

COL11A1 as a potential prognostic target for oral squamous cell carcinoma

Short title: Prognostic target for oral squamous cell carcinoma

Yong-bin Di

The First Hospital of Hebei Medical University

Yang Bao

The Fourth Hospital of Hebei Medical University

Jie Guo

The Fourth Hospital of Hebei Medical University

Wei Liu

The Fourth Hospital of Hebei Medical University

Su-xin Zhang

The Fourth Hospital of Hebei Medical University

Guan-hua Zhang

The Fourth Hospital of Hebei Medical University

Tian-ke Li (✉ litianb112@163.com)

The Fourth Hospital of Hebei Medical University

Research Article

Keywords: Oral squamous cell carcinoma, COL11A1, Prognostic, Targets, Survival time

Posted Date: May 26th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1671839/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Oral squamous cell carcinoma (OSCC) is a malignant tumor of epithelial origin. Explore correlation between COL11A1 and prognosis of OSCC, provide a new direction for treatment and prognosis of patients.

Methods

Bioinformatics was used to screen and identify role of COL11A1 on OSCC. 200 patients with oral squamous cell carcinoma were recruited. Clinical and follow-up data were recorded and COL11A1 expression levels were tested. Pearson chi-square test and Spearman correlation coefficient were used to analyze relationship between prognosis and related parameters in patients with oral squamous cell carcinoma. Univariate and multivariate Logistic regression, univariate and multivariate Cox proportional risk regression were used for further analysis, survival curve was drawn.

Results

Through bioinformatics analysis, OSCC patients with higher expression of COL11A1 have poor overall survival compare with OSCC patients with lower expression of COL11A1 (HR = 1.32, P = 0.047). Pearson chi-square test showed that age (P = 0.011), tumor grade (P = 0.023), COL11A1 (P < 0.001) was significantly correlated with prognosis of oral squamous cell carcinoma. Spearman correlation coefficient analysis showed that age ($\rho = 0.179$, P = 0.011), tumor grade ($\rho = 0.161$, P = 0.023), COL11A1 ($\rho = 0.561$, P < 0.001) were significantly correlated with prognosis of patients with oral squamous cell carcinoma. Univariate Logistic regression analysis showed age (OR = 2.102, 95%CI: 1.180–3.746, P = 0.012), tumor grade (OR = 1.919, 95%CI: 1.093–3.372, P = 0.023) and COL11A1 (OR = 12.775, 95%CI: 6.509–25.071, P < 0.001). Multivariate Logistic regression analysis showed that COL11A1 (OR = 12.066, 95%CI: 6.042–24.096, P < 0.001) was significantly associated with prognosis of patients with oral squamous cell carcinoma. Univariate Cox regression analysis showed that age (HR = 1.592, 95%CI: 1.150–2.205, P = 0.005), tumor grade (HR = 1.460, 95%CI: 1.067–1.999, P = 0.018) and COL11A1 (HR = 1.848, 95%CI: 1.340–2.548, P < 0.001) were significantly correlated with survival time of oral squamous cell carcinoma patients. Multivariate Cox regression analysis showed that tumor grade (HR = 1.466, 95%CI: 1.064–2.020, P = 0.019) and COL11A1 (HR = 1.645, 95%CI: 1.164–2.325, P = 0.005) were significantly correlated with survival time of oral squamous cell carcinoma patients.

Conclusion

COL11A1 is significantly correlated with occurrence of oral squamous cell carcinoma.

Introduction

Oral squamous cell carcinoma (OSCC) is a malignant tumor occurring in the oral cavity, mainly squamous cells, which is the most malignant tumor of the head and neck [1]. OSCC has good activity, rich blood circulation, high rate of lymphatic metastasis and poor prognosis. In addition to local proliferation and erosion of surrounding tissue function, lymphatic metastasis in the neck drainage area often occurs in oral squamous cell carcinoma [2]. Therefore, the treatment of oral squamous cell carcinoma usually requires the treatment of neck lymphatic tissue in addition to the expanded resection of the local primary lesion. Generally speaking, the malignant degree of oral squamous cell carcinoma is relatively low, the growth rate is relatively slow, and the treatment is mainly surgical treatment, postoperative adjuvant chemotherapy can be used [3]. In the early stage, oral squamous cell carcinoma is mostly manifested as ulceration, erosion and hyperplasia. In the later stage, it gradually increases into vegetable pattern masses or crater like masses, forming infiltrating masses. Metastasis may occur later, often in lymph nodes and lungs [4]. If oral squamous cell carcinoma can be found at an early stage, it can be completely cured in the clinical sense by surgical expansion of the tumor, cervical lymph node dissection, combined with systemic chemotherapy and local radiotherapy after surgery. At present, clinical treatment methods have become more and more advanced and mature, and more and more treatment methods have been applied in clinical practice, so the survival rate and cure rate of oral squamous cell carcinoma are also increasing [5].

Collagen Type XI Alpha 1 (COL11A1) is a protein-coding gene and a secondary Collagen, belonging to the fibrocollagen gene family, whose main physiological function is to regulate the diameter of main Collagen fibers [6]. Collagen is the framework of the extracellular matrix, which is found in all organs and tissues in the body [7]. COL11A1 gene is located in chromosome 1p21 region and contains 68 exons. At present, at least 4 isomers have been found [8]. At present, it is generally believed that the generation process of COL11A1 is: The intracellular COL11A1 gene encodes the synthesis of pre-COL11A1 mRNA, which is translated into pre-COL11A1 polypeptide, pre-COL11A1 polypeptide chain and pre-COL11A1 in sequence on the endoplasmic reticulum, and then secreted into mature COL11A1 by enzymolysis. COL11A1 encodes one of the two alpha chains of type XI collagen, a mild form of fibrous collagen. Collagen type is a heterotrimer, but the third α chain is posttranslatively modified α 1 type II chain. Single nucleotide polymorphisms in this gene have also been associated with susceptibility to lumbar disc herniation and have been identified as multiple transcriptional variants [9]. The gene body (GO) annotations associated with the COL11A1 gene include extracellular matrix structural composition and extracellular matrix binding. COL11A1 may play an important role in fiber formation by controlling the lateral growth of collagen fibrils, and may also affect the occurrence and development of cancer by regulating several downstream genes involved in biological processes such as cell proliferation, apoptosis, metastasis and signal transduction[10]. However, the relationship between COL11A1 and oral squamous cell carcinoma remains unclear.

This study hypothesized that during the occurrence and development of oral squamous cell carcinoma, the higher the content of COL11A1, the worse the prognosis and shorter the survival time of patients with

oral squamous cell carcinoma. Based on the above hypothesis, we recruited 200 patients with oral squamous cell carcinoma. These results may reveal COL11A1 as a potential molecular target of oral squamous cell carcinoma and provide new ideas for its molecular mechanism.

Methods

Dataset of the OSCC and difference analysis

GSE41613 was downloaded from the GENE EXPRESSION OMNIBUS (GEO) database. And the GSE41613 consisted of ninety-seven OSCC individuals, including fifty-six OSCC samples with high grade and forty-one OSCC samples with low grade. Intra-group correlation analysis was used for the repeatability test of the samples. GEO2R was implied to screen the differently expressed genes (DEGs) between the OSCC with high grade and OSCC with low grade. The DEGs between the two groups were screened by criteria of $\text{Log}_2(\text{fold change}) \geq 1$ or ≤ -1 and $P\text{-value} \leq 0.05$. Finally, COL11A1 was identified with the highest fold change ($P < 0.05$).

Analysis of COL11A1 on the OSCC

The TCGA dataset was used to verify the expression of COL11A1 in the OSCC. Comparative Toxicogenomics Database (CTD) database was implied to explore the relationship among the expression of COL11A1, mouth neoplasms, head and neck neoplasms. The effect of COL11A1 on the pathological stage OSCC was analyzed by using the TCGA database. The heatmap was drew for presenting the expression of DEGs between the OSCC with low grade and high grade. And the overall survival analysis of COL11A1 on the OSCC was made by the Kaplan Meier plotter.

Enrichment analysis for the DEGs related with the OSCC

Gene Set Enrichment Analysis (GSEA) is used to evaluate the distribution trend of genes in a pre-defined gene set in the gene table ranked by phenotype correlation, thereby judging their contribution to phenotype. The input data consists of two parts, one is the gene set with known functions (can be GO annotation, MsigDB annotation or other gene set definitions that conform to the format), the other is the expression matrix (it can also be a sorted list). The software will sort the genes according to their correlation with the phenotype (which can be understood as the change in expression value) from large to small, and then judge whether the genes under each annotation in the gene set are enriched in the upper part of the gene table after the phenotype correlation is sorted or the lower part, so as to judge the effect of cooperative changes of genes in this gene set on phenotypic changes. The GSEA was performed to made enrichment analysis for the DEGs related with the OSCC.

Patients and ethics

A total of 200 patients diagnosed with oral squamous cell carcinoma in Fourth Hospital of Hebei Medical University Hospital from March 2015 to June 2020 were selected.

Inclusion criteria: 18-80 years old, diagnosed with oral squamous cell carcinoma, Cardiopulmonary function was normal, Normal coagulation.

Exclusion criteria: patients requiring emergency surgery, the patients and their families did not agree to participate in the study.

This study was approved by the Ethics Committee of Fourth Hospital of Hebei Medical University Hospital, and all patients signed informed consent.

The parameter

According to clinical data, patients were classified by Gender (Male/Female), Age (≤ 65 / > 65), Tumor size (≤ 1 cm/ > 1 cm), Family history (No/Yes), Tumor grade (Low/High), COL11A1 (Low/High), Tumor stage (Low/High), and Prognosis (Good/Poor).

