Construction of a new predictive model in head and neck squamous cell carcinoma based on the investigation of extracellular matrix-associated genes

LIANGQIAN TU[1], YIYIN QIU[2], HANDAN ZHENG[3], BIXIN FANG[4] and MING CHEN[4]

[1]Department of Otolaryngology, The Liangzhu Street Community Health Service Center, Hangzhou, Zhejiang 311113, P.R. China; [2]Department of Otolaryngology, Ninghai First Hospital, Ninghai, Ningbo, Zhejiang 315600, P.R. China; [3]Department of Otolaryngology, The Second Affiliated Hospital, Zhejiang Chinese Medical University, Hangzhou, Zhejiang 310005, P.R. China; [4]Department of Otolaryngology, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang 310009, P.R. China

Received October 23, 2025; Accepted July 16, 2026

DOI: 10.3892/ol.2026.15839

Abstract. A key aspect influencing immune cell infiltration is the composition of the extracellular matrix (ECM). Therefore, investigating the association between ECM-associated proteins and immune cell infiltration is key for the identification of new biomarkers to distinguish 'immune-hot' solid tumors and predict patient prognosis. A total of 513 head and neck squamous cell carcinoma (HNSCC) cases as training samples from The Cancer Genome Atlas and an additional 270 as testing samples from the Gene Expression Omnibus were obtained for use in the present study. Using a single-sample Gene Set Enrichment Analysis method, the 513 training samples were divided into Cluster 1 and Cluster 2. Subsequently, the present analysis uncovered 1,573 differentially expressed genes distinguishing the two clusters. After performing an intersection analysis with 751 ECM-associated genes, 103 differentially expressed ECM-associated genes were identified. Least absolute shrinkage and selection operator-Cox and multivariate Cox regression analyses were employed to identify candidate ECM risk genes ($P<0.05$) and to construct a predictive model. Finally, a nomogram and a three gene (cerebellin 2, galectin-10 and cathepsin G) predictive model were developed. Therefore, the present prognostic risk score model can evaluate the immune infiltration, predict the prognosis of HNSCC, and potentially guide more personalized immunotherapy interventions.

Correspondence to: Dr Ming Chen, Department of Otolaryngology, The Second Affiliated Hospital, School of Medicine, Zhejiang University, 88 Jiefang Road, Shangcheng, Hangzhou, Zhejiang 310009, P.R. China
E-mail: chenming23@zju.edu.cn

Key words: head and neck squamous cell carcinoma, extracellular matrix, predictive model, immunotherapy, tumor microenvironment

Introduction

The global incidence of head and neck squamous cell carcinoma (HNSCC) is predicted to escalate, with ~890,000 new cases (representing 4.5% of all cancer cases) and 450,000 deaths annually, with a projected 30% increase by 2030 (1). Furthermore, a stagnation has been observed in the 5-year survival rates for patients with HNSCC over the past three decades, with negligible advancements recorded. For patients undergoing salvage surgery, the 5-year overall survival rate is 42.2% (2). Immunotherapy, such as genetically modified immune cells or antitumoral cytokines, has been extensively investigated as a cancer treatment (3,4). However, due to the immunosuppressive tumor microenvironment (TME), the dysfunction of immune effector cells, HNSCC shows therapeutic resistance to a number of immunotherapies, including checkpoint inhibitors. Therefore, novel biomarkers are needed to identify 'immune-hot' solid tumors, characterized by high immune infiltration, and to predict the prognosis.

As a key element of TME, the extracellular matrix (ECM) serves pivotal roles in cancer cell proliferation, the establishment of an immunosuppressive TME, immune evasion, angiogenesis, migration and invasion (5). Furthermore, ECM remodeling directly modulates the recruitment and activity of immune cells by creating physical barriers that impede infiltration and providing biochemical cues that shape immune cell function including tumor-associated macrophages, myeloid-derived suppressor cells, regulatory T cells and $CD8^+$ T cells, thereby contributing to the formation of an immunosuppressive TME (6,7). Conversely, certain ECM fragments, known as matrikines, have been shown to be associated with enhanced T-cell infiltration, contributing to tumor growth suppression (8).

Dense ECM in HNSCC can impede the infiltration of immune cells (9). Tenascin C has been shown to accelerate tumor progression in a murine model possessing intact immune function, by fostering the development of an immunosuppressive stromal environment (10). The ECM serves a key role in influencing resistance to immune checkpoint inhibitor (ICI) therapy (11). The interaction of matrisome-associated proteins

(including ECM-affiliated proteins, ECM regulators and secreted factors) with the core ECM serves an important role in ECM restructuring (12). Therefore, identifying the association between ECM-associated proteins and the infiltration of immune cells will not only help predict the immune microenvironment and prognosis of patients but also the response rate of patients to immunotherapy.

Materials and methods

Data collection. A total of 513 HNSCC sample data from The Cancer Genome Atlas (TCGA) website was downloaded and used as the training dataset (https://portal.gdc.cancer.gov/). In addition, 270 HNSCC sample data from the GSE65858 dataset (ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65858) was downloaded from the Gene Expression Omnibus (GEO) database and used as the independent external test dataset (13). Clinical data (including age and TNM stage) regarding all patients were acquired for further examination and analysis (Table I).

Immune subtype analysis. To ascertain the infiltration levels of 28 immune cell populations within the training dataset, the Gene Set Variation Analysis tool implemented in R software (version 4.3.2; Posit Software, PBC) was used. This process depended upon the application of the single-sample gene set enrichment analysis (ssGSEA) methodology (14). The cell signatures of 28 immune cells were downloaded from the tumor-immune system interaction database datasets (cis.hku.hk/TISIDB/data/download/CellReports.txt). Based on immune infiltration data, cluster analysis was conducted for the training dataset using 'ConsensusClusterPlus' R package (version 1.64.0) (15). Utilizing the cumulative distribution function (CDF), the optimal number of clusters for the present analysis were determined. In the present study, the optimal number of clusters with the highest within-cluster average consensus was K=2. Estimation of stromal and immune cells in malignant tumors using expression data (ESTIMATE) score and tumor purity were calculated utilizing the 'estimate' R package (version 1.0.13) (16). Cluster 1 was characterized by exhausted immune cells and Cluster 2 was rich in the infiltration of immune cells. Results were visualized using the 'ggplot2' R package (version 3.4.4), and 'pheatmap' R package (CRAN.R-project.org/package=pheatmap, version 1.0.12) (17).

Identifying ECM genes potentially associated with immunity. The 'DESeq2' package (version 1.32.0) was utilized to identify the differentially expressed genes (DEGs) in TCGA training dataset (https://portal.gdc.cancer.gov/) (18). After excluding genes with a proportion of expression values of 0 that were >50%, 1,573 DEGs were identified that exhibited significant differences [false discovery rate (FDR) <0.05 and $\log_2$ fold change >1.25] between the two clusters. A total of 275 genes encoding core ECM and 751 genes encoding ECM-associated proteins were downloaded from the Molecular Signature Database (19). Intersection analysis revealed 103 differentially expressed ECM-associated genes that may exhibit an association with the infiltration of immune cells in HNSCC.

Differentially expressed ECM genes and immunity. Subsequently, the biological functions and pathways to explore the association between these differentially expressed ECM genes and the immune microenvironment were investigated. Gene set enrichment analysis (version 3.0) and the Database for Annotation, Visualization and Integrated Discovery were employed for conducting Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses (david.abcc.ncifcrf.gov) (20). A FDR threshold of <0.25 was adopted as the inclusion criterion for enriched gene sets. This threshold is recommended by the developers of GSEA to balance the discovery of potentially relevant biological pathways with the control of false positives, particularly when analyzing a relatively small number of gene sets (21).

To explore the characteristics of the 103 overlapping differentially expressed ECM-associated genes, the 103 genes were first imported into the Search Tool for the Retrieval of Interacting Genes/Proteins database to obtain the corresponding protein-protein interaction (PPI) data (string-db.org/). The PPI data were then imported into Cytoscape software (version 3.7.2; cytoscape.org/), and the 'cytoHubba' plugin was used to screen for hub genes (22). The Maximal Clique Centrality algorithm, which offers high accuracy in identifying core regulatory genes in PPI networks, was selected for node scoring (22). Based on the scoring results, the top 10 genes were identified as hub genes: CXCL10, CXCL9, IL13, CCL22, CXCL13, CCL4, IL17A, IFNG, IL10 and CCL19.

Statistical analysis. Least absolute shrinkage and selection operator (LASSO)-Cox regression analysis, 10-fold cross-validation, and multivariate Cox regression analysis were performed to seek candidate ECM risk genes (P<0.05) and constructed a prognostic risk score model. Cox regression analysis with LASSO regularization was implemented utilizing the 'glmnet' package (version 4.1.8) (23). The formula for the risk score was constructed as follows:

Risk score=Σ (Expression$_i$ x β_i) where expression$_i$ denotes the normalized expression level of the respective gene, and β_i is the corresponding regression coefficient derived from the multivariate Cox regression analysis.

