysbiosis of salivary microbiome and cytokines influence oral squamous cell carcinoma through inflammation. Arch Microbiol. 2021;203:137–52.
[17] Laliberté C, Ng N, Eymael D, et al. Characterization of oral squamous cell carcinoma associated inflammation: a Pilot Study. Front Oral Health. 2021;2:740469.
[18] Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. Cell. 2010;140:883–99.
[19] Colotta F, Allavena P, Sica A, et al. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. Carcinogenesis. 2009;30:1073–81.
[20] Serre G, Mils V, Haftek M, et al. Identification of late differentiation antigens of human cornified epithelia, expressed in re-organized desmosomes and bound to cross-linked envelope. J Invest Dermatol. 1991;97:1061–72.
[21] Lundström A, Serre G, Haftek M, et al. Evidence for a role of corneodesmosin, a protein which may serve to modify desmosomes during cornification, in stratum corneum cell cohesion and desquamation. Arch Dermatol Res. 1994;286:369–75.
[22] Ishida-Yamamoto A, Furio L, Igawa S, et al. Inflammatory peeling skin syndrome caused by homozygous genomic deletion in the PSORS1 region encompassing the CDSN gene. Exp Dermatol. 2014;23:60–3.
[23] Jonca N, Caubet C, Leclerc EA, et al. Protease sensitivity of corneodesmosin variants encoded by the six more common CDSN haplotypes. J Invest Dermatol. 2011;131:1381–4.
[24] Siva Subramaniam N, Morgan E, Bottomley S, et al. Predictive mutational bioinformatic analysis of variation in the skin and wool associated corneodesmosin (CDSN) gene in sheep. Anim Sci J. 2012;83:386–93.
[25] Wada T, Matsuda Y, Muraoka M, et al. Alu-mediated large deletion of the CDSN gene as a cause of peeling skin disease. Clin Genet. 2014;86:383–6.
[26] Levy-Nissenbaum E, Betz RC, Frydman M, et al. Hypotrichosis simplex of the scalp is associated with nonsense mutations in CDSN encoding corneodesmosin. Nat Genet. 2003;34:151–3.
[27] Williams MR, Costa SK, Zaramela LS, et al. Quorum sensing between bacterial species on the skin protects against epidermal injury in atopic dermatitis. Sci Transl Med. 2019;11:eaat8329.
[28] Nagarsheth N, Wicha MS, Zou W. Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. Nat Rev Immunol. 2017;17:559–72.
[29] Singh N, Baby D, Rajguru JP, et al. Inflammation and cancer. Ann Afr Med. 2019;18:121–6.
[30] Mantovani A, Allavena P, Sica A, et al. Cancer-related inflammation. Nature. 2008;454:436–44.
[31] Iyengar NM, Gucalp A, Dannenberg AJ, et al. Obesity and cancer mechanisms: tumor microenvironment and inflammation. J Clin Oncol. 2016;34:4270–6.
[32] Landskron G, De la Fuente M, Thuwajit P, et al. Chronic inflammation and cytokines in the tumor microenvironment. J Immunol Res. 2014;2014:149185.