

# The matrix metalloproteinase gene family: a significant prognostic gene lineage correlated with immune infiltrates in laryngeal squamous cell carcinoma

Chuan YANG, Huan CAO, Jian-Wang YANG, Tao LIU, Jian-Tao WANG, Bao-Shan WANG\*

Department of Otorhinolaryngology, The Second Hospital of Hebei Medical University, Shijiazhuang, Hebei, China

\*Correspondence: [hebwangbs@163.com](mailto:hebwangbs@163.com)

Received May 12, 2021 / Accepted July 1, 2021

This study aimed to elucidate the potential genes of the matrix metalloproteinase (MMP) family, responsible for the progression of laryngeal squamous cell carcinoma (LSCC). Besides, we ascertained the changes in common malignant behaviors *in vitro* by knocking down MMP1. TCGA, GEO, Oncomine, and microarray data were conducted to analyze the expression levels of MMPs and to find tissue-specific genes in LSCC. Univariate and multivariate Cox regression analyses were established in the construction of a prognostic model based on expression profiles and clinical information of LSCC in TCGA. We then comprehensively analyzed survival, co-expression network, and immune infiltration based on a prognostic model by Kaplan-Meier analysis, WGCNA, and CIBERSORT. Thereafter, qRT-PCR, proliferation, Transwell, and wound-healing assays were used to assess the accuracy of the bioinformatics data. A total of seven genes in the MMP family were identified as differentially expressed genes (DEGs) by integrating three public databases and microarray data. Additionally, multivariate Cox regression was used to establish a four-gene (MMP1/3/8/10) prognostic model, which exhibited a better predictive accuracy than the TNM (tumors/nodes/metastases) based model. The prognostic model was related to plasma cells, CD8+ T cells, follicular helper T cells, resting NK cells, and M0 macrophages infiltration. The expression of MMP1, MMP3, and MMP10 was the highest in head and neck squamous cell carcinoma (HNSC) compared to other cancer in the Oncomine and GEPIA dataset. Further, MMP1 demonstrated significant upregulation in 40 paired LSCC tissues. Eventually, MMP1 downregulation inhibited cell viability, colony formation, and cell migration in TU686 and FaDu cells. Our findings suggest that the four-gene signature might be associated with the prognosis. Further, we revealed that MMP1 is a pivotal biomarker for the biotherapy and prognostic evaluation of patients with LSCC.

*Key words: LSCC, MMP family genes, bioinformatics analysis, biomarker, prognosis, MMP1*

Laryngeal cancer is the second most prevalent tumor in otorhinolaryngology – head and neck surgery, with the squamous cell carcinoma subtype accounting for 95% of larynx cancer [1]. The causes of laryngeal cancer are complex and multifaceted. Besides, the significance of early clinical symptoms remains unclear, and survival rates at later stages are still low [2]. With the implementation of smoking bans, the total number of new cases of LSCC has slightly reduced over the past three years. However, the incidence remains high, accounting for approximately 1% of all cancers [3]. Therefore, it is essential to further explore novel therapeutic targets and effective prognostic biomarkers.

Matrix metalloproteinase (MMP) family are highly conserved zinc-dependent endopeptidases and primary proteases participating in the degradation of extracellular matrix (ECM) and basement membranes. A total of 25

MMP genes have been described in humans and classified into six types based on the structural domain composition and substrate specificity. Irregular expression of MMP genes has been implicated in numerous diseases, including cardiovascular diseases, rheumatic diseases, and cancer [4]. Several studies have attributed disordered MMP expression in different types of cancers to poor performance, specifically in breast cancer and ovarian cancer [5].

To the best of our knowledge, no systematic analysis has been reported on the expression and specific tissue distribution of the MMP family genes in different tumors. Moreover, the functions and prognostic capacities of individual MMPs in LSCC are less described. As such, we explored the potential oncogenic values of distinct MMP family members in LSCC. We aimed to identify MMP family genes indicating poor clinical outcomes. We also investigated the correlations of tumor immunity with a prognosis of LSCC,

which might help to evaluate therapeutic decisions and the prognosis for LSCC patients.