Detection of blood related parameters

Venous blood samples from patients were immediately sent for examination to detect the expression level of COL11A1.

Immunofluorescence assay

Paraffin sections were dewaxed. The tissue sections were placed in a retrieval box filled with EDTA antigen retrieval buffer (PH8.0) for antigen retrieval in a microwave oven. After natural cooling, the slides were washed in PBS (pH 7.4) with shaking on a destaining shaker. After the sections were slightly dried, use a histochemical pen to draw circles around the tissue (to prevent the antibody from flowing away), spin dry PBS, add BSA dropwise, and seal for 30 minutes. Gently shake off the blocking solution, drop the primary antibody prepared in a certain proportion of PBS on the slices, and incubate the slices in a wet box at 4°C overnight. After the slices were slightly dried, the secondary antibody of the corresponding species of the primary antibody was added dropwise in the circle to cover the tissue, and incubated at room temperature for 50 min in the dark. DAPI counterstained nuclei: the slides were placed in PBS (pH 7.4) and washed 3 times with shaking on a destaining shaker, 5 min each time. After the sections were slightly dried, DAPI staining solution was added dropwise to the circle, and incubated at room temperature for 10 min in the dark. Add autofluorescence quencher to the circle for 5 min, and rinse with running water for 10 min. After drying, the sections were mounted with anti-fluorescence quenching mounting medium. Sections were observed under a fluorescence microscope and images were collected.

RT-qPCR assay

Take the homogenization tube, and add 1 ml of RNA extraction solution, and place it on ice to pre-cool. Take approximately 20 mg of tissue and add to a homogenization tube. The homogenizer was thoroughly ground until no visible tissue clumps were observed. Place the centrifuge tube on the ultra-clean table and blow for 3 minutes, add 15 μ l RNA lysis solution to dissolve the RNA, and incubate at

55°C for 5 minutes. RNA concentration and purity were detected by Nanodrop 2000. The main steps included the followed procedures. DNA denaturation (90°C-96°C): The double-stranded DNA template breaks hydrogen bonds under the action of heat to form single-stranded DNA. Annealing (renaturation) (40°C-65°C): the system temperature decreases, and the primer binds to the DNA template , forming a partial double strand. Extension (68°C-75°C): under the action of Taq enzyme (the best activity at around 72°C), using dNTP as raw material, extend from the 5' end → 3' end of the primer to synthesize DNA strand complementary to the template. The DNA content doubles with each cycle of denaturation, annealing, and extension. The $\Delta\Delta$ CT method was used to calculate the relative expression of COL11A1.

Statistical methods

Receiver operating characteristic (ROC) analysis was made to explore the sensitivity and specificity of the expression level of COL11A1 on the overall survival time of OSCC. And the Kaplan-Meier analysis was performed to explored the effect of COL11A1 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 COL11A1 was set as the input value, and the survival time of OSCC was set as the output value.

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 oral squamous cell carcinoma. Univariate and multivariate Logistic regression analysis were used to calculate the OR values of prognostic variables in oral squamous cell carcinoma patients. Univariate and multivariate Cox proportional risk regression analysis were used to explore the potential factors of survival time.

All statistical analyses were performed using SPSS software, version 21.0 (IBM, Armonk, NY, USA). $P < 0.05$ was considered statistically significant.

Results

Repeatability test of the samples and analysis of the COL11A1

In the GSE41613, there were strong relationships among all the samples, which manifested that the repeatability of the data is good (Figure 1A). Compared with the control group, the expression of COL11A1 in the tumor group was higher ($P < 0.05$, Figure 1B). There existed high inference score of COL11A1 in the mouth neoplasms, head and neck neoplasms, squamous cell carcinoma of head and neck through the CTD database (Figure 1C). The higher expression of COL11A1 was, the more severe pathological stage of OSCC was (Figure 1D). Based on the different expression of DEGs, the tumor grade of OSCC could be classified (Figure 1E). The OSCC patients with higher expression of COL11A1 have the

poor overall survival compare with the OSCC patients with lower expression of COL11A1 (HR=1.32, P=0.047) (Figure 1F).

GO enrichment analysis for the OSCC

Butterfly plot could present that the high grade OSCC and low grade OSCC could be separated based on the score (signal 2 noise) (Figure 2A). The DEGs related with the OSCC were mainly enriched in the phosphatidylinositol binding, which was higher in the high grade OSCC (Figure 2B). Furthermore, the DEGs were also enriched in the epidermal cell differentiation (Figure 2C), establishment or maintenance of epithelial cell apical basal polarity (Figure 2D), which were down-regulated in the high grade OSCC. The NES vs. Significance could be presented in the Figure 2E, which showed the relationship between FDR q-value (or P-value) and the NES (Figure 2E). Ranked gene list correlation profile showed that area bias to high =49.3% and zero crossing at rank 11548 (55.9%) (Figure 2F).

Random ES distributions of the phosphatidylinositol binding, epidermal cell differentiation, establishment or maintenance of epithelial cell apical basal polarity could classify the two groups of low grade and high grade OSCC (Figure 3A-C). In addition, there were plenty of DEGs based on the expression, and the expression heat map could classify the OSCC (Figure 3D-E).

KEGG enrichment analysis for the OSCC

Through the GSEA, the DEGs were mainly enriched in the P53 signaling pathway, alpha linolenic acid metabolism, amino sugar and nucleotide sugar metabolism, arachidonic acid metabolism, beta alanine metabolism, cysteine and methionine metabolism (Figure 4A-F). Random ES distributions of these KEGG pathways could classify the two groups of low grade and high grade OSCC (Figure 5A-F).

Verification of the expression of COL11A1 with the OSCC clinical samples

Through the immunofluorescence assay, the expression of COL11A1 was higher in the high grade OSCC than the low grade OSCC (P<0.05) (Figure 6A). Furthermore, the result also be verified by the PCR assay. Compared with the low grade OSCC, the mRNA expression of COL11A1 was higher in the high grade OSCC (Figure 6B).

Strong predictive value of COL11A1 for the OSCC survival time

Through the BP-neural network, best training performance is 0.0027083 at epoch 3000, and the training relevance (R) is 0.99409 between COL11A1 and the OSCC survival time. Furthermore, the percentage errors of the BP-neural network were small (Figure 6C-E). In addition, the strong predictive value of COL11A1 for the OSCC survival time was also verified by the SVM model ($y=0.1834x+42.4856$, R=0.6809) (Figure 7).

High sensitivity and specificity of COL11A1 for diagnosing the OSCC

The expression of COL11A1 might be molecular target for the diagnosis of the OSCC, and the sensitivity and specificity were high (AUC=0.781, $P < 0.05$). However, the sensitivity and specificity of tumor grade (AUC=0.581) and tumor size (AUC=0.540) for diagnosing the OSCC were low. The joint effect of all factors on the diagnosing the OSCC was strong (AUC=0.815, $P < 0.05$) (Figure 8).

Overall survival analysis for the OSCC

Based on the clinical samples, the OSCC patients with higher expression of COL11A1 have the poor overall survival compare with the OSCC patients with lower expression of COL11A1 (HR=1.645, $P = 0.005$). However, the other factors were not related with the survival time of OSCC ($P > 0.05$, Figure 9)

Pearson's Chi-square test was used to analyze the correlation between prognosis and related factors of oral squamous cell carcinoma

Pearson's Chi-square test was used to summarize the relationship between the prognosis of oral squamous cell carcinoma and related clinical factors. Age ($P = 0.011$), tumor grade ($P = 0.023$), COL11A1 ($P < 0.001$) was significantly correlated with the prognosis of oral squamous cell carcinoma. However, gender ($P = 0.887$), tumor size ($P = 0.261$), family history ($P = 0.418$), and tumor stage ($P = 0.429$) were not significantly correlated with the prognosis of oral squamous cell carcinoma. (Table 1).

Spearman correlation coefficient was used to analyze the correlation between prognosis and related factors of oral squamous cell carcinoma

Further analysis of Spearman correlation coefficient showed that prognosis of oral squamous cell carcinoma was correlated with age ($\rho = 0.179$, $P = 0.011$), tumor grade ($\rho = 0.161$, $P = 0.023$), COL11A1 ($\rho = 0.561$, $P < 0.001$) significantly correlated. However, gender ($\rho = 0.010$, $P = 0.888$) and tumor size ($\rho = 0.079$, $P = 0.263$), and family history ($\rho = 0.057$, $P = 0.420$), tumor staging ($\rho = 0.056$, $P = 0.432$) had no significant correlation with the prognosis of oral squamous cell carcinoma. (Table 2).

Univariate Logistic regression analysis of prognosis and related factors of oral squamous cell carcinoma

Logistic regression was used to determine the relationship between relevant parameters and prognosis, odds ratio (OR), and 95% confidence interval (95% CI) for oral squamous cell carcinoma. Table 3 describes the OR and 95%CI of the subjects at the univariate Logistic regression level, and the results show that age (OR = 2.102, 95%CI: 1.180-3.746, $P = 0.012$), tumor grade (OR = 1.919, 95%CI: 1.093-3.372, $P = 0.023$) and COL11A1 (OR = 12.775, 95%CI: 6.509-25.071, $P < 0.001$) was significantly correlated with the prognosis of oral squamous cell carcinoma. However, gender (OR = 1.041, 95%CI: 0.597-1.815, $P = 0.887$), tumor size (OR = 1.376, 95%CI: 0.788-2.404, $P = 0.261$), and family history (OR = 1.264, 95%CI: 0.717-2.229, $P = 0.418$), tumor stage (OR = 1.253, 95%CI: 0.716-2.190, $P = 0.429$) had no significant correlation with the prognosis of oral squamous cell carcinoma. (Table 3).