The 'rms' R package (CRAN.R-project.org/package=rms, version 6.7.1) was employed to construct a nomogram model. Both the training dataset and the external test datasets were divided into high-risk and low-risk subsets (Table II) and each of these subsets was subjected to Kaplan-Meier (K-M) analyses and the construction of receiver operating characteristic (ROC) curves. The R packages 'pROC' (version 1.17.0.1), 'survival' (CRAN.R-project.org/package=survival, version 3.5.7) and 'survminer' (CRAN.R-project.org/package=survminer, version 0.4.9) were used for visualization (24).

Collection of clinical HNSCC samples. A total of 17 HNSCC clinical tissues and corresponding adjacent non-tumor tissue specimens were collected from patients (15 male and 2 female; median age, 70 years; age range 58-77 years) who underwent surgical resection at the Department of Otolaryngology, The Second Affiliated Hospital (Hangzhou, China). Of these, 8 retrospective cases were obtained from patients who underwent

Table I. Clinical characteristics of patients with HNSCC.

Variable	Total (n=783)	GEO (n=270)	TCGA (n=513)
Median overall survival time, years (IQR)	758.00 (419.00-1140.50)	837.00 (598.25-1073.75)	654.00 (379.00-1190.00)
Median age, years (IQR)	60.00 (53.00-68.57)	58.64 (52.41-68.42)	61.00 (53.00-69.00)
Survival status, n (%)			
Alive	473 (60.41)	176 (65.19)	297 (57.89)
Dead	310 (39.59)	94 (34.81)	216 (42.11)
Sex, n (%)			
Female	181 (23.12)	47 (17.41)	134 (26.12)
Male	602 (76.88)	223 (82.59)	379 (73.88)
T, n (%)			
T1	70 (9.13)	35 (12.96)	35 (7.04)
T2	229 (29.86)	80 (29.63)	149 (29.98)
T3	191 (24.90)	58 (21.48)	133 (26.76)
T4	277 (36.11)	97 (35.93)	180 (36.22)
N, n (%)			
N0	336 (44.15)	94 (34.81)	242 (49.29)
N1	114 (14.98)	32 (11.85)	82 (16.70)
N2	291 (38.24)	131 (48.52)	160 (32.59)
N3	20 (2.63)	13 (4.81)	7 (1.43)
M, n (%)			
M0	745 (98.28)	263 (97.41)	482 (98.77)
M1	13 (1.72)	7 (2.59)	6 (1.23)
AJCC/UICC stage, n (%)			
I	38 (4.94)	18 (6.67)	20 (4.01)
II	133 (17.30)	37 (13.70)	96 (19.24)
III	141 (18.34)	37 (13.70)	104 (20.84)
IV	457 (59.43)	178 (65.93)	279 (55.91)

HNSCC, head and neck squamous cell carcinoma; GEO, Gene Expression Omnibus; TCGA, The Cancer Genome Atlas; AJCC, American Joint Committee on Cancer; UICC, Union for International Cancer Control.

surgery prior to July 2025, while the remaining 9 prospective cases were collected from patients treated from April 2026 to June 2026.

Inclusion criteria were as follows: i) Newly diagnosed patients without prior radiotherapy, chemotherapy or other anti-cancer therapies; ii) Surgically resectable tumors; iii) Complete clinical and pathological data; iv) Patients who were informed of the study content and provided written informed consent. Exclusion criteria were: i) Histologically non-squamous cell carcinoma; ii) Recurrent or metastatic disease; iii) Prior anti-cancer treatment before surgery; iv) Concurrent malignancies; v) Severe organ dysfunction or uncontrolled systemic diseases; vi) Patients who declined to receive study information or refused to provide informed consent.

Immunohistochemistry (IHC). Cathepsin G (CTSG) rabbit anti-human polyclonal antibody (cat. no. ab282105) was purchased from Abcam Inc. Galectin 10 (CLC) polyclonal antibody (cat. No. 25225-1-AP) was purchased from Proteintech Group, Inc. The cerebellin 2 (CBLN2) antibody (cat. no. 34571) was purchased from Signalway Antibody LLC. The CD3 antibody (ZSJQ Bio Co., cat. no. Zm-0417) and associated reagents were used as prediluted, ready-to-use reagents. The Rabbit Two-Step Immunohistochemistry kit (cat. no. PV-6001) was purchased from ZSJQ Bio Co.

Immunohistochemical staining was performed according to the instructions as follows: Tissue specimens were fixed in 4% formaldehyde solution at 4˚C for 24 h. Paraffin-embedded tissues were sliced to ~4 μm thickness and baked at 68˚C for 1 h. Dewaxing and rehydration were performed using an automated system with xylene and graded ethanol. Antigen retrieval was performed at 110˚C for 150 sec. Following dewaxing and antigen retrieval, tissue sections were washed three times with PBS (5 min each), incubated in 3% hydrogen peroxide for 15 min to quench endogenous peroxidase activity and washed again three times with PBS (5 min each). Sections were then blocked with 5% BSA for 2 h at room temperature prior to primary antibody incubation. Sections were subsequently incubated for 16 h with primary antibodies (CTSG, 1:5,000; CLC, 1:250; CBLN2, 1:100; CD3,1:50) at 4˚C, followed by incubation with a HRP-labeled secondary antibody provided in

Table II. Clinical characteristics of GSE65858 samples and TCGA samples in different subsets.

Variable	H (n=135)	L (n=135)	H (n=257)	L (n=256)
Overall survival time, mean ± SD	845.06±434.42	921.00±466.91	772.95±711.64	1073.73±987.71
Age, mean ± SDyears	61.17±10.40	59.08±10.22	61.34±11.97	60.59±11.75
Survival status, n (%)				
Alive	79 (58.52)	97 (71.85)	127 (49.42)	170 (66.41)
Dead	56 (41.48)	38 (28.15)	130 (50.58)	86 (33.59)
Gender, n (%)				
Female	25 (18.52)	22 (16.30)	74 (28.79)	60 (23.44)
Male	110 (81.48)	113 (83.70)	183 (71.21)	196 (76.56)
T (%)				
T1	13 (9.63)	22 (16.30)	12 (4.82)	23 (9.27)
T2	32 (23.70)	48 (35.56)	66 (26.51)	83 (33.47)
T3	34 (25.19)	24 (17.78)	67 (26.91)	66 (26.61)
T4	56 (41.48)	41 (30.37)	104 (41.77)	76 (30.65)
N, n (%)				
N0	47 (34.81)	47 (34.81)	123 (50.20)	119 (48.37)
N1	12 (8.89)	20 (14.81)	38 (15.51)	44 (17.89)
N2	70 (51.85)	61 (45.19)	80 (32.65)	80 (32.52)
N3	6 (4.44)	7 (5.19)	4 (1.63)	3 (1.22)
M, n (%)				
M0	130 (96.30)	133 (98.52)	239 (98.76)	243 (98.78)
M1	5 (3.70)	2 (1.48)	3 (1.24)	3 (1.22)
AJCC/UICC stage, n (%)				
I	8 (5.93)	10 (7.41)	7 (2.80)	13 (5.22)
II	16 (11.85)	21 (15.56)	48 (19.20)	48 (19.28)
III	18 (13.33)	19 (14.07)	53 (21.20)	51 (20.48)
IV	93 (68.89)	85 (62.96)	142 (56.80)	137 (55.02)

H, high-risk subset; L, low-risk subset; GEO, Gene Expression Omnibus; TCGA, The Cancer Genome Atlas; AJCC, American Joint Committee on Cancer; UICC, Union for International Cancer Control.

the Rabbit Two-Step Immunohistochemistry kit for 20 min at 37˚C. DAB was used for color visualization, including 90 sec for CTSG and 120 sec for CBLN2, CLC and CD3. Slides were then rinsed in distilled water to terminate the reaction. The sections were counterstained with hematoxylin at room temperature for 50 secs. Following dehydration and clearing (through a graded ethanol series and xylene), sections were mounted with a neutral resin. The stained sections were then digitally scanned using a Jiangfeng KF-PRX-040 digital pathology slide scanner (light microscope). Cell counting was performed and expression scores were determined using ImageJ (National Institutes of Health, version 1.54g).

Immunohistochemical results quantification. Quantification of CTSG, CBLN2, CLC and CD3 expression in tissue sections was performed using a composite scoring system. This score was derived from the product of two independently assessed parameters: i) Proportion of immunopositive cells, scored as follows: 1 point for ≤10%, 2 points for 10-50%, 3 points for 50-75% and 4 points for >75%; and ii) the staining intensity, scored as: 0 points for negative (-), 1 points for weak (+), 2 points for moderate (++) and 3 points for strong (+++). The final expression score for each marker within each tissue sample was then calculated as the product of its respective proportion score and intensity score.