## Patients and methods

**Data acquisition from a public database.** FPKM data of the LSCC with 123 samples and relevant clinical data were obtained from The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>). The mRNA levels of MMPs family members in different types of cancer were investigated by Oncomine ([www.oncomine.org](http://www.oncomine.org)). The genes with  $|\log_2\text{fold-chang}(\log_2\text{FC})| > 1$  and adjusted  $p < 0.05$  were regarded as DEGs. Gene matrix expression profiles of LSCC were obtained from Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/gds/?term=>) (GSE51985, GSE59102, GSE59652, GSE84957, GSE117005, GSE143224, see Table 1 for specific information). All genes in each data sets were sorted by  $\log_2\text{FC}$  and integrated using the RobustRankAggreg R package. Given the TCGA database included complete gene expression profiles and clinical information of LSCC, the MMPs-related prognostic prediction analysis was performed using TCGA data. The other databases were only used for differential gene expression analysis.

**Prognosis-related genes identification.** R-package ‘survival’ was used to perform univariate and multifactorial Cox regression analysis. A four-gene-based prognostic model was performed to estimate the risk score of each patient following the equation:  $\text{risk score} = \text{MMP1} \times 0.121 + \text{MMP3} \times 0.040 + \text{MMP8} \times 0.800 + \text{MMP10} \times 0.0709$ .

The patients were then subdivided into a high-risk and a low-risk group based on the median risk score. The receiver operating characteristic (ROC) curve was drawn and the area under curve (AUC) was calculated by R-package ‘survivalROC’. The AUC value indicated the accuracy of prediction and was significant when it exceeded 0.60 [6]. Kaplan-Meier procedure was used to estimate the survival time difference between the low-risk and high-risk groups by R-package ‘Survminer’. Meanwhile, the ‘forestplot’ R-package was used to visualize the Cox results. Nomograms were built by the ‘rms’ package of R. Moreover, the risk-score-based DEGs were analyzed in the high- and the low-risk groups in LSCC using the R-package ‘edgeR’.

**Co-expression network and tumor microenvironment analysis.** Risk-score-based DEGs and MMP family genes were evaluated via the Weighted Gene Co-Expres-

sion Network Analysis (WGCNA) following the tutorial on the official website [7]. The networks were visualized by Cytoscape 3.7.2. Stromal score, immune score, estimate score, and tumor purity were analyzed by R package ‘estimate’. Immune infiltration data were estimated by CIBERSORT (<https://cibersort.stanford.edu>).

**Clinical tissue samples.** A total of 40 pairs of LSCC tissues and noncancerous tissues were frozen at the Institute of Otorhinolaryngology – Head and Neck Surgery (Hebei, China). Clinical features and pathological diagnosis were exported from corresponding data records. No chemotherapy and/or radiotherapy treatment was received before surgery. The Ethical Committee approved this research of the Second Hospital of Hebei Medical University (Hebei, China), and all patients provided informed consent.

**Cell culture and silencing of MMP1.** The LSCC cell lines (TU686, TU177, and AMC-HN-8), hypopharyngeal cell lines (FaDu), and 293T cell lines were purchased from Beijing Beina Chuanglian Institute of Biotechnology (Beijing, China). These cell lines were cultured as described in previous protocols [8]. Small interfering RNA (siRNA) was used to knock down MMP1 expression *in vitro*. The siRNA of MMP1 and negative control (si-NC) were obtained by RiboBio (Guangzhou, China). The sequences are listed in Table 2. The Lipofectamine 3000 (Invitrogen, USA) was adopted to accomplish the cell transfections following the manufacturer’s instructions. Cells transfected with siRNAs for 24 h proceeded to a quantitative real-time polymerase chain reaction (qRT-PCR) and cell function assays. The 18S was used as an internal control to determine the expression of interesting genes.

**Cell biological behavior assays.** Previous research described the experimental process for cell proliferative assay, cell migration and invasion, and the colony formation assays [8]. Here, we briefly introduce the steps of the wound-healing assay. When the cells in 6-well plates grew to 70–80% confluence, a 200  $\mu\text{l}$  pipette tip was used to scrape them. The scratched areas were photographed and measured at 0, 24, and 48 h.

**Statistical analysis.** Statistical data were conducted using R-4.0.2-win. Wilcoxon rank-sum test was used to analyze continuous skewed variables, while the parametric data were performed using t-tests. A value of  $p < 0.05$  was considered statistically significant. GraphPad Prism 8 was applied for bar and line graphs.