Multivariate Logistic regression analysis of prognostic factors of oral squamous cell carcinoma

Multivariate Logistic regression was used to describe the OR and 95%CI of the subjects at the multivariate level. COL11A1 (OR = 12.066, 95%CI: 6.042-24.096, $P < 0.001$) was significantly correlated with the prognosis of oral squamous cell carcinoma. However, gender (OR= 0.847, 95%CI: 0.422-1.699, $P = 0.640$), age (OR= 1.774, 95%CI: 0.874-3.603, $P = 0.113$), tumor size (OR= 1.465, 95%CI: 0.738-2.906, $P = 0.275$), family history (OR = 1.112, 95%CI: 0.552-2.240, $P = 0.766$), tumor grade (OR = 1.561, 95%CI: 0.786-3.098, $P = 0.203$) and tumor stage (OR = 1.009, 95%CI: 0.503-2.022, $P = 0.981$) were not significantly correlated with the prognosis of oral squamous cell carcinoma. (Table 4).

Univariate Cox regression analysis

Table 5 shows hazard ratios (HRs) and 95% confidence intervals (95%CI) for prognosis of oral squamous cell carcinoma. Age (HR=1.592, 95%CI: 1.150-2.205, $P = 0.005$), tumor grade (HR=1.460, 95%CI: 1.067-1.999, $P = 0.018$) and COL11A1 (HR=1.848, 95%CI: 1.340-2.548, $P < 0.001$) were significantly associated with patient survival time. However, gender (HR = 1.275, 95%CI: 0.932-1.743, $P = 0.129$) tumor size (HR = 1.066, 95%CI: 0.781-1.454, $P = 0.687$), family history (HR = 0.929, 95%CI: 0.675-1.278, $P = 0.650$), and tumor stage (HR = 1.252, 95%CI: 0.915-1.714, $P = 0.160$) had no significant correlation with patient survival time. (Table 5).

Multivariate Cox regression analysis

In order to effectively control the influence of confounding factors, all factors were included into the multivariate Cox regression model. Multivariate Cox proportional regression analysis showed that tumor grade (HR = 1.466, 95%CI: 1.064-2.020, $P = 0.019$) and COL11A1 (HR = 1.645, 95%CI: 1.164-2.325, $P = 0.005$) were significantly correlated with patient survival time. However, gender (HR = 1.097, 95%CI: 0.785-1.535, $P = 0.588$), age (HR = 1.379, 95%CI: 0.976-1.949, $P = 0.068$), tumor size (HR = 1.068, 95%CI: 0.779-1.464, $P = 0.682$), family history (HR = 0.959, 95%CI: 0.694-1.326, $P = 0.801$), and tumor stage (HR = 1.242, 95%CI: 0.902-1.711, $P = 0.184$) had no significant correlation with patient survival time. (Table 6)

Discussion

Pearson's Chi-square test showed that age ($P = 0.011$), tumor grade ($P = 0.023$), COL11A1 ($P < 0.001$) was significantly correlated with the prognosis of oral squamous cell carcinoma. Spearman correlation coefficient analysis showed age ($\rho = 0.179$, $P = 0.011$), tumor grade ($\rho = 0.161$, $P = 0.023$), COL11A1 ($\rho = 0.561$, $P < 0.001$) was significantly correlated with the prognosis of oral squamous cell carcinoma. Univariate Logistic regression analysis showed that age (OR = 2.102, 95%CI: 1.180–3.746, $P = 0.012$), tumor grade (OR = 1.919, 95%CI: 1.093–3.372, $P = 0.023$) and COL11A1 (OR = 12.775, 95%CI: 6.509–25.071, $P < 0.001$) was significantly correlated with the prognosis of oral squamous cell carcinoma. In addition, multivariate Logistic regression analysis showed that COL11A1 (OR = 12.066, 95%CI: 6.042–24.096, $P < 0.001$) was significantly correlated with the prognosis of oral squamous cell carcinoma. Univariate Cox regression analysis showed that age (HR = 1.592, 95%CI: 1.150–2.205, $P = 0.005$), tumor grade (HR = 1.460, 95%CI: 1.067–1.999, $P = 0.018$) and COL11A1 (HR = 1.848, 95%CI: 1.340–2.548, $P < 0.001$) were significantly associated with survival time of oral squamous cell carcinoma patients.

Multivariate Cox regression analysis showed that tumor grade (HR = 1.466, 95%CI: 1.064–2.020, P = 0.019) and COL11A1 (HR = 1.645, 95%CI: 1.164–2.325, P = 0.005) were significantly correlated with survival time of oral squamous cell carcinoma patients.

Oral squamous cell carcinoma is the most common malignant tumor of the head and neck. It refers to malignant transformation of the oral squamous epithelium, resulting in malignant tumor [11]. Early oral squamous cell carcinoma after treatment, the cure rate is high. The prognosis and quality of life of advanced patients need to be improved. Oral squamous cell carcinoma is prone to distant metastasis, such as lung metastasis, bone metastasis, lymphatic metastasis, blood transport metastasis and so on.

COL11A1 gene, encoding gene of $\alpha 1$ subunit of collagen, was mainly expressed in cartilage and was a small amount of fibrous collagen expressed by chondrocytes and osteoblasts [12]. COL11A1 forms heterotrimers with COL11A2 and COL2A1 to assemble type XI collagen.

COL11A1, an important member of the collagen family, is mainly expressed and secreted by cancer-associated fibroblast subpopulations and regulates tumor-matrix interactions and the mechanical properties of the extracellular matrix. COL11A1 also promotes cancer cell migration, metastasis and treatment resistance by activating pro-survival pathways and regulating tumor metabolic phenotypes [13]. Related research results suggest that COL11A1 is secreted by Extracellular matrix and ECM plays an important role in tumor development. Collagen is the most abundant component of ECM and is involved in the biological formation of cancer [14]. COL11A1 plays a role in epithelial mesenchymal transformation, metastasis and invasion, a process of tumor transformation. As one of the important components of tumor microenvironment, COL11A1 has attracted more and more attention in recent years. It is highly expressed in most human tumor cell lines and tissues, and may play a role in promoting cancer and affecting tumor cell proliferation by regulating cell cycle. It affects cancer progression by regulating several downstream genes involved in biological processes such as cell proliferation, apoptosis, metastasis, and signal transduction [15]. COL11A1 overexpression has been observed in many cancer types [16]. Studies have shown that COL11A1 is highly expressed in patients with gastric cancer, which may be a potential therapeutic target [17]. COL11A1 has also been shown to promote tumor progression and predict adverse clinical outcomes in ovarian cancer [18]. COL11A1 is overexpressed in recurrent NSCLC and promotes cell proliferation, migration, invasion and drug resistance [19].

COL11A1 is associated with cancer progression and low survival rate, and its expression is positively correlated with progression and lymph node metastasis. COL11A1 expression is not present in benign pathological conditions, including adhesion hyperplasia, fibrosis, cirrhosis, pancreatitis, and inflammatory bowel disease or precancerous lesions [20]. Under normal circumstances, COL11A1 expression is low or almost non-existent in most tissues [21]. Normal fibroblasts secrete some collagen, including COL11A1, when trauma occurs in human tissue. In clinical samples, high COL11A1 expression tends to predict distant metastasis, poor prognosis, and short overall survival. COL11A1 is believed to be involved in many intracellular signaling pathways that tend to accumulate in tumor tissues and contribute to the malignant progression of human cancers [10]. COL11A1 overexpression is associated with poor

prognosis and mainly occurs in solid tumors. Role of COL11A1 as a cancer-specific but not inflammation-specific biomarker, but COL11A1 overexpression has been demonstrated in chronic inflammatory diseases such as osteoarthritis [22]. COL11A1 is expressed not only in tumor cells, but also in other tumor-associated stromal cells. COL11A1 promotes cancer progression, metastasis, and drug resistance by binding to specific receptors and activating several key cell survival signaling pathways. COL11A1 is the only up-regulated collagen gene in the stroma of invasive cancer [23]. Therefore, it is speculated that the expression level of COL11A1 may be related to oral squamous cell carcinoma.

However, there are some shortcomings in this study. Although clinical data have been examined and analyzed, the molecular mechanisms by which COL11A1 expression levels affect outcomes in patients with oral squamous cell carcinoma have not been validated in animal models. Therefore, future studies should focus on animal experiments to explore the molecular pathway and mechanism of COL11A1 in oral squamous cell carcinoma.

Conclusion

COL11A1 expression level was significantly correlated with the occurrence of oral squamous cell carcinoma and the survival time of patients. High COL11A1 expression is associated with poor prognosis and short survival time in patients with oral squamous cell carcinoma. COL11A1, as a potential prognostic target of oral squamous cell carcinoma, provides a new direction for the molecular mechanism of its occurrence and development.

Abbreviations

Oral squamous cell carcinoma (OSCC), differently expressed genes (DEGs), Comparative Toxicogenomics Database (CTD), Gene Set Enrichment Analysis (GSEA), Receiver operating characteristic (ROC), support vector machine (SVM), hazard ratios (HRs), 95% confidence intervals (95%CI)

Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Fourth Hospital of Hebei Medical University Hospital, and all patients signed informed consent.

Consent for publication

All authors reviewed the paper and agreed to publish it.

Availability of data and materials

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

Competing interests

The authors declare that there is no conflict of interest.

Funding

This project was funded by the Key Science and Technology Research Program of Hebei Provincial Health Commission, No.202110801.

Authors' contributions

Yong-bin Di participated in the topic selection and design of the paper. Yang Bao and Jie Guo participated in data analysis and interpretation. Wei Liu, Guan-hua Zhang and Su-xin Zhang drafted and revised key theories in the paper. Tian-ke Li answered academic questions.