Results

Identification of two immune subtypes with distinct prognosis in HNSCC. After calculating the immune cell infiltration score (IIs) of 28 immune cell types in the training dataset samples using ssGSEA, a clustering analysis was conducted utilizing 'ConsensusClusterPlus' (Fig. 1A). The k-means clustering method was applied, using a distance metric based on the 1-Pearson correlation coefficient. In addition, 80% of the dataset was resampled in every iteration, amounting to a total of 100 iterations. According to the CDF, the most suitable number of clusters was K=2, which exhibited the highest within-cluster average consensus (Fig. 1B). The heatmap presented in Fig. S1 shows the IIs of the training dataset samples, as well as age and sex clinical data, for C2 (immunity-high) and C1 (immunity-low; Fig. S1).

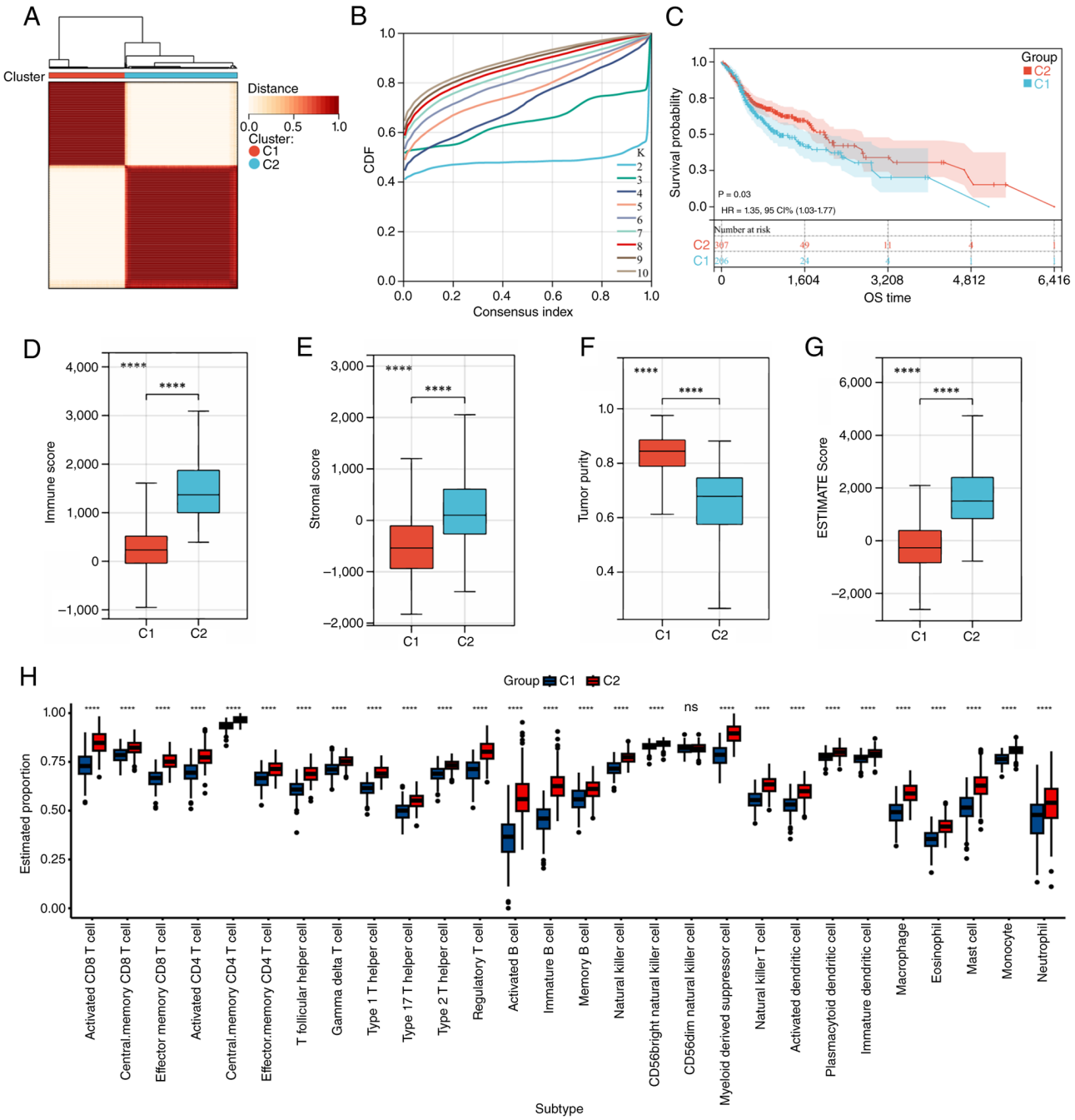

Figure 1. Clusters identified by single-sample gene set enrichment analysis. (A) Sample subtype analysis cluster diagram. (B) CDF curve of The Cancer Genome Atlas samples. (C) Kaplan-Meier survival curves of the C2 (immunity-high) and C1 (immunity-low). (D) Immune score in two clusters. (E) Comparison of the Stroma Score in two clusters. (F) Tumor Purity in two clusters. (G) Comparison of the ESTIMATE Score in two clusters. (H) Box plot showing a statistical difference in immune cell infiltration between the two clusters. ****P<0.0001. OS, overall survival; HR, hazard ratio; ESTIMATE, Estimation of Stromal and Immune Cells in Malignant Tumors using Expression data; CDF, cumulative distribution function.

All assessed immune cell populations, except CD56-low natural killer cells, exhibited a higher level of infiltration in Cluster 2 (Fig. 1H). The ESTIMATE score and tumor purity between the two clusters were significantly different (Fig. 1D-G). Cluster 2 displayed a higher immune score and stroma score, concurrently accompanied by a lower tumor purity level. K-M survival analysis showed that patients in Cluster 2 exhibited significantly improved overall survival compared with those in the Cluster 1 (Fig. 1C), which indicated that the high immune infiltration subtype was associated with improved clinical outcomes.

Differentially expressed ECM genes and immunity. A volcano plot of the 1,573 DEGs, including 339 upregulated genes and 1234 downregulated genes, is shown in Fig. 2A. Gene set enrichment analysis (version 3.0) was performed to investigate the differentially enriched hallmark gene sets between Cluster 1 and Cluster 2. Genes highly expressed in Cluster 2 were significantly enriched in 7 Gene Ontology (GO) and 25 common Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (FDR<0.25), including cytokine-cytokine receptor interaction, T-cell receptor signaling pathway, pathways in cancer, apoptosis and kinase binding protein domain specific