**Table 1.** Details of the GEO (Gene Expression Omnibus) LSCC data sets.

Reference	Tissue	GEO	Platform	Normal	Tumor	Biotype
Lian et al. (2013)	LSCC	GSE51985	GPL10558	3	10	mRNA
Wilson (2014)	LSCC	GSE59102	GPL6480	13	25	mRNA
Shen et al. (2014)	LSCC	GSE59652	GPL13825	7	6	mRNA
Feng et al. (2016)	LSCC	GSE84957	GPL17843	9	9	mRNA
Liu et al. (2020)	LSCC	GSE117005	GPL20115	5	4	mRNA
Nicolau et al. (2020)	LSCC	GSE143224	GPL5175	11	14	mRNA

## Results

**DEGs of MMPs in LSCC patients based on different databases.** A total of 19 differential expression genes of the MMP family were screened from the TCGA database, 18 genes were considerably highly expressed in LSCC tissues, while 1 gene was decreased in LSCC compared with adjacent non-tumor tissues (Figure 1A). Besides, eight differential genes were in the GEO database. Moreover, MMP12 and MMP1 were significantly upregulated in LSCC tissues compared to surrounding non-tumor tissues, with a fold change of 4.76 and 4.67, respectively (Figure 1B). From the Oncomine dataset, we analyzed the mRNA expression of MMPs between HNSC and adjacent normal tissues. MMP1, MMP2, MMP3, MMP7, MMP9, MMP10, MMP11, MMP12, MMP13, MMP14, and MMP16 were highly expressed in tissues compared with adjacent non-tumor tissues. MMP21, MMP23A, MMP24, MMP26, and MMP28 have no difference in HNSC and normal tissues. Among them, MMP1 had the largest number of datasets with statistically significant alterations in mRNA expression (Supplementary Figure S1). These results were almost consistent with those from TCGA. The combination of microarray results between four pairs of LSCC tissues and adjacent normal tissues [8] showed MMP1, MMP3, MMP8, MMP10, MMP11, MMP12, MMP13, MMP14, and MMP25 had both elevated expressions in LSCC tissues. MMP1 and MMP12 were the most significant upregulated genes, with the fold change of 6.47 and 9.22, respectively, in accordance with GEO datasets (Figure 1C). Further, seven common DEGs in the intersection of the four databases mentioned above were identified, including MMP1, MMP3, MMP9, MMP10, MMP11, MMP12, and MMP13 (Figure 1D). Together, these data suggest an important role of the seven common DEGs overexpression in LSCC patients.

**Table 2.** Primer sequences used in this study.

Gene	Sequence
MMP1	F:5'-TGTTCTGGGGTGTGGTGTCT-3' R:5'-CTCCGCTTTTCAACTTGCCTC-3'
MMP3	F:5'-GGTTCATGCTGGTGTCTCA-3' R:5'-AGGCAAGACAGCAAGGCATA-3'
MMP8	F:5'-AGACGCTTCCATTTCTGCTCT-3' R:5'-TTTCCAGGTAGTCCTGAACAGT-3'
MMP10	F:5'-TGCTTTGTCCTTCGATGCCA-3' R:5'-AAACGGTGTCCCTGCTGTTA-3'
PRSS23	F:5'-GCTCGGCGCGAACAG-3' R:5'-AATTGAGGGTAGACTGGGGC-3'
18S	F:5'-ATCCTCAGTGAGTTCTCCCG-3' R:5'-CTTTGCCATCACTGCCATTA-3'
si-MMP1-1	5'-GCTTGAAGCTGCTTACGAA-3'
si-MMP1-2	5'-GGACCATGCCATTGAGAAA-3'
si-MMP1-3	5'-GCACATGACTTCTCGAA-3'

**Construction of an MMPs-related prognostic model based on MMP1, MMP3, MMP8, and MMP10.** First, univariate Cox regression analysis was applied with the expression profiles of the 25 MMPs genes. Hazards ratio and p values of the 25 genes were displayed using heatmaps (Supplementary Figure S2). MMP1, MMP3, MMP8, and MMP10 were screened using  $p < 0.05$ . Then, multi-gene prediction models were constructed based on the univariate Cox outcomes to evaluate the involved effect of the screened genes on patient overall survival. In total, MMP1, MMP3, MMP8, and MMP10 were eventually selected to form the models, and risk scores of the patients were calculated using the formulas mentioned above in the final analysis (Figure 2A).

Additionally, we compared the AUC for each ROC curve to assess the MMPs-related prognostic model performance. The AUC of the MMP gene-based model (area under red = 0.641) was higher than the model based on TNM stages (area under blue curve = 0.536), the model based on grade (area under green = 0.523), and the model based on gender (area under yellow = 0.374) (Figure 2B). Moreover, patients were divided into the high- and low-risk groups by the median value, and the corresponding risk value was listed in ascending order (Figure 2C). Survival curves presented that patients with high-risk scores were significantly correlated with shorter survival rates (Figure 2D). Patient survival status demonstrated that patients with high-risk scores had greater mortality than patients with low-risk scores (Figure 2E). Then, a heatmap was drawn to display the gene expression profiles between high-risk and low-risk groups. And, increased expression of the four prognostic genes was observed as the risk value increased (Figure 2F).