Acknowledgements

Not applicable

References

1. Hasegawa K, Fujii S, Matsumoto S, Tajiri Y, Kikuchi A, Kiyoshima T. YAP signaling induces PIEZO1 to promote oral squamous cell carcinoma cell proliferation. *J Pathol.* 2021;253(1):80–93.
2. Vitória JG, Duarte-Andrade FF, Dos Santos Fontes Pereira T, Fonseca FP, Amorim L, Martins-Chaves RR, et al. Metabolic landscape of oral squamous cell carcinoma. *Metabolomics.* 2020;16(10):105.
3. McCord C, Kiss A, Magalhaes MA, Leong IT, Jordan T, Bradley G. Oral Squamous Cell Carcinoma Associated with Precursor Lesions. *Cancer Prev Res (Phila).* 2021;14(9):873–84.
4. Qian DC, Magliocca KR, Aiken AH, Baugnon KL, Brandon DC, Stokes WA, et al. Outcomes and Predictive Value of Post-adjuvant Therapy PET/CT for Locally Advanced Oral Squamous Cell Carcinoma. *Laryngoscope.* 2020;130(12):E850-850E857.
5. Fan L, Hu X, Lin S, Zhou W, Fu S, Lv H. Concurrent preoperative chemotherapy and three-dimensional conformal radiotherapy followed by surgery for oral squamous cell carcinoma: a retrospective analysis of 104 cases. *Oncotarget.* 2017;8(43):75557–67.
6. Truong DQ, Ho BT, Chau GC, Truong DK, Pham T, Nakagawara A, et al. Collagen XI Alpha 1 (COL11A1) Expression in the Tumor Microenvironment Drives Neuroblastoma Dissemination. *Pediatr Dev Pathol.* 2021:10935266211039200.
7. Kolácná L, Bakesová J, Varga F, Kostáková E, Plánka L, Necas A, et al. Biochemical and biophysical aspects of collagen nanostructure in the extracellular matrix. *Physiol Res.* 2007;56 Suppl 1:S51-51S60.
8. Mladenova M, Todorov T, Grozdanova L, Mitev V, Todorova A. Novel Mutation in the COL11A1 Gene Causing Marshall-Stickler Syndrome in Three Generations of a Bulgarian Family. *Balkan J Med*

- Genet. 2021;24(1):95–8.
9. Mio F, Chiba K, Hirose Y, Kawaguchi Y, Mikami Y, Oya T, et al. A functional polymorphism in COL11A1, which encodes the alpha 1 chain of type XI collagen, is associated with susceptibility to lumbar disc herniation. *Am J Hum Genet.* 2007;81(6):1271–7.
 10. Liu Z, Lai J, Jiang H, Ma C, Huang H. Collagen XI alpha 1 chain, a potential therapeutic target for cancer. *FASEB J.* 2021;35(6):e21603.
 11. Panarese I, Aquino G, Ronchi A, Longo F, Montella M, Cozzolino I, et al. Oral and Oropharyngeal squamous cell carcinoma: prognostic and predictive parameters in the etiopathogenetic route. *Expert Rev Anticancer Ther.* 2019;19(2):105–19.
 12. Vázquez-Villa F, García-Ocaña M, Galván JA, García-Martínez J, García-Pravia C, Menéndez-Rodríguez P, et al. COL11A1/(pro)collagen 11A1 expression is a remarkable biomarker of human invasive carcinoma-associated stromal cells and carcinoma progression. *Tumour Biol.* 2015;36(4):2213–22.
 13. Nallanthighal S, Heiserman JP, Cheon DJ. Collagen Type XI Alpha 1 (COL11A1): A Novel Biomarker and a Key Player in Cancer. *Cancers (Basel).* 2021;13(5).
 14. Raglow Z, Thomas SM. Tumor matrix protein collagen Xla1 in cancer. *Cancer Lett.* 2015;357(2):448–53.
 15. Wu YH, Huang YF, Chang TH, Wu PY, Hsieh TY, Hsiao SY, et al. miR-335 Restrains the Aggressive Phenotypes of Ovarian Cancer Cells by Inhibiting COL11A1. *Cancers (Basel).* 2021;13(24).
 16. Jia D, Liu Z, Deng N, Tan TZ, Huang RY, Taylor-Harding B, et al. A COL11A1-correlated pan-cancer gene signature of activated fibroblasts for the prioritization of therapeutic targets. *Cancer Lett.* 2016;382(2):203–14.
 17. Li A, Li J, Lin J, Zhuo W, Si J. COL11A1 is overexpressed in gastric cancer tissues and regulates proliferation, migration and invasion of HGC-27 gastric cancer cells in vitro. *Oncol Rep.* 2017;37(1):333–40.
 18. Wu YH, Chang TH, Huang YF, Huang HD, Chou CY. COL11A1 promotes tumor progression and predicts poor clinical outcome in ovarian cancer. *Oncogene.* 2014;33(26):3432–40.
 19. Shen L, Yang M, Lin Q, Zhang Z, Zhu B, Miao C. COL11A1 is overexpressed in recurrent non-small cell lung cancer and promotes cell proliferation, migration, invasion and drug resistance. *Oncol Rep.* 2016;36(2):877–85.
 20. Zheng X, Liu X, Zheng H, Wang H, Hong D. Integrated bioinformatics analysis identified COL11A1 as an immune infiltrates correlated prognosticator in pancreatic adenocarcinoma. *Int Immunopharmacol.* 2021;90:106982.
 21. Fischer H, Salahshor S, Stenling R, Björk J, Lindmark G, Iselius L, et al. COL11A1 in FAP polyps and in sporadic colorectal tumors. *BMC Cancer.* 2001;1:17.
 22. Fernández-Torres J, Martínez-Nava GA, Zamudio-Cuevas Y, Martínez-Flores K, Mijares-Díaz F. Multifactor dimensionality reduction reveals a strong gene-gene interaction between STC1 and COL11A1 genes as a possible risk factor of knee osteoarthritis. *Mol Biol Rep.* 2020;47(4):2627–34.

23. Wu YH, Huang YF, Chang TH, Chen CC, Wu PY, Huang SC, et al. COL11A1 activates cancer-associated fibroblasts by modulating TGF- β 3 through the NF- κ B/IGFBP2 axis in ovarian cancer cells. *Oncogene*. 2021;40(26):4503–19.

Tables

Table 1. Relevant characteristics of patients with oral squamous cell carcinoma

Characteristics			Prognosis of OSCC		P
			Good	Poor	
Gender	Male	100	47 (23.5%)	53 (26.5%)	0.887
	Female	100	46 (23.0%)	54 (27.0%)	
Age	≤65	78	45 (22.5%)	33 (16.5%)	0.011*
	65	122	48 (24.0%)	74 (37.0%)	
Tumor size	≤ 1cm	99	50 (25.0%)	49 (24.5%)	0.261
	1cm	101	43 (21.5%)	58 (29.0%)	
Family history	No	80	40 (20.0%)	40 (20.0%)	0.418
	Yes	120	53 (26.5%)	67 (33.5%)	
Tumor grade	Low	101	55 (27.5%)	46 (23.0%)	0.023*
	High	99	38 (19.0%)	61 (30.5%)	
COL11A1	Low	99	74 (37.0%)	25 (12.5%)	<0.001*
	High	101	19 (9.5%)	82 (41.0%)	
Tumor stage	Low	108	53 (26.5%)	55 (27.5%)	0.429
	High	92	40 (20.0%)	52 (26.0%)	

Pearson's Chi-square test. OSCC: oral squamous cell carcinoma. *P 0.05.

Table 2. Relationship between patient characteristics and prognosis of oral squamous cell carcinoma		
Characteristics	Prognosis of OSCC	
	ρ	P
Gender	0.010	0.888
Age	0.179	0.011*
Tumor size	0.079	0.263
Family history	0.057	0.420
Tumor grade	0.161	0.023*
COL11A1	0.561	<0.001*
Tumor stage	0.056	0.432
Spearman correlation analysis. OSCC: oral squamous cell carcinoma.*P<0.05		

Table 3. Effects of related parameters on the prognosis of oral squamous cell carcinoma based on univariate Logistic regression analysis

Characteristics			Prognosis of OSCC		
			OR	95% CI	P
Gender	Male	100	1		0.887
	Female	100	1.041	0.597-1.815	
Age	≤65	78	1		0.012*
	65	122	2.102	1.180-3.746	
Tumor size	≤ 1cm	99	1		0.261
	1cm	101	1.376	0.788-2.404	
Family history	No	80	1		0.418
	Yes	120	1.264	0.717-2.229	
Tumor grade	Low	101	1		0.023*
	High	99	1.919	1.093-3.372	
COL11A1	Low	99	1		<0.001*
	High	101	12.775	6.509-25.071	
Tumor stage	Low	108	1		0.429
	High	92	1.253	0.716-2.190	

OR: odds ratio; 95% CI: 95% confidence interval. OSCC: oral squamous cell carcinoma. * P < 0.05

Table 4. Multivariate Logistic regression analysis of the prognostic characteristics and influence of oral squamous cell carcinoma

Characteristics	Prognosis of OSCC		
	OR	95%CI	P
Gender	0.847	0.422-1.699	0.640
Age	1.774	0.874-3.603	0.113
Tumor size	1.465	0.738-2.906	0.275
Family history	1.112	0.552-2.240	0.766
Tumor grade	1.561	0.786-3.098	0.203
COL11A1	12.066	6.042-24.096	<0.001*
Tumor stage	1.009	0.503-2.022	0.981

OR: odds ratio; 95% CI: 95% confidence interval. OSCC: oral squamous cell carcinoma. * P < 0.05

Table 5. Influence of correlation characteristics on patient survival time based on multivariate Cox regression analysis