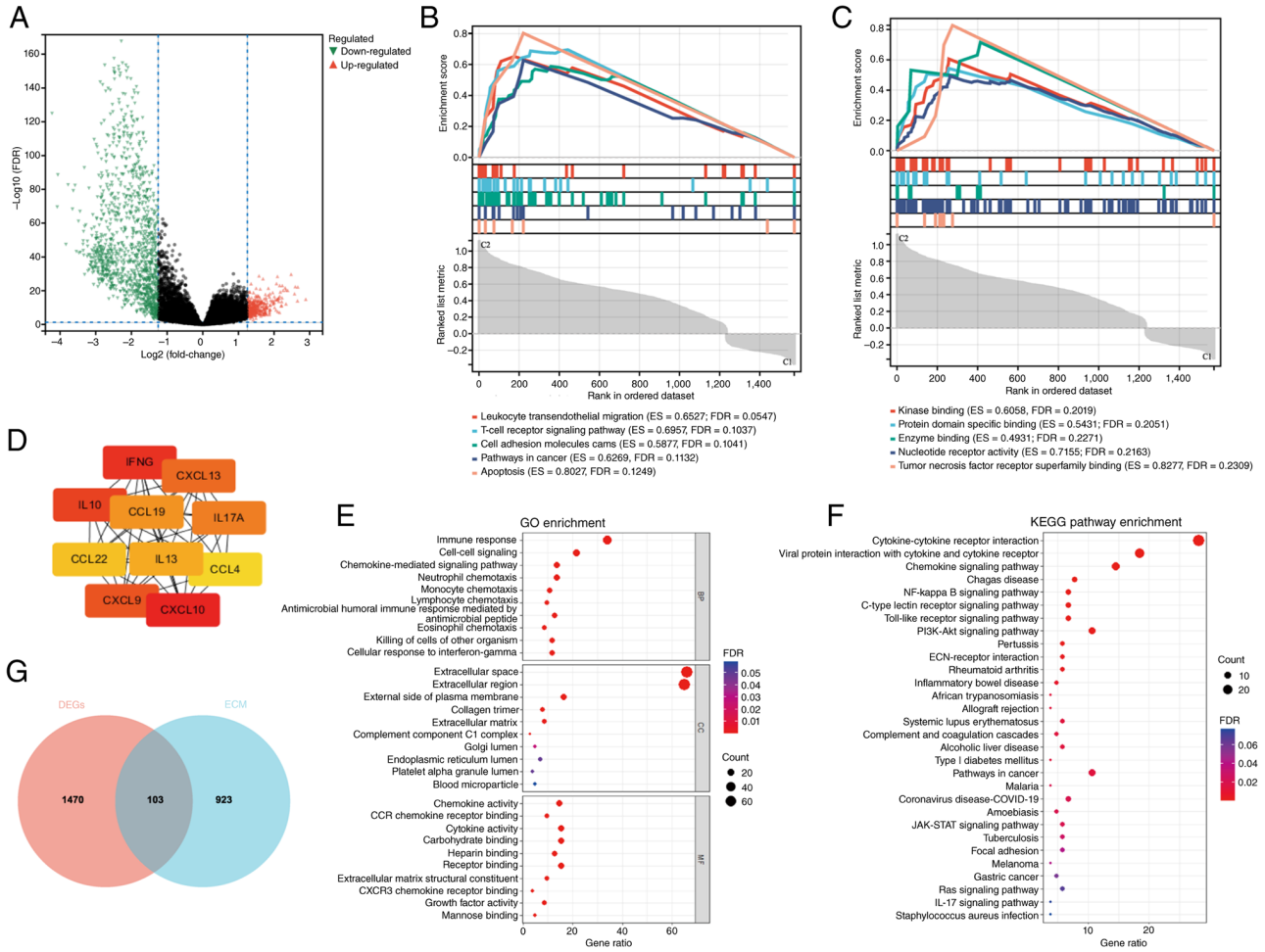

Figure 2. Analysis of DEGs. (A) Volcano plot of the DEGs. (B) GSEA snapshots of the top 5 KEGG terms enrichment analysis. (C) Top 5 GO pathways enrichment analysis. (D) Top 10 hub genes identified using the Maximal Clique Centrality algorithm from 'cytoHubba'. Bubble maps of (E) GO and (F) KEGG signaling pathway enrichment analysis associated with the 103 significantly DEGs. (G) Venn diagram identifying overlap genes of the 1,573 DEGs and 1,026 ECM genes. DEG, differentially expressed genes; FDR, false discovery rate; ECM, extracellular matrix; GSEA, Gene Set Enrichment Analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology.

binding. The top 5 GO and the top 5 KEGG signaling pathways were selected for visualization (Fig. 2B and C).

The Venn diagram indicates 103 overlapping genes between the DEGs and 1,026 ECM genes (Fig. 2G). A total of 10 hub genes (CXCL10, CXCL9, IL13, CCL22, CXCL13, CCL4, IL17A, IFNG, IL10 and CCL19) were identified by 'cytoHubba' in Cytoscape software (version 3.7.2; Fig. 2D). All 10 hub genes were immune cytokines or chemokines, which are derived from the ECM-associated gene set, indicating that the ECM exerts its regulatory role in HNSCC through these immune-associated factors. These 103 overlapping genes, including 12 upregulated and 91 downregulated, were significantly enriched in 61 biological process (BP) terms, 10 cellular component (CC) terms, 33 molecular function (MF) terms and 40 KEGG signaling pathways (FDR<0.25). The BPs were primarily associated with immune response, lymphocyte chemotaxis and chemokine-mediated signaling pathway, the CCs were mainly associated with the extracellular space, extracellular region for example, and the MFs were mostly associated with chemokine activity, CXCR3 chemokine receptor binding, cytokine activity and similar. The top 10, after ranking the FDR, enriched BPs, CCs and MFs were selected for visualization and the top 30 KEGG signaling pathways were selected for visualization (Fig. 2E and F). All GO and the KEGG signaling pathways results can be found in Table SI.

Construction and validation of prognostic models. To screen potential genes associated with prognosis, the LASSO-Cox regression method was utilized. This approach integrated survival time, survival status and gene expression data, enabling authors to screen for genes that have a profound impact on prognosis (25). After performing 10-fold cross-validation, it was found with high probability that 17 genes were involved (ANGPTL1, CBLN2, CCL22, CHRDL1, CLC, CLEC2D, CLEC4C, CLEC9A, CTSG, F13A1, FGF7, IL17A, PAPPA2, RSPO1, SFTPA2, SFTPC and SPOCK2; Fig. 3A). A prognostic risk signature incorporating three genes (CBLN2, CLC and CTSG) was derived through the conduction of a multivariate Cox regression analysis. The hazard ratio (HR) for the three prognostic genes was as follows (Fig. 3B): CBLN2 (HR=0.77), CTSG (HR=0.79), and CLC (HR=1.27). According to standard prognostic criteria, an HR<1 indicates a protective effect: CBLN2 and CTSG, with their higher expression was

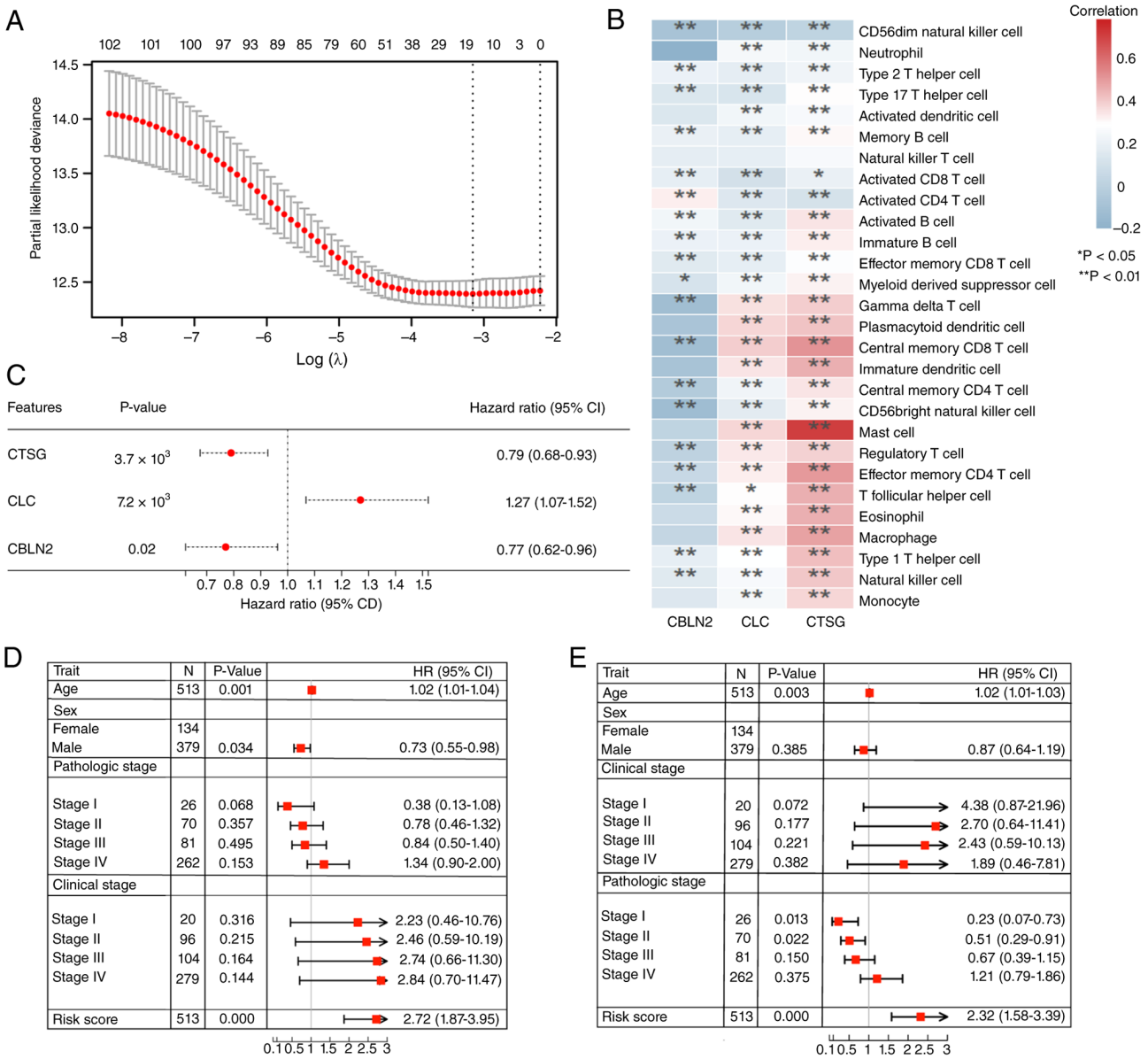

Figure 3. Screening for the potential ECM genes associated with prognosis. (A) Plot showing the cross-validation curve along with upper and lower standard deviation. The left vertical line indicates the gene number when the value of λ gives a minimal mean squared error. Based on the first vertical line, the model included 17 genes with a non-zero prediction value. (B) Heatmap of the correlations between genes and 28 immune cell subsets. (C) Multivariate Cox regression analysis of CBLN2, CLC and CTSG. (D) Univariate and (E) multivariate Cox regression analysis in TCGA-HNSCC. *P<0.05 and **P<0.01. ECM, extracellular matrix; CBLN2, cerebellin 2; CLC, galectin-10; CTSG, cathepsin G; HR, hazard ratio; TCGA, The Cancer Genome Atlas; HNSCC, head and neck squamous cell carcinoma.

associated with improved patient survival (26). By contrast, CLC with HR>1 was a risk prognostic factor and its higher expression was associated with poorer patient survival. Spearman's correlation values between the mRNA expression of CBLN2, CLC, CTSG and 28 immune cell subsets were visualized in a heatmap, using the 'Hmisc' package (version 5.1.1) and 'pheatmap' package (version 1.0.12; Fig. 3C).