**MMP1, MMP3, MMP8, and MMP10-prognostic model evaluated as independent prognostic factor and construction of nomogram model.** The risk score, age, gender, grade, and TNM stage were estimated by multivariate Cox regression analysis. Consequently, risk score ( $p = 0.015$ ) was significantly associated with prognosis, dominated independent prognostic factor for overall survival (Figure 2G). Moreover, the nomogram was constructed based on the risk score, stage, gender, age, and grade for predicting prognoses of the patients. The score of each factor and the total scores of all factors can be obtained from the nomogram. The total scores predicted the 1-, 3-, and 5-year survival rates (Figure 2H).

**Weighted gene co-expression network analysis of MMP1, MMP3, MMP8, and MMP10.** The co-expression network of MMP genes and the risk-score-based DEGs were obtained in modules were significant ( $p < 0.05$ ). Red nodes showed MMP family genes, and the genes belonging to the 4-gene prediction model were labeled with bigger red nodes. The blue nodes indicated the co-expressed genes. The four prognostic model genes (MMP1, MMP3, MMP8, and MMP10) were all present in the co-expression network. In addition, MMP1 was collectively correlated with MMP3 and PRSS23, MMP8 was collectively correlated with MMP9, and MMP10 was collectively correlated with FN1 and IL24

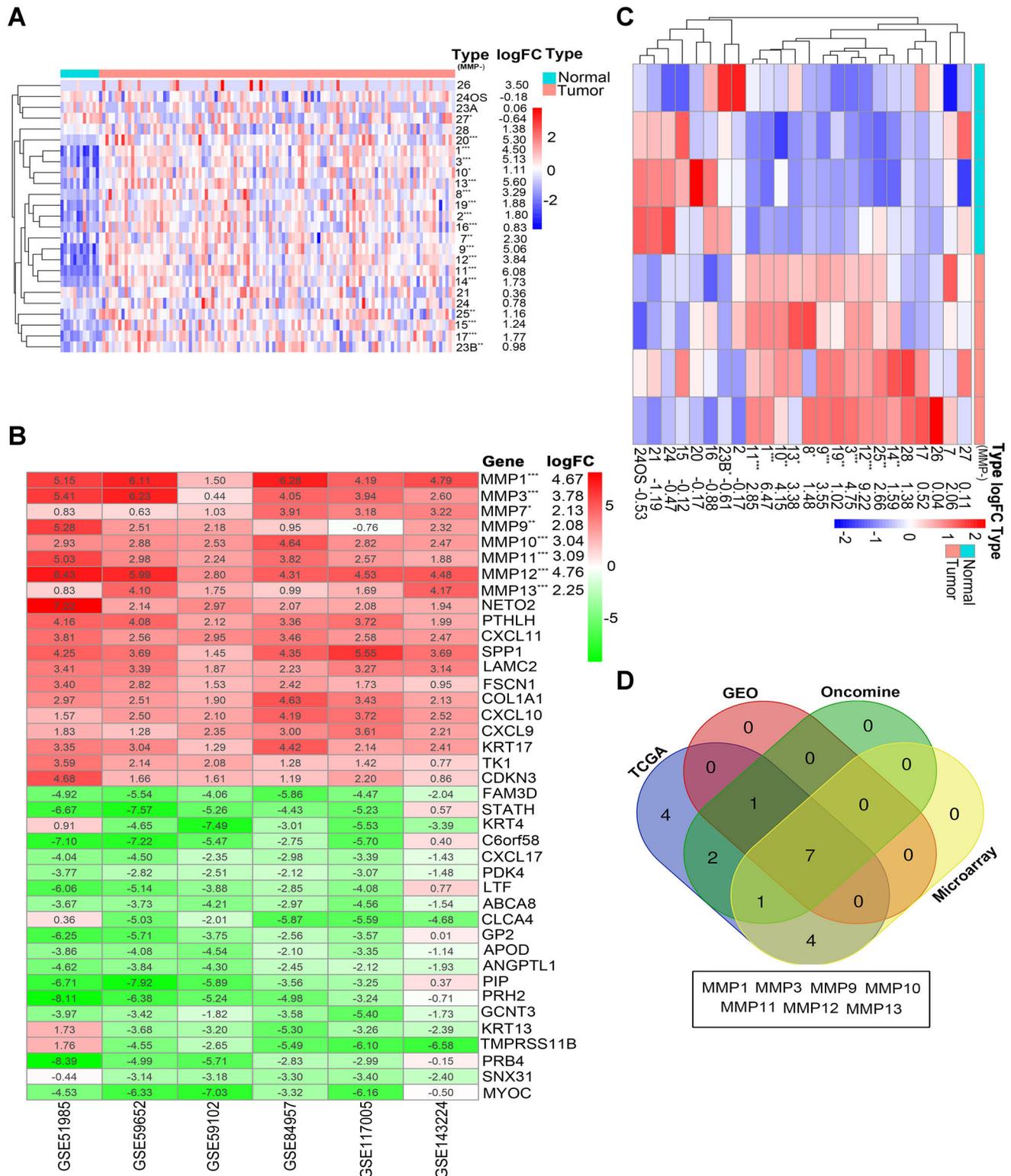


Figure 1. mRNA expression of MMPs in LSCC. The heatmap showing the fold changes and p-value of MMP family members compared LSCC patients with normal people from the TCGA (A), GEO (B), and microarray (C) data sets (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001); D) Venn-diagram of common DEGs in TCGA, GEO, Oncomine, and microarray data