Characteristics			Survival time		
			HR	95% CI	P
Gender	Male	100	1		0.129
	Female	100	1.275	0.932-1.743	
Age	≤65	78	1		0.005*
	65	122	1.592	1.150-2.205	
Tumor size	≤ 1cm	99	1		0.687
	1cm	101	1.066	0.781-1.454	
Family history	No	80	1		0.650
	Yes	120	0.929	0.675-1.278	
Tumor grade	Low	101	1		0.018*
	High	99	1.460	1.067-1.999	
COL11A1	Low	99	1		<0.001*
	High	101	1.848	1.340-2.548	
Tumor stage	Low	108	1		0.160
	High	92	1.252	0.915-1.714	

HR: hazard ratios; 95% CI: 95% confidence intervals. *P < 0.05

Table 6. Influence of correlation characteristics on patient survival time based on multivariate Cox regression analysis

Characteristics	Survival time		
	HR	95% CI	P
Gender	1.097	0.785-1.535	0.588
Age	1.379	0.976-1.949	0.068
Tumor size	1.068	0.779-1.464	0.682
Family history	0.959	0.694-1.326	0.801
Tumor grade	1.466	1.064-2.020	0.019*
COL11A1	1.645	1.164-2.325	0.005*
Tumor stage	1.242	0.902-1.711	0.184

HR: hazard ratios; 95% CI: 95% confidence intervals. *P < 0.05

Figures

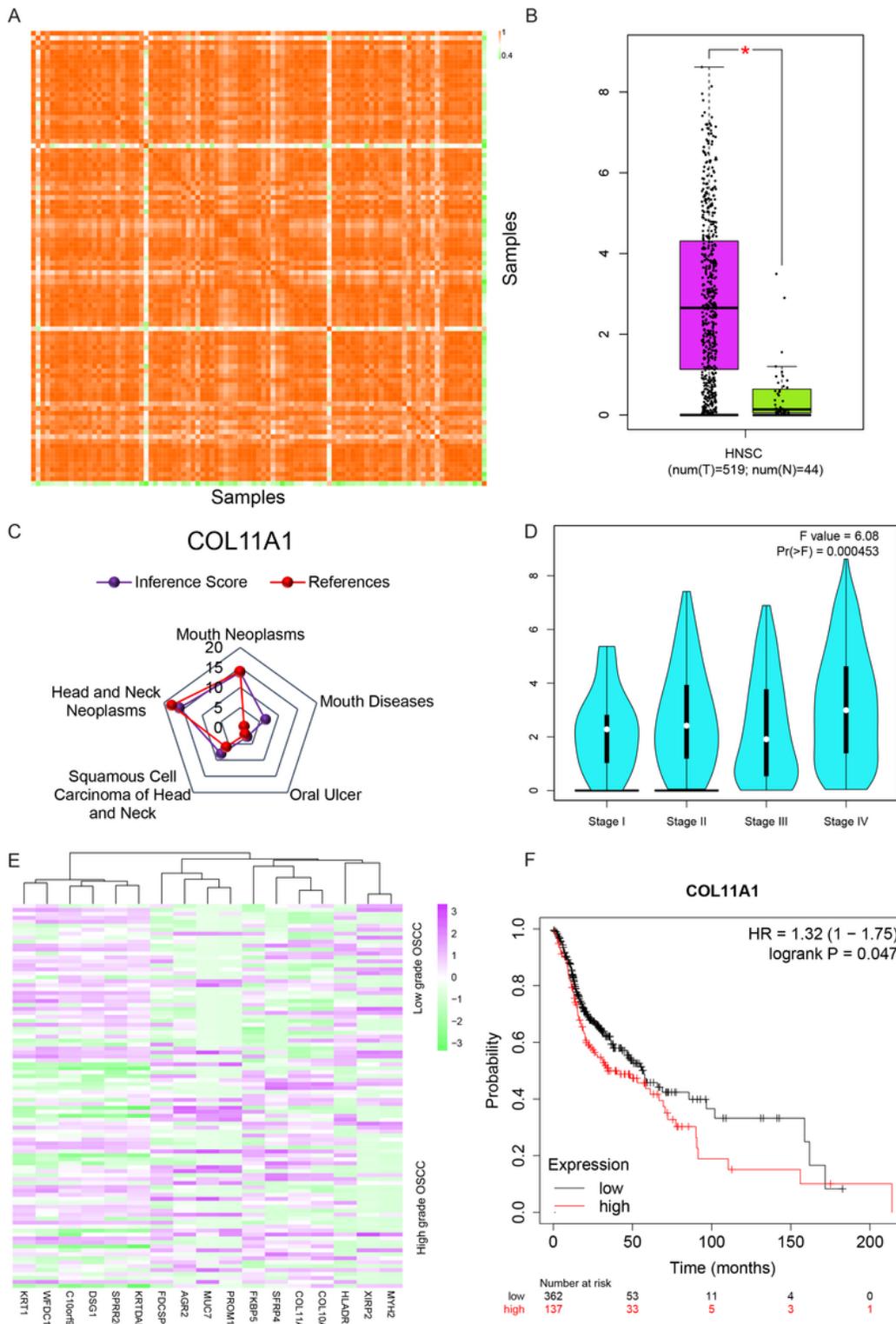


Figure 1

Repeatability test of the samples and analysis of the COL11A1. (A) In the GSE41613, there were strong relationships among all the samples, which manifested that the repeatability of the data is good. (B) Compared with the control group, the expression of COL11A1 in the tumor group was higher ($P < 0.05$) based on the TCGA database. (C) CTD analysis for the inference score of COL11A1 in the diseases related with OSCC. (D) Pathological stage of OSCC is higher when the expression of COL11A1 was

higher. (E) The heatmap manifested the expression of DEGs related with the OSCC. (F) Survival analysis of COL11A1 for the OSCC based on the TCGA.

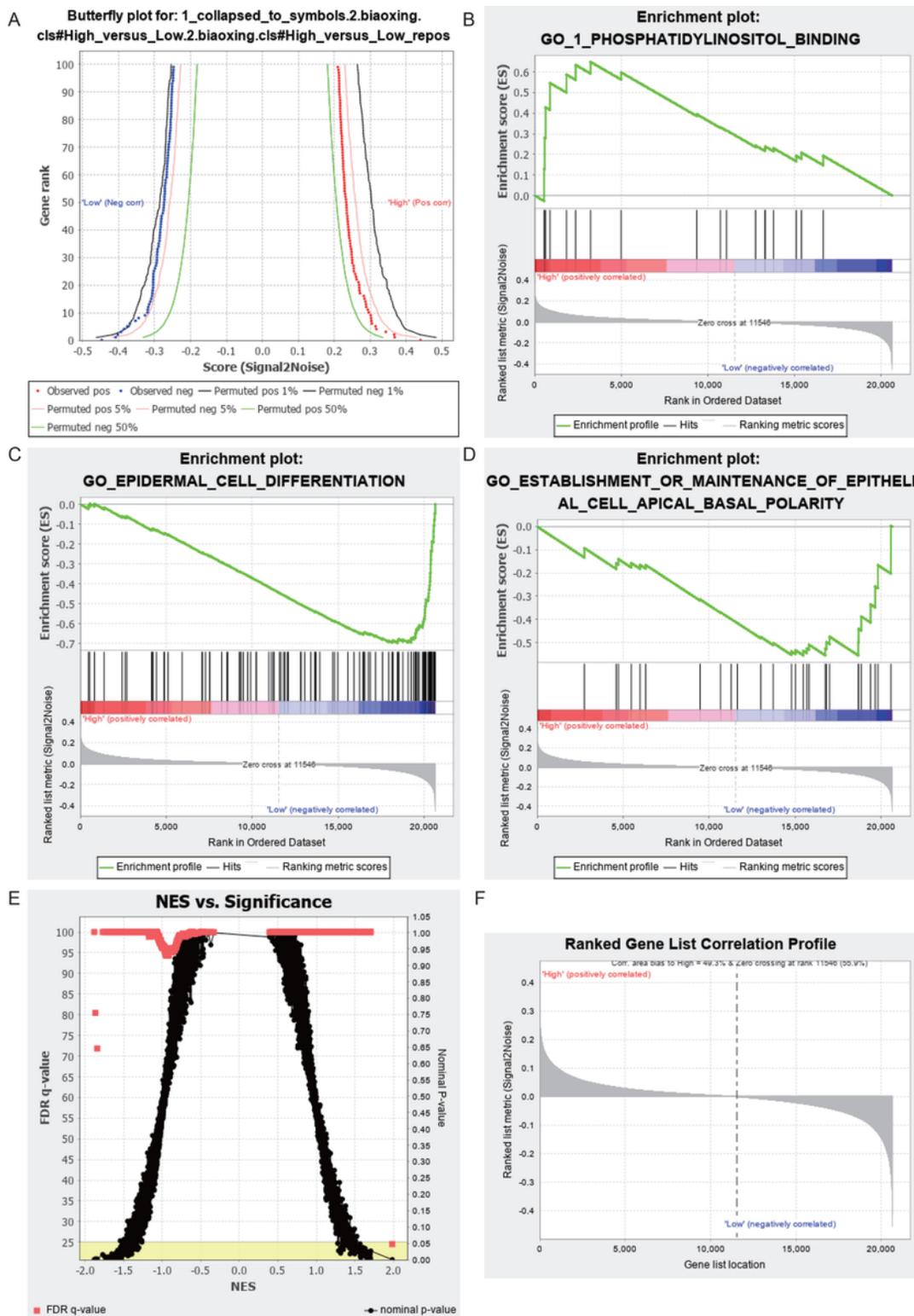


Figure 2

GO enrichment analysis for the OSCC. (A) Butterfly plot. (B) Phosphatidylinositol binding. (C) The epidermal cell differentiation. (D) Establishment or maintenance of epithelial cell apical basal polarity. (E)

NES vs. Significance. (F) Ranked gene list correlation profile.