The regression coefficients of prognostic genes were determined through multivariate Cox regression analysis. The coefficients were -0.257145966422451 for CBLN2, 0.242592579400633 for CLC and -0.234662389278031 for CTSG. Thus, the risk score was computed as:: Risk score=expCBLN2 x (-0.257145966422451) + expCLC x (0.242592579400633) + expCTSG x (-0.234662389278031) where exp is the normalized expression of the respective gene. Univariate and multivariate Cox regression analyses in the training dataset were employed to investigate the prognostic significance of the risk score (Fig. 3D and E). The Cox analysis of univariate and multivariate was conducted with risk score, age, sex and stage, indicating that the risk score may be an independent risk factor for the prognosis of HNSCC. The median risk score was then used as a threshold, dividing the samples in the training dataset and in the testing dataset into two distinct groups: A high-risk group and a low-risk group. The log-rank test was used to analyze survival differences between the two risk groups. K-M curves were employed to visualize the survival data (Fig. 4A and C). A time-dependent ROC curve was also drawn, to evaluate the predictive

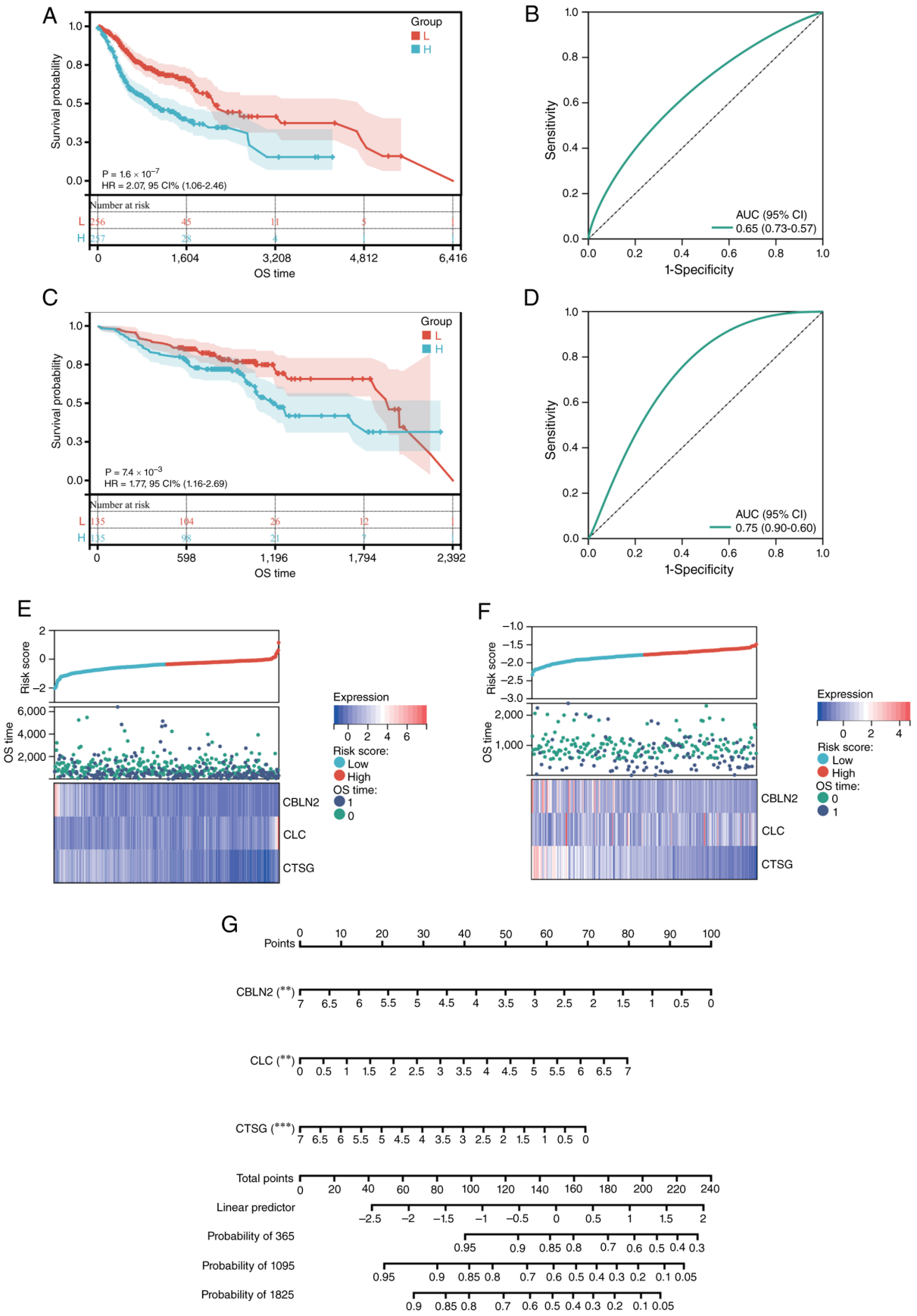

Figure 4. A three gene predictive model and nomogram. (A) Kaplan-Meier survival curves of the training high- and low-risk groups. (B) ROC curve of the training dataset. (C) Kaplan-Meier survival curves of the testing high- and low-risk groups. (D) ROC curve of the testing dataset. (E) Risk score, overall survival status and gene expression profiles of the three ECM genes of the training samples. (F) Risk score, overall survival status and gene expression profiles of the three ECM genes of the testing samples. (G) Nomogram of the training dataset. **P<0.01 and ***P<0.001. L, low risk; H, high risk; OS, overall survival; HR, hazard risk; ROC, receiver operating characteristic; ECM, extracellular matrix; CBLN2, cerebellin 2; CLC, galectin-10; CTSG, cathepsin G; AUC, area under the curve.

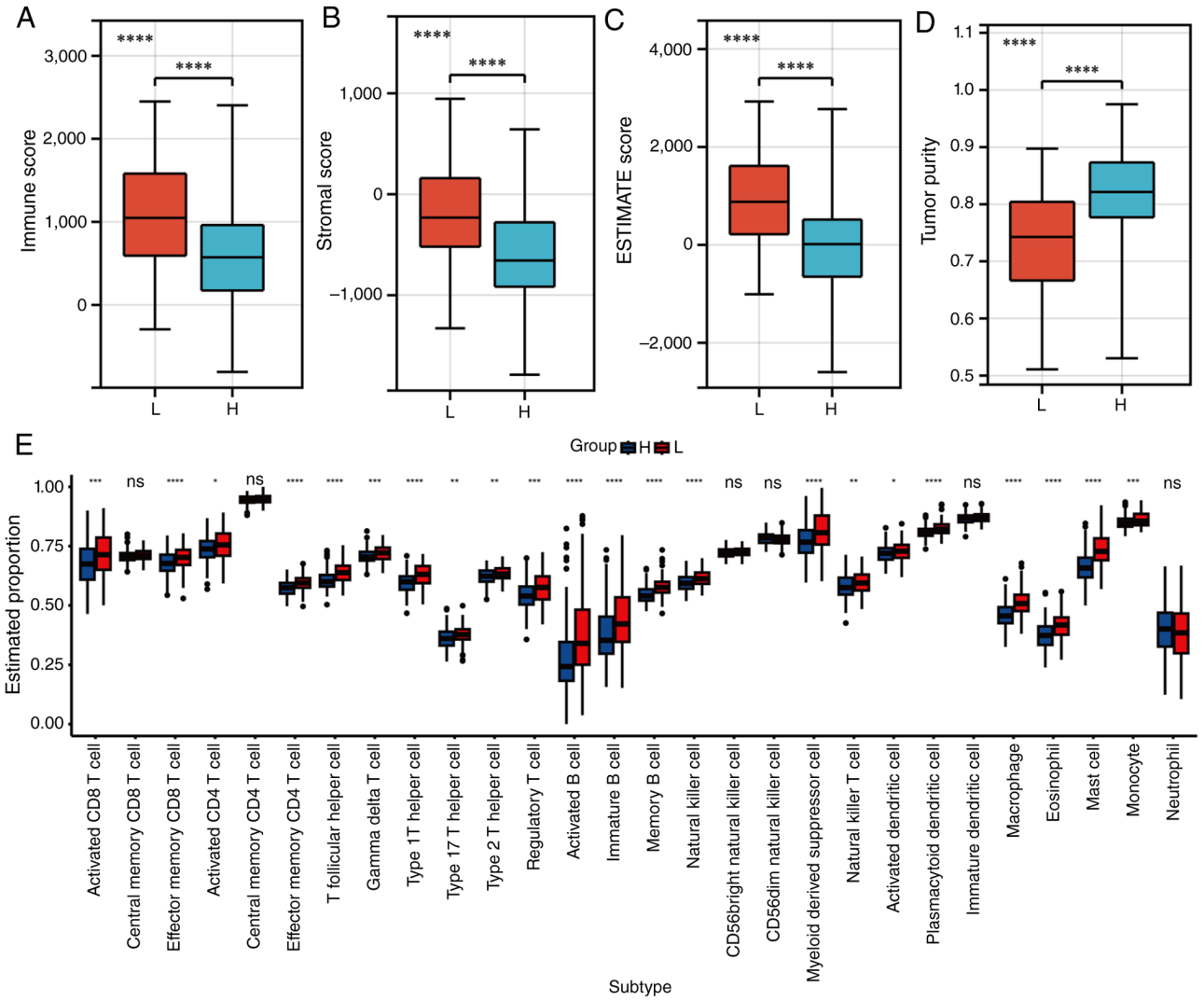

Figure 5. Immune landscape of the test dataset. (A) Immune Score in high- and low-risk groups. (B) Comparison of the Stroma Score in high- and low-risk groups. (C) Comparison of the ESTIMATE Score in high- and low-risk groups. (D) Comparison of the Tumor Purity in high- and low-risk groups. (E) Box plot showing the statistical difference in immune cell infiltration between the two groups. *P<0.05, **P<0.01, ***P<0.001 and ****P<0.0001. L, low risk; H, high risk; ESTIMATE, Estimation of Stromal and Immune Cells in Malignant Tumors using Expression data; ns, not significant.