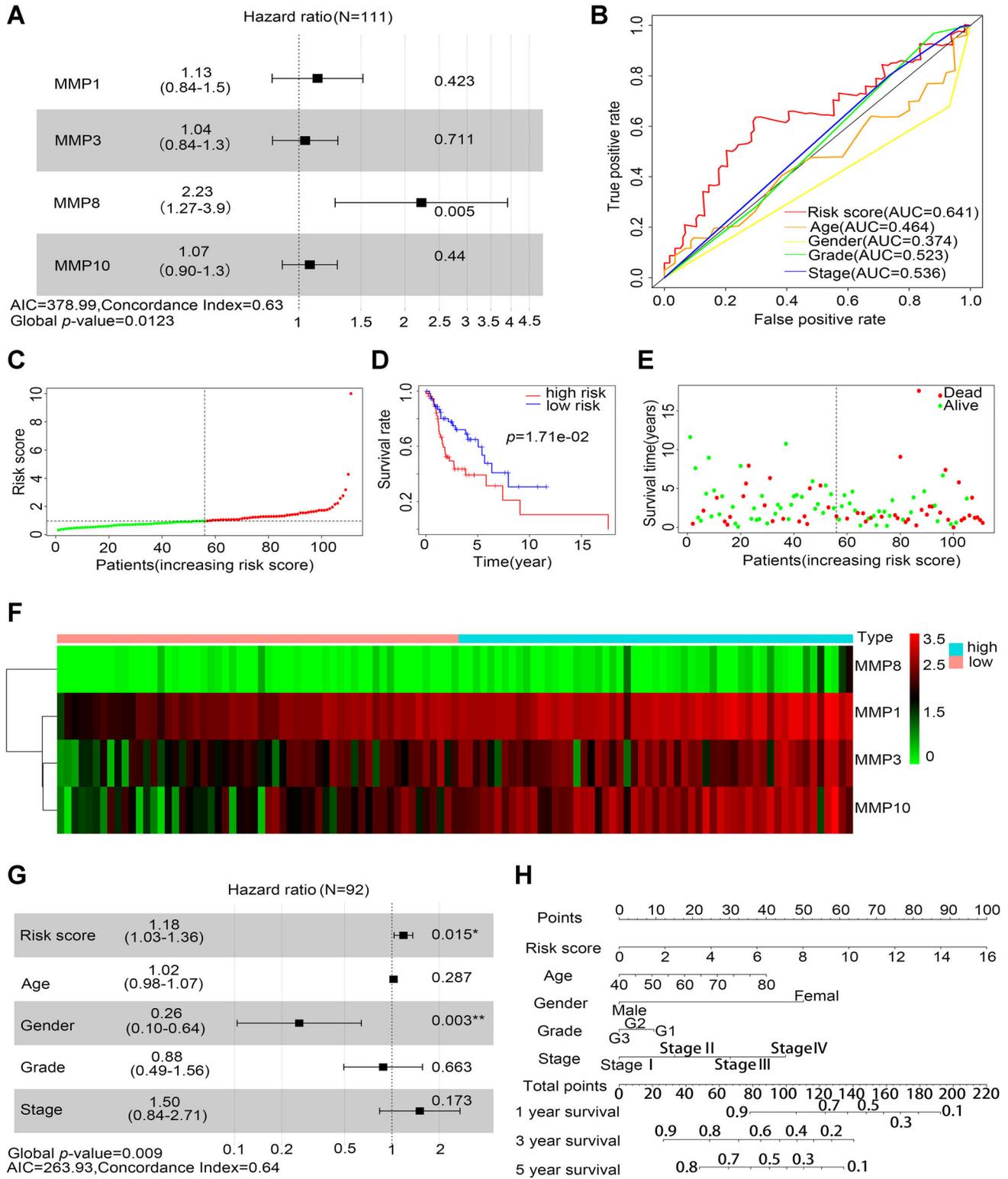


Figure 2. The multivariate Cox regression analysis. A) Forest plot of the multivariate Cox regression model showing four prognosis-related signature genes; B) The ROC curves and AUC of risk score (red curve), age (orange curve), gender (yellow curve), grade (green curve), and TNM stage (blue curve); the risk score distribution (C), Kaplan-Meier curve (D), patient survival status (E), and heatmap of the four MMP genes of the LSCC patients in the high- and the low-risk groups (F); G) Forest plot presenting the multivariate risk factors of LSCC patients; H) Nomogram for the 1, 3, 5-year OS prediction based on risk score, age, gender, grade, and TNM stage

(Figure 3A). Notably, the four prognostic model genes (MMP1, MMP3, MMP8, and MMP10) did not exist in isolation but instead, as a complex interconnected network, so MMP1 or MMP10 might fulfill pivotal roles in LSCC patients.

**Tumor microenvironment estimation of subtypes of LSCC based on MMP1, MMP3, MMP8, and MMP10-prognostic model.** The ESTIMATE algorithm was used to calculate the stromal score, immune score, and tumor purity of the high- and low-risk groups. The high-risk group showed a higher stromal component and lower tumor purity than those in the low-risk group. Our results show that the more advanced the risk, the higher the immune score will be, despite being no statistical significance (Figure 3B). The difference of immune cells infiltration was compared between the low- and high-risk groups. Five types of immune cells with differences in infiltration were detected between the two groups, such as plasma cells, CD8+ T cells, follicular helper T cells, resting NK cells, and M0 macrophages (Figure 3C). These findings suggested that tumor microenvironment, especially immune cells might be involved in MMPs-affected prognosis in LSCC.

**qRT-PCR validated results of selected MMPs.