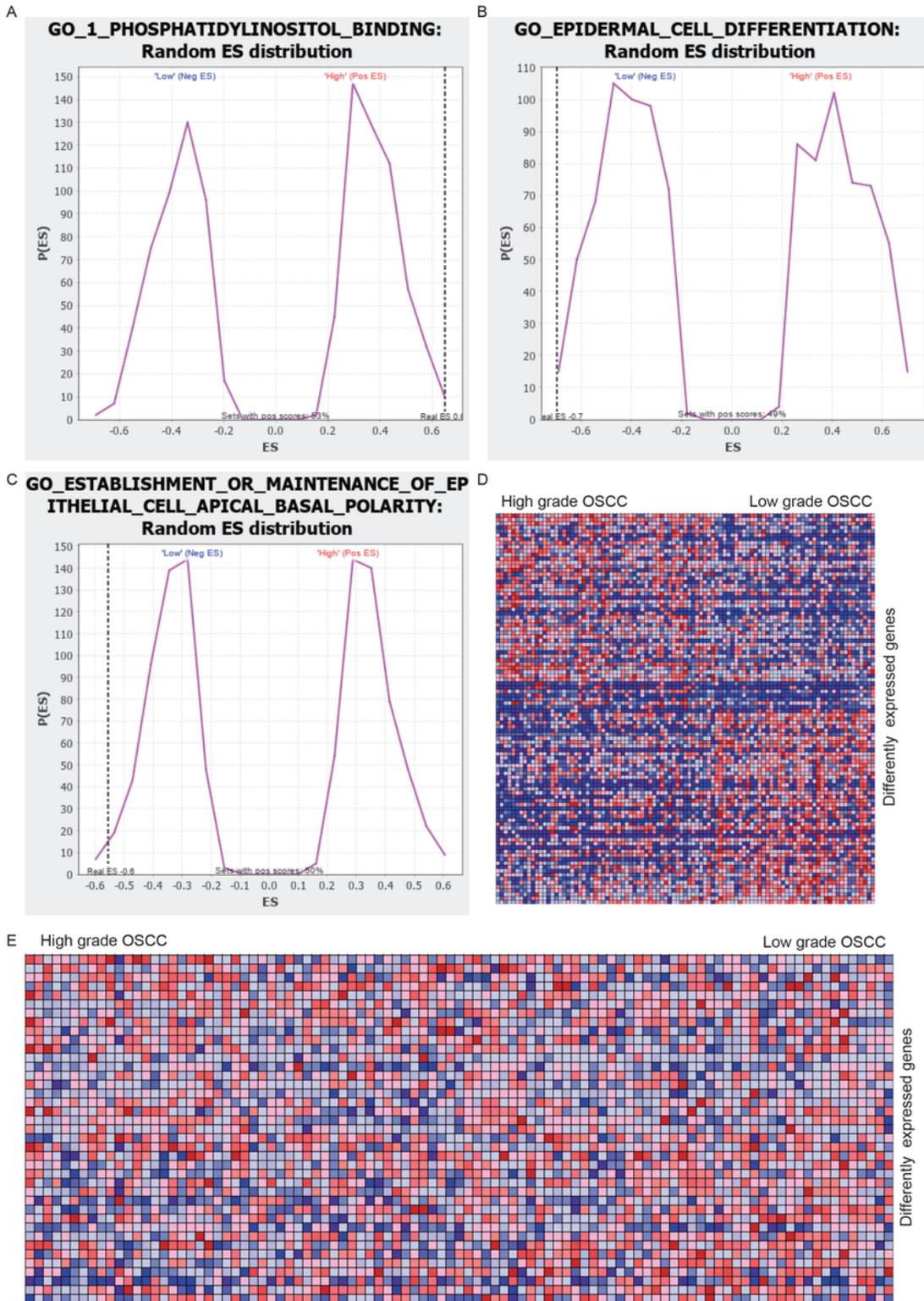


Figure 3

Random ES distributions and the heatmap for the expression of DEGs. (A) Random ES distribution of phosphatidylinositol binding. (B) Random ES distribution of epidermal cell differentiation. (C) Random ES

distribution of establishment or maintenance of epithelial cell apical basal polarity. (D, E) There were plenty of DEGs based on the expression, and the expression heat map could classify the OSCC

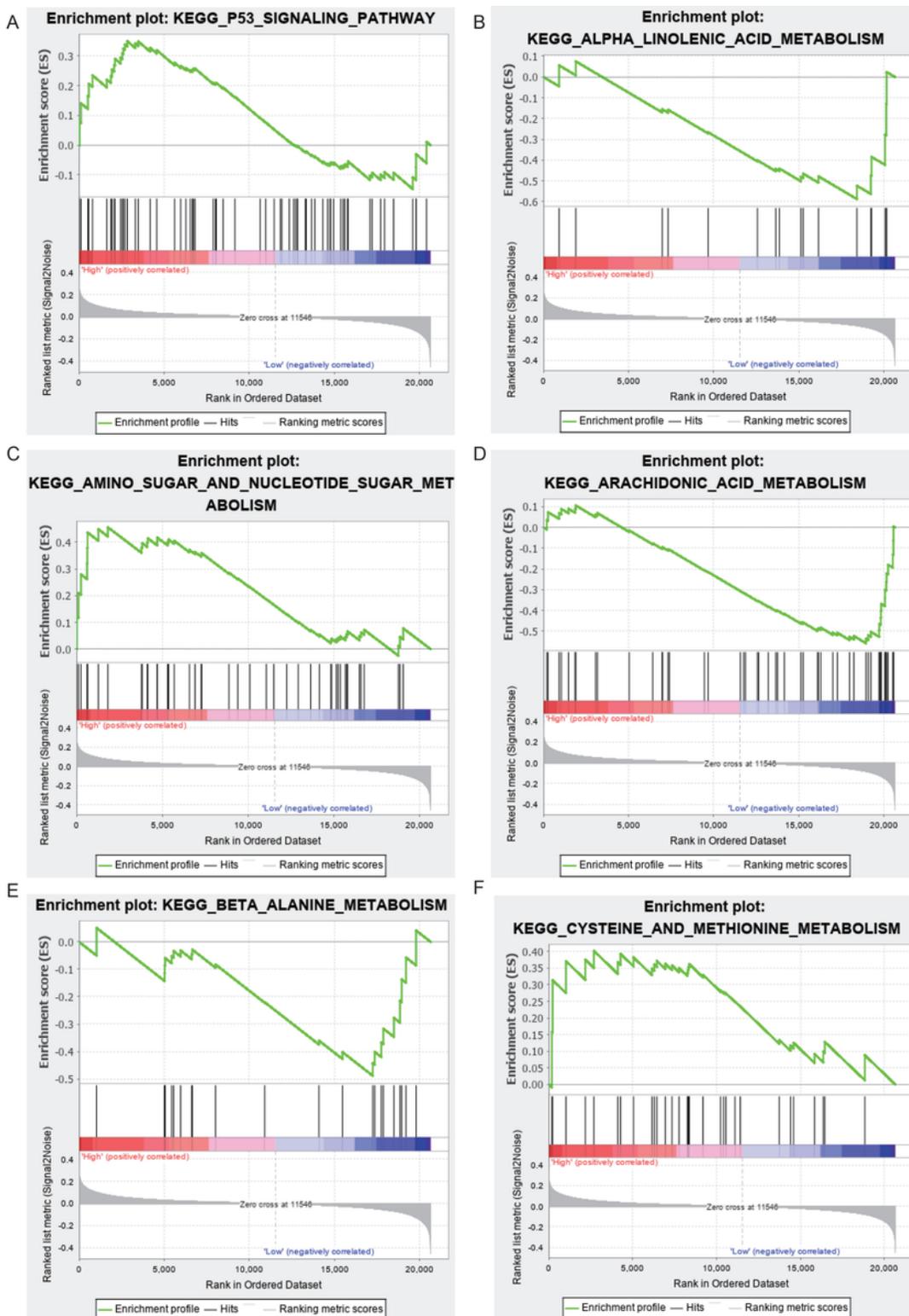


Figure 4

KEGG enrichment analysis for the OSCC. (A) P53 signaling pathway. (B) Alpha linolenic acid metabolism. (C) amino sugar and nucleotide sugar metabolism. (D) arachidonic acid metabolism. (E) beta alanine

metabolism. (F) cysteine and methionine metabolism.