accuracy of the prognostic models, and the 5-year area under the ROC curve (AUC) value was calculated as 0.65 in the training dataset and 0.75 in the testing dataset, respectively (Fig. 4B and D). The survival time, survival status and three characteristics were integrated to establish a nomogram using the Cox method, which assessed the prognostic significance of these characteristics in 513 samples (Fig. 4G). The results demonstrated that the prognosis for the low-risk group was significantly superior to that of the high-risk group.

The same method as used in the training dataset was employed to evaluate the infiltration levels of the immune cell types, the ESTIMATE score and tumor purity in GSE65858 samples (testing dataset). Samples in the high-risk group exhibited a suppressed TME (Fig. 5).

Based on data from the Clinical Proteomic Tumor Analysis Consortium, CTSG and CLC expression levels were higher in HNSCC samples compared with normal samples. CTSG, CBLN2, CLC and CD3 protein levels were further detected by IHC in 17 paired HNSCC tumor and adjacent non-tumor tissues. Compared with adjacent normal tissues, tumor tissues exhibited significantly higher expression scores for CTSG and CBLN2, whereas CLC expression did not differ notably between groups. Consistent with this, IHC staining revealed that regions with high CTSG or CBLN2 expression frequently exhibited abundant $CD3^{+}$ T cell infiltration, while regions with high CLC expression tended to show lower levels of $CD3^{+}$ T cells. However, owing to the limited sample size (n=17), these results lack statistical power and should be interpreted as preliminary observations. All results are presented in Fig. 6, and the clinical characteristics and expression scores of the 17 samples are provided in Table SII.

Discussion

Immunotherapy has notably advanced cancer treatment by manipulating the immune system to improve its recognition and destruction of malignant cells. Therapies such as antitumoral cytokines and chimeric antigen receptor (CAR)-modified immune cells have exhibited marked therapeutic success in clinical outcomes (27-29). In HNSCC, EGFR-chimeric

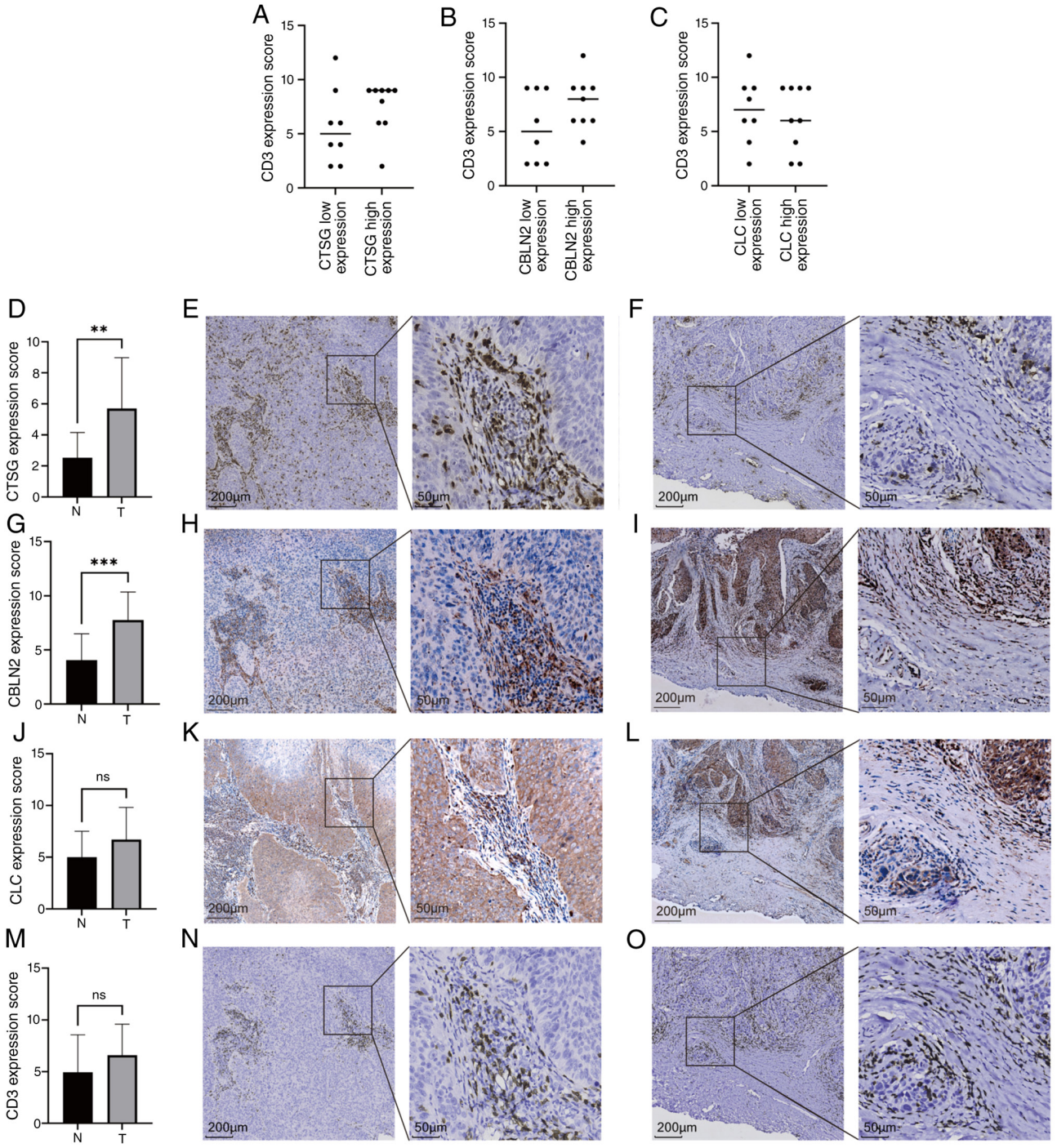

Figure 6. Immunohistochemistry analysis of clinical HNSCC samples. Statistics of the CD3 expression score in the (A) CTSG, (B) CBLN2 and (C) CLC high expression group and low expression group. (D) CTSG expression score of 17 clinical HNSCC samples (P<0.01). (E) CTSG expression in HNSCC tissue (patient B) by IHC. (F) CTSG expression in HNSCC tissue (patient O) by IHC. (G) CBLN2 expression score of 17 clinical HNSCC samples (P<0.001). (H) CBLN2 expression in HNSCC tissue (patient B) by IHC (Scale bar: 200 μm; 50 μm). (I) CBLN2 expression in HNSCC tissue (patient O) by IHC (Scale bar: 200 μm; 50 μm). (J) CLC expression score of 17 clinical HNSCC samples. (K) CLC expression in HNSCC tissue (patient B) by IHC (Scale bar: 200 μm; 50 μm). (L) CLC expression in HNSCC tissue (patient O) by IHC (Scale bar: 200 μm; 50 μm). (M) CD3 expression score of 17 clinical HNSCC samples. (N) CD3 expression in HNSCC tissue (patient B) by IHC (Scale bar: 200 μm; 50 μm). (O) CD3 expression in HNSCC tissue (patient O) by IHC (Scale bar: 200 μm; 50 μm). **P<0.01 and ***P<0.001. ns, not significant; HNSCC, head and neck squamous cell carcinoma; CBLN2, cerebellin 2; CLC, galectin-10; CTSG, cathepsin G.

antigen receptor (CAR)-T cells (genetically engineered T cells for specific killing of EGFR-positive tumor cells) have been constructed and validated (30). Haist *et al* (31) engineered an optimized cetuximab-based CAR construct with high transduction efficiency and EGFR-independent cytotoxicity (31). In addition, anti-CD44v6 CAR-natural killer cells have demonstrated a marked improvement in cytotoxicity, showing a killing efficiency that is 2-3 times higher compared with that of cytokine-expanded primary natural killer cells when tested against a number of HNSCC cell lines (32). Yang *et al* (33) reported that the combined inhibition of IL-6 and CCR2 pathways markedly enhanced the ability of natural killer cells to

fight tumor cells in human papillomavirus-negative HNSCC. Despite this, immunotherapy still faces numerous challenges, such as the scarcity of tumor-specific antigens, notable tumor heterogeneity, immunosuppressive TME and the potential for treatment-associated toxicities in the context of solid tumor treatment (34). Therefore, it is important to identify 'hot' tumors that are more likely to respond to immunotherapy. The present study aimed to predict the prognosis and the infiltration of the immune cells of HNSCC based on accurate models which may facilitate the provision of more personalized immunotherapy interventions for patients with HNSCC. A three gene (CBLN2, CLC and CTSG) predictive model was constructed which can evaluate the immune infiltration with stability and effectiveness.