** The mRNA levels of MMP1, MMP3, MMP8, and MMP10 in different tumors were identified by the Oncomine database (Figure 4A). This analysis revealed that the MMP1, MMP3,

and MMP10 expression, but not MMP8 was the highest in HNSC than other cancers, and similar trends were observed in the GEPIA dataset (<http://gepia.cancer-pku.cn/>) (Supplementary Figure S3). Thus, MMP1, MMP3, and MMP10 were selected for the next qRT-PCR validation in 40 paired LSCC and adjacent nontumor tissues with consideration of tissue-specific expression pattern. Patient detailed information is shown in Table 3. The results demonstrated that MMP3 and MMP10 were significantly upregulated in 34/40, 31/40 paired LSCC tissues. MMP1 indicated significant upregulation (40/40) in paired LSCC tissues and had the highest basal expression of LSCC samples in TCGA (Figures 4B, 4C). Moreover, the TCGA database analysis indicated that MMP1 overexpression was correlated with lower progression-free survival and overall survival of patients with LSCC (Supplementary Figure S4). Therefore, MMP1 as a candidate of MMPs was selected for further functional validation.

**The expression of MMP1 and the relationship between MMP1 and clinical features in LSCC cohort of the Second Hospital of Hebei Medical University.** The immunohistochemistry results revealed that the expression levels of MMP1 were markedly upregulated in LSCC tissues compared with adjacent non-tumor tissues (Figure 4D). The relationship between the expression of MMP1 and clinical parameters was also investigated in 40 paired LSCC patients. Statistical results revealed that expression level of MMP1 was significantly associated with smoking ( $p < 0.05$ ), TNM stage ( $p < 0.001$ ), lymphatic metastasis ( $p < 0.001$ ), and pathological differentiation ( $p < 0.01$ ). Moreover, no relationship was observed between the age, alcohol use, and location of carcinoma (Figure 5A). Based on the qPCR ( $2^{-\Delta\