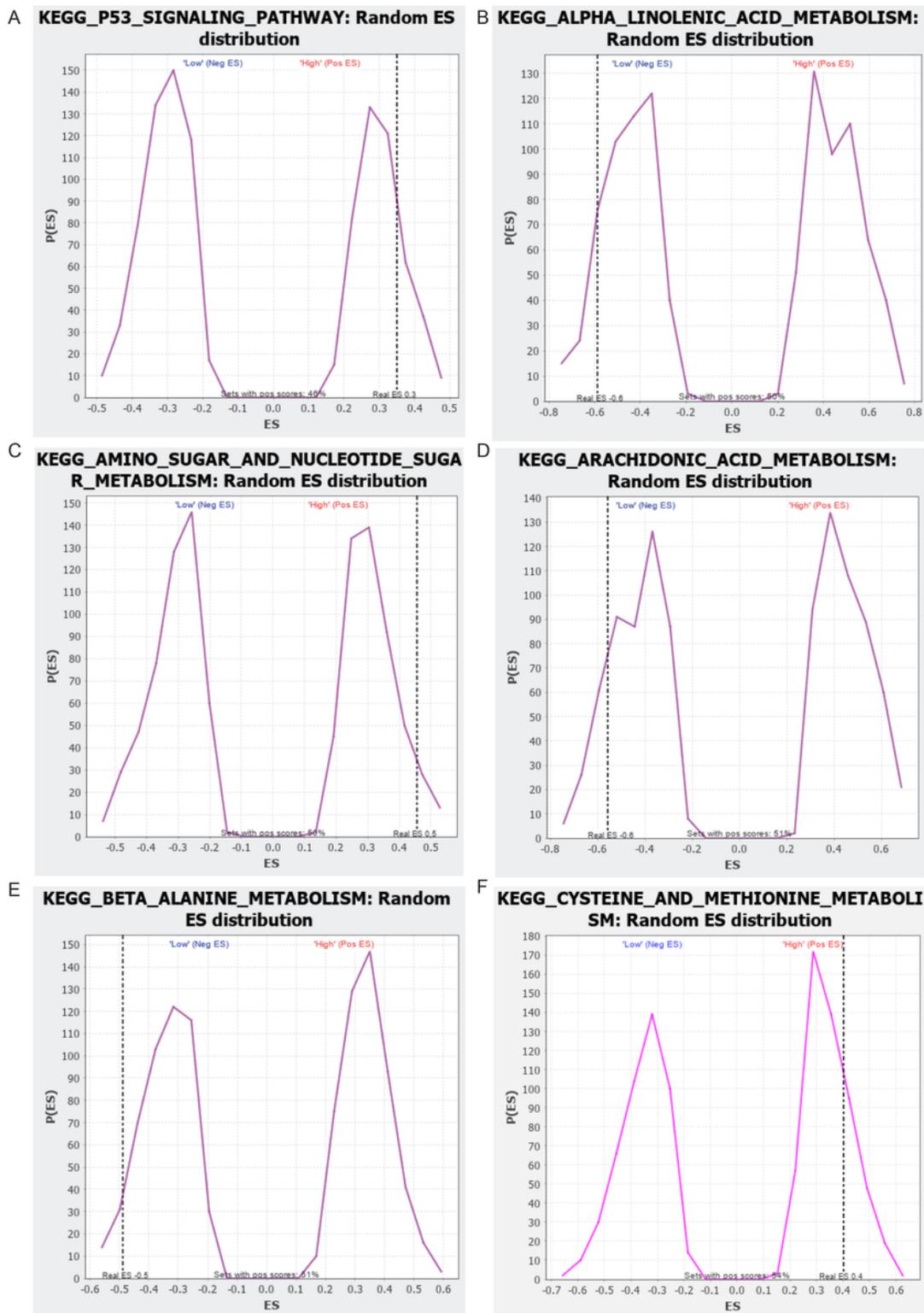


Figure 5

Random ES distributions of these KEGG pathways could classify the two groups of low grade and high grade OSCC. (A) P53 signaling pathway. (B) Alpha linolenic acid metabolism. (C) amino sugar and

nucleotide sugar metabolism. (D) arachidonic acid metabolism. (E) beta alanine metabolism. (F) cysteine and methionine metabolism.

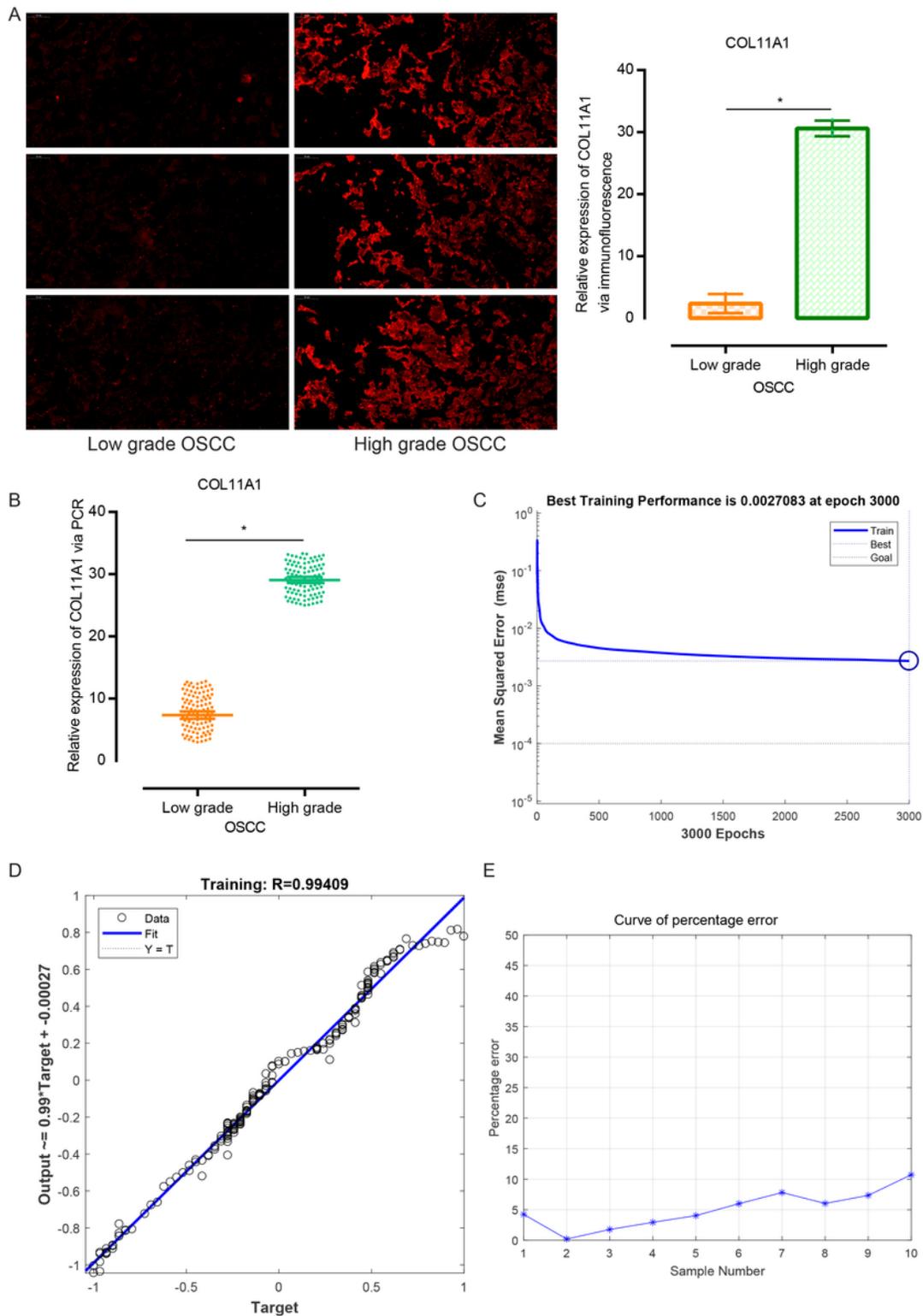


Figure 6

Verification of the expression of COL11A1 with the OSCC clinical samples and the BP-neural network. (A) Through the immunofluorescence assay, the expression of COL11A1 was higher in the high grade OSCC

than the low grade OSCC ($P < 0.05$). (B) Compared with the low grade OSCC, the mRNA expression of COL11A1 was higher in the high grade OSCC via the PCR assay. (C-E) Through the BP-neural network, best training performance is 0.0027083 at epoch 3000, and the training relevance (R) is 0.99409 between COL11A1 and the OSCC survival time. Furthermore, the percentage errors of the BP-neural network were small.