The ECM acts as a reservoir for cytokines and chemokines, which are secreted by immune cells infiltrating damaged or infected regions, thereby establishing a chemoattractive and immune-modulating gradient that both recruits and activates additional immune cells (35). Following the application of the ssGSEA approach to evaluate the immune cell profiles in patients with HNSCC from TCGA, a total of 1,573 DEGs were pinpointed. GSEA suggested the DEGs may affect the development of HNSCC and the changes in the immune microenvironment through intercellular signaling pathways, including cytokine-cytokine receptor interaction and T cell receptor signaling pathway. Additionally, 103 ECM-associated DEGs were identified between two clusters, which have an impact on immune infiltration in HNSCC. KEGG pathway analysis revealed that the 103 ECM-associated DEGs were implicated in cytokine-cytokine receptor interaction pathways, suggesting that ECM genes serves a role in regulating the immune response through cytokines in HNSCC.

In the present study, three ECM-associated genes were identified as risk measures. Among them, CTSG reversibly binds to and activates B cells, as well as $CD4^+$ and $CD8^+$ T cells. It also reversibly binds to natural killer cells and augments their cytotoxicity by exerting its protease activity (36,37). In colorectal cancer, CTSG has been found to suppress tumor growth by inactivating anti-apoptotic signaling pathways (38). CTSG is a potential immune-associated biomarker, functioning as an independent biomarker and possibly serving as an immunotherapeutic target in oral squamous cell carcinoma treatment (39). Recently, immune infiltration and apoptosis have been found to be positively associated with CTSG expression in HNSCC (40,41). CBLN2 was hypothesized to serve a key role in preserving synaptic structure integrity and facilitating spontaneous synaptic transmission. According to Zhu *et al* (42), CBLN2, along with transmembrane protein 220, are prospective prognostic indicators in colorectal adenocarcinoma cases. Separately, a study by Belotti *et al* (43) identified a panel consisting of 19 genes (such as CBLN2, chondroadherin and MMP17) associated with the TME with prognostic relevance in high-grade serous ovarian cancer. The high expression of CLC is positively associated with a higher risk score, suggesting that this gene may be a risk factor for sarcoma development and clinical deterioration. This may be attributed to its ability to orchestrate immune responses and it's pivotal role in maintaining the suppressive capabilities of CD25-positive regulatory T cells (44). The three gene (CBLN2, CLC and CTSG) predictive model that was built demonstrated stability and effectiveness and has been further validated in independent external datasets. In the present study, CTSG and CBLN2 served as tumor suppressor genes in HNSCC and were associated with the infiltration of immune cells. Risk score can predict the infiltration of immune cells in the TME.

CTSG is secreted by monocytes and activated neutrophils (36). In the present IHC analysis of 17 paired HNSCC tissues, CTSG expression was significantly elevated in tumor compared with adjacent normal tissues, and CTSG-high regions tended to colocalize with more abundant $CD3^+$ T cell infiltration. This observation suggests a possible positive association between CTSG expression and $CD3^+$ T cell density in the tumor microenvironment. Similarly, CBLN2 expression was found to be significantly higher in tumor tissues and regions with strong CBLN2 staining also frequently exhibited increased $CD3^+$ T cell infiltration. CBLN2 encodes cerebellin 2 precursor, a member of the C1q/TNF superfamily of secreted glycoproteins. Previous studies have largely focused on its role in the neuronal synaptic cleft; however, its function in the tumor microenvironment is gradually being uncovered. CBLN2 contains a C1q-like domain (45), and C1q is closely involved in immune opsonization and synaptic pruning (46), suggesting a potential role for CBLN2 in mediating communication between tumor cells and immune cells. By contrast, CLC expression did not differ notably between tumor and normal tissues, and high CLC expression areas were more often accompanied by relatively low $CD3^+$ T cell density, hinting at a potential inverse relationship with immune infiltration. However, for all three genes, the statistical power to compare CD3 scores between high- and low-expression groups was insufficient due to the limited sample size (n=17). Therefore, these observations should be regarded as preliminary and interpreted with caution. Whether elevated CTSG or CBLN2 promotes immune cell recruitment, or whether infiltrating immune cells in turn stimulate their expression, and whether CLC serves an immunosuppressive role in HNSCC, remain to be explored in larger cohorts with appropriate functional experiments.

A notable limitation of the present study lies in the relatively modest AUC value for survival prediction in the prognostic model that was constructed. This implies that the predictive model does not offer exceptionally precise forecasts of the survival durations of patients. Despite this, the prognostic model demonstrated a survival difference between high- and low-risk groups in the independent external validation cohort. In addition, it effectively differentiated the tumor immune microenvironment among patients. The present prognostic model can therefore hopefully serve as a reference for determining whether patients are suitable candidates for further immunotherapy and that it may facilitate the exploration of novel therapeutic targets based on ECM research. It should also be noted that the cohorts in the present study were all derived from publicly available datasets, lacking validation through entirely new clinical cohorts. Furthermore, prospective predictive data from clinical settings were unavailable, but further validation using clinical data is anticipated.

Notably, the observed correlation between the three genes and CD3 was derived from a small cohort (n=17), which inherently limits statistical power and may introduce potential sampling bias. Therefore, this finding should be interpreted

with caution, as the limited sample size makes it difficult to fully validate the reliability and reproducibility of this association. To address this limitation and further elucidate the biological functions of the target genes, a follow-up research plan was formulated. This plan includes two main components: First, *in vitro* gene knockout experiments will be conducted to validate the regulatory roles of the target genes (CBLN2, CLC and CTSG) in immune cell recruitment; second, the clinical cohort will be expanded by constructing a new, independent cohort of patients with HSCC, enabling a more robust exploration of the association between CTSG and immune cell infiltration in a larger population.

It should also be noted that the ESTIMATE algorithm has been reported to be influenced by stromal abundance and tumor heterogeneity (16). In the present study, immune cell infiltration was primarily assessed by ssGSEA, with IHC providing additional validation, and the clustering of patients was based on this ssGSEA methodology. ESTIMATE was subsequently used to verify that the clusters differed in their immune and stromal content. However, given the known limitations of ESTIMATE (16), these confirmatory results should be interpreted with caution. The association between the prognostic model, the involved ECM-associated genes, and immune infiltration still requires further validation in independent clinical cohorts or animal models.

In conclusion, a prognostic model was constructed and the risk score showed a marked associated between the prognosis of HNSCC and the immune microenvironment. This evaluation of the progression and TME may facilitate the provision of more personalized immunotherapy interventions for patients with HNSCC. CTSG acted as tumor suppressor genes in HNSCC and exhibited an association with immune cell infiltration, which suggested a potential approach to enhance the efficacy of immunotherapy in HNSCC.

Acknowledgements

The authors would like to thank Mr. Jiayu Yang (Center for Basic Research, The Second Affiliated Hospital of Zhejiang University, Hangzhou, China) for help in sample processing, technical guidance on immunohistochemistry and analysis. The authors would also like to thank Mr. Haolong Shi (Center for Basic Research, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China) for his technical support in pathological experiments, including paraffin embedding, section preparation and tissue microscopic imaging.

Funding

The present study was supported by the Joint TCM Science & Technology Projects of National Demonstration Zones for Comprehensive TCM Reform (grant no. GZY-KJS-ZJ-2025-009) and the Natural Science Foundation of Zhejiang Province, China (grant no. LY23H130002).

Availability of data and materials

The data generated in the present study may be requested from the corresponding author.