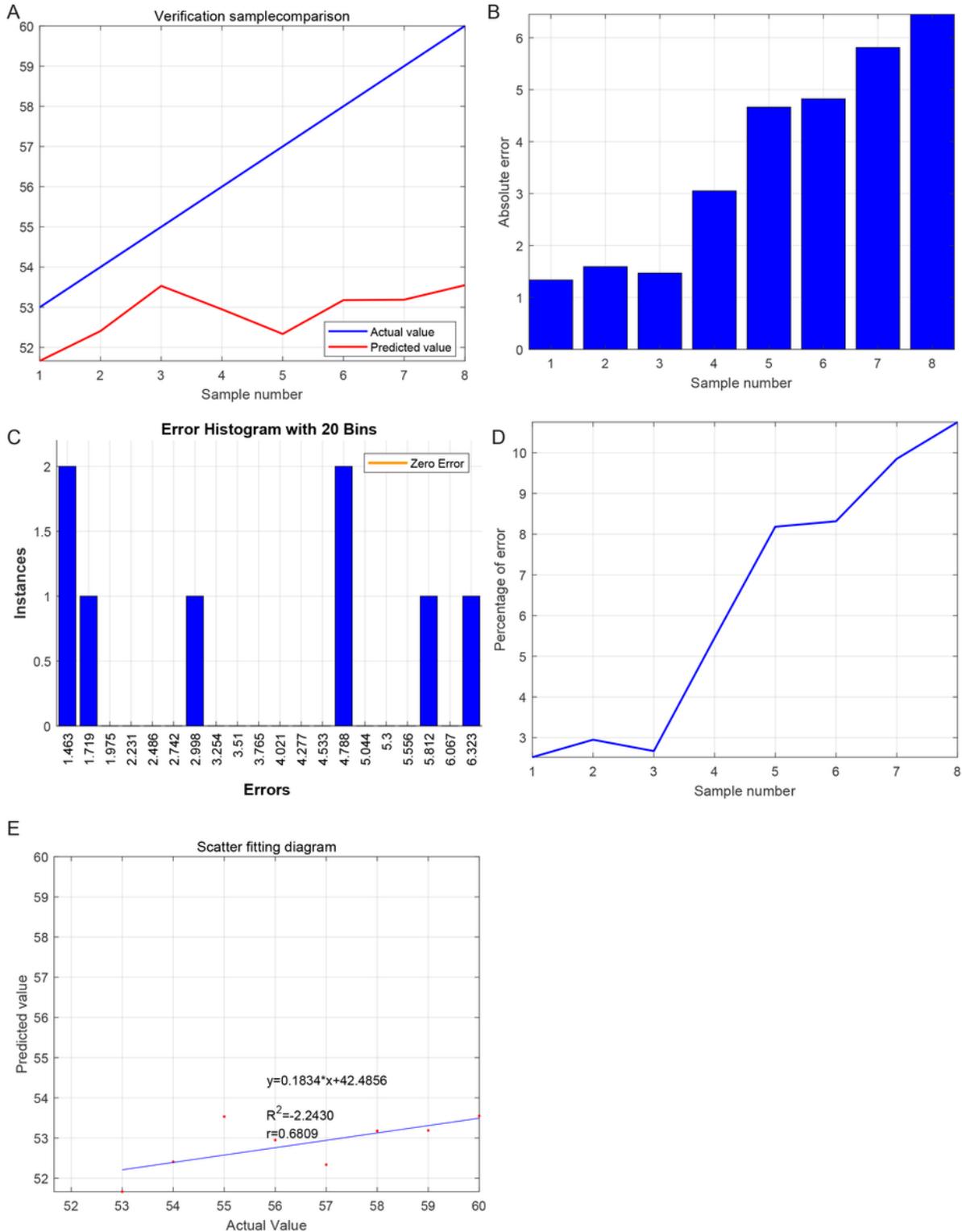


Figure 7

support vector machine of expression of COL11A1 for the survival time of OSCC.

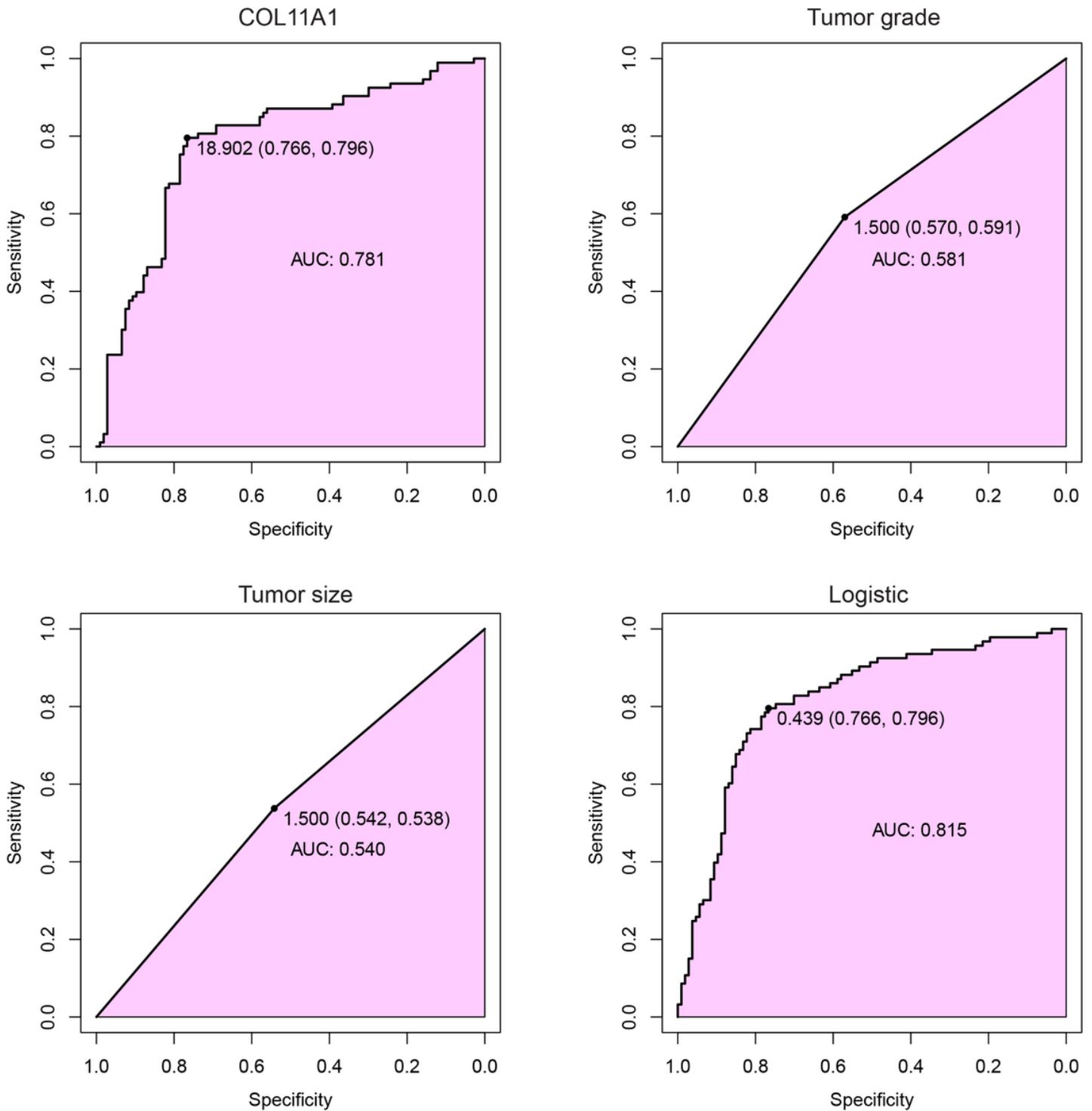


Figure 8

High sensitivity and specificity of COL11A1 for diagnosing the OSCC based on the ROC analysis.

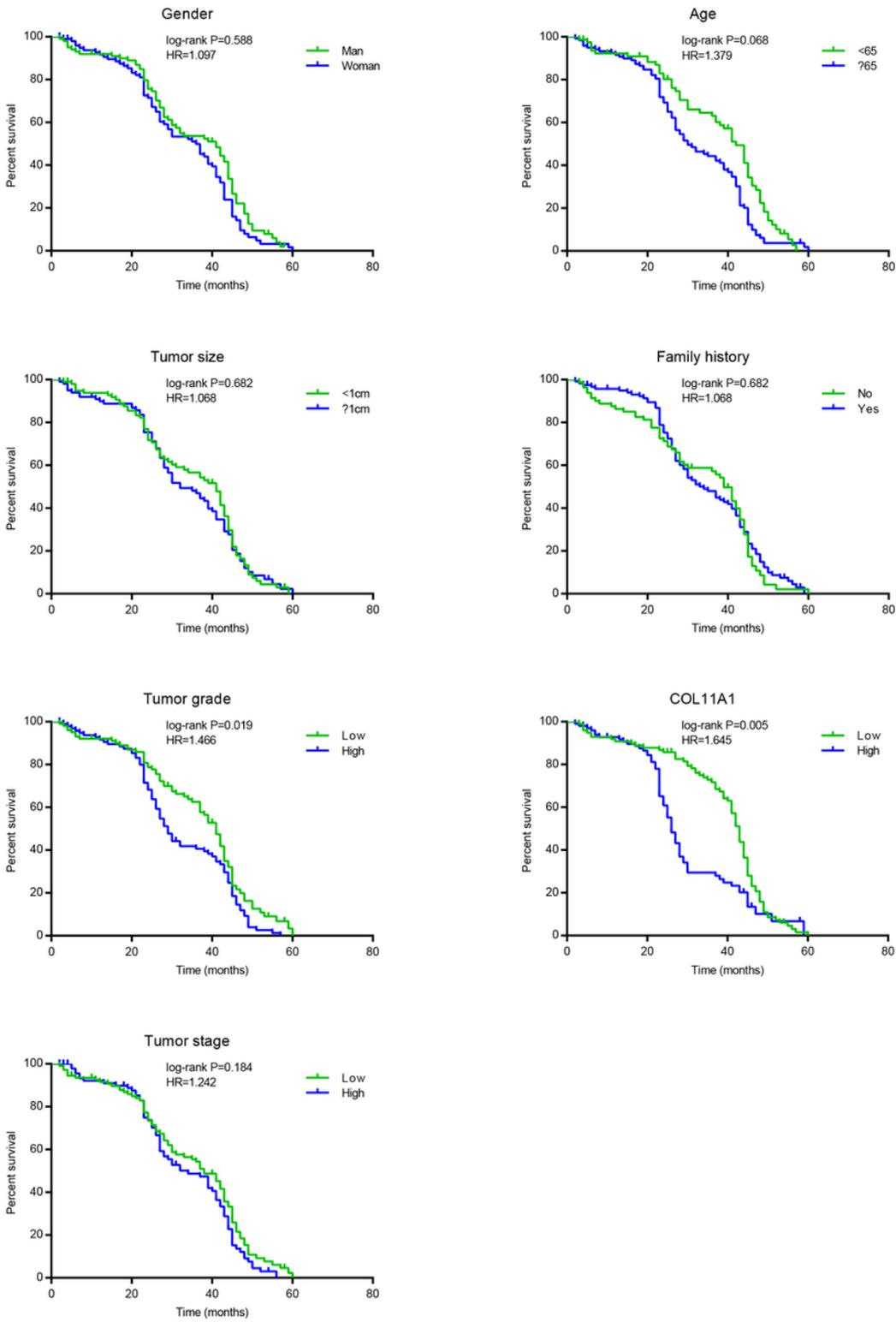


Figure 9

Overall survival analysis for the OSCC. Based on the clinical samples, the OSCC patients with higher expression of COL11A1 have the poor overall survival compare with the OSCC patients with lower expression of COL11A1 (HR=1.645, P=0.005).