Authors' contributions

LT and MC confirm the authenticity of all the raw data. LT designed the study, analyzed and interpreted data, constructed figures, and wrote and revised the manuscript. YQ designed the study and revised the manuscript. HZ contributed to sample collection and preliminary processing of the specimens. HZ also contributed to acquisition of data. BF designed the study and constructed figures. MC designed the study and edited the manuscript. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

The present study was approved by the ethics committee of The Second Affiliated Hospital, Zhejiang University School of Medicine (Hangzhou, China; approval no. 2025-0873) and was conducted in strict conformity with the ethical guidelines of the Declaration of Helsinki.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71: 209-249, 2021.
2. Williamson A, Burton S, Rajan Z, Kwong E, Tengku S, Anthony G, Montgomery J, Devabalan Y, Winter SC, Moen C, *et al*: Salvage surgery for residual and recurrent head and neck squamous cell carcinoma (RESCUE): An IReC multicentre consecutive cohort study. Oral Oncol 179: 107997, 2026.
3. Gabizon-Peretz S and Kluger HM: Advances in cancer immunotherapy for solid tumors. Adv Sci (Weinh) 13: e76020, 2026.
4. Han M, Lou J, Pan M, Xu S, Jiang X and Jing W: Engineering immune cells for next-generation therapies. J Control Release 395: 115012, 2026.
5. Jiang Y, Zhang H, Wang J, Liu Y, Luo T and Hua H: Targeting extracellular matrix stiffness and mechanotransducers to improve cancer therapy. J Hematol OncolJ Hematol Oncol 15: 34, 2022.
6. Harper J and Sainson RC: Regulation of the anti-tumour immune response by cancer-associated fibroblasts. Semin Cancer Biol 25: 69-77, 2014.
7. Huang J, Zhang L, Wan D, Zhou L, Zheng S, Lin S and Qiao Y: Extracellular matrix and its therapeutic potential for cancer treatment. Signal Transduct Target Ther 6: 153, 2021.
8. Pires A, Burnell S and Gallimore A: Exploiting ECM remodelling to promote immune-mediated tumour destruction. Curr Opin Immunol 74: 32-38, 2022.
9. El Herch I, Tornaas S, Dongre HN and Costea DE: Heterogeneity of cancer-associated fibroblasts and tumor-promoting roles in head and neck squamous cell carcinoma. Front Mol Biosci 11: 1340024, 2024.
10. Spenlé C, Loustau T, Murdamoothoo D, Erne W, Beghelli-de la Forest Divonne S, Veber R, Petti L, Bourdely P, Mörgelin M, Brauchle EM, *et al*: Tenascin-C orchestrates an immune-suppressive tumor microenvironment in oral squamous cell carcinoma. Cancer Immunol Res 8: 1122-1138, 2020.
11. Gavrielatou N, Vathiotis I, Aung TN, Shafi S, Burela S, Fernandez AI, Moutafi M, Burtness B, Economopoulou P, Anastasiou M, *et al*: Digital spatial profiling links beta-2-microglobulin expression with immune checkpoint blockade outcomes in head and neck squamous cell carcinoma. Cancer Res Commun 3: 558-563, 2023.

12. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA and Hynes RO: The matrisome: In Silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. Mol Cell Proteomics 11: M111.014647, 2012.
13. Wichmann G, Rosolowski M, Krohn K, Kreuz M, Boehm A, Reiche A, Scharrer U, Halama D, Bertolini J, Bauer U, *et al*: The role of HPV RNA transcription, immune response-related gene expression and disruptive TP53 mutations in diagnostic and prognostic profiling of head and neck cancer. Int J Cancer 137: 2846-2857, 2015.
14. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C, *et al*: Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462: 108-112, 2009.
15. Wilkerson MD and Hayes DN: ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. Bioinformatics 26: 1572-1573, 2010.
16. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, *et al*: Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 4: 2612, 2013.
17. Wickham H: ggplot2: Elegant Graphics for Data Analysis. Springer Cham, New York, NY, 2016.
18. Love MI, Huber W and Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550, 2014.
19. Naba A, Clauser KR, Ding H, Whittaker CA, Carr SA and Hynes RO: The extracellular matrix: Tools and insights for the 'Omics' Era. Matrix Biol 49: 10-24, 2016.
20. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, Imamichi T and Chang W: DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res 50: W216-W221, 2022.
21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102: 15545-15550, 2005.
22. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY: cytoHubba: Identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 8 (Suppl 4): S11, 2014.
23. Engebretsen S and Bohlin J: Statistical predictions with glmnet. Clin Epigenetics 11: 123, 2019.
24. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC and Müller M: pROC: An open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12: 77, 2011.
25. Tibshirani R: The lasso method for variable selection in the Cox model. Stat Med 16: 385-395, 1997.
26. Hazra A and Gogtay N: Biostatistics series module 8: Assessing risk. Indian J Dermatol 62: 123-129, 2017.
27. Safaei S, Yari A, Pourbagherian O and Maleki LA: The role of cytokines in shaping the future of Cancer immunotherapy. Cytokine 189: 156888, 2025.
28. Bayat M and Nahand JS: CAR-engineered cell therapies: Current understandings and future perspectives. Mol Biomed 7: 7, 2026.
29. Wang M, Wang Z, Zhang G and Fan J: Interleukin-enhanced CAR-engineered immune cells in tumor immunotherapy: Current insights and future perspectives. Cytokine 192: 156973, 2025.
30. Dong YH, Ding YM, Guo W, Huang JW, Yang Z, Zhang Y and Chen XH: The functional verification of EGFR-CAR T-cells targeted to hypopharyngeal squamous cell carcinoma. OncoTargets Ther 11: 7053-7059, 2018.
31. Haist C, Poschinski Z, Bister A, Hoffmann MJ, Grunewald CM, Hamacher A, Kassack M, Wiek C, Scheckenbach K and Hanenberg H: Engineering a single-chain variable fragment of cetuximab for CAR T-cell therapy against head and neck squamous cell carcinomas. Oral Oncol 129: 105867, 2022.
32. Ciulean IS, Fischer J, Quaiser A, Bach C, Abken H, Tretbar US, Fricke S, Koehl U, Schmiedel D and Grunwald T: CD44v6 specific CAR-NK cells for targeted immunotherapy of head and neck squamous cell carcinoma. Front Immunol 14: 1290488, 2023.
33. Yang F, Yuan C, Chen F, Qin ZS, Schmitt NC, Lesinski GB, Saba NF and Teng Y: Combined IL6 and CCR2 blockade potentiates antitumor activity of NK cells in HPV-negative head and neck cancer. J Exp Clin Cancer Res 43: 76, 2024.
34. Safarzadeh Kozani P, Safarzadeh Kozani P and Rahbarizadeh F: Addressing the obstacles of CAR T cell migration in solid tumors: Wishing a heavy traffic. Crit Rev Biotechnol 42: 1079-1098, 2022.
35. McQuitty CE, Williams R, Chokshi S and Urbani L: Immunomodulatory role of the extracellular matrix within the liver disease microenvironment. Front Immunol 11: 574276, 2020.
36. Hase-Yamazaki T and Aoki Y: Stimulation of human lymphocytes by cathepsin G. Cell Immunol 160: 24-32, 1995.
37. Yamazaki T and Aoki Y: Cathepsin G binds to human lymphocytes. J Leukoc Biol 61: 73-79, 1997.
38. Chan S, Wang X, Wang Z, Du Y, Zuo X, Chen J, Sun R, Zhang Q, Lin L, Yang Y, *et al*: CTSG suppresses colorectal cancer progression through negative regulation of Akt/mTOR/Bcl2 signaling pathway. Int J Biol Sci 19: 2220-2233, 2023.
39. Huang GZ, Wu QQ, Zheng ZN, Shao TR, Li F, Lu XY, Ye HY, Chen GX, Song YX, Zeng WS, *et al*: Bioinformatics analyses indicate that cathepsin G (CTSG) is a potential immune-related biomarker in oral squamous cell carcinoma (OSCC). OncoTargets Ther 14: 1275-1289, 2021.
40. Shen Y, Chen H, Gong X, Wang Z, Chen M and Chen D: Identification of lysosome-related genes in connection with prognosis and immune cell infiltration for drug candidates in head and neck cancer. Open Life Sci 18: 20220660, 2023.
41. Hua H, Yang X, Meng D, Gan R, Chen N, He L, Wang D, Jiang W, Si D, Wang X, *et al*: CTSG restraines the proliferation and metastasis of head and neck squamous cell carcinoma by blocking the JAK2/STAT3 pathway. Cell Signal 127: 111562, 2025.
42. Zhu L, Sun H, Tian G, Wang J, Zhou Q, Liu P, Tang X, Shi X, Yang L and Liu G: Development and validation of a risk prediction model and nomogram for colon adenocarcinoma based on methylation-driven genes. Aging (Albany NY) 13: 16600-16619, 2021.
43. Belotti Y, Lim EH and Lim CT: The role of the extracellular matrix and tumor-infiltrating immune cells in the prognostication of high-grade serous ovarian cancer. Cancers (Basel) 14: 404, 2022.
44. Kubach J, Lutter P, Bopp T, Stoll S, Becker C, Huter E, Richter C, Weingarten P, Warger T, Knop J, *et al*: Human CD4+CD25+ regulatory T cells: Proteome analysis identifies galectin-10 as a novel marker essential for their anergy and suppressive function. Blood 110: 1550-1558, 2007.
45. Tom Tang Y, Hu T, Arterburn M, Boyle B, Bright JM, Palencia S, Emtage PC and Funk WD: The complete complement of C1q-domain-containing proteins in Homo sapiens. Genomics 86: 100-111, 2005.
46. Bulla R, Tripodo C, Rami D, Ling GS, Agostinis C, Guarnotta C, Zorzet S, Durigutto P, Botto M and Tedesco F: C1q acts in the tumour microenvironment as a cancer-promoting factor independently of complement activation. Nat Commun 7: 10346, 2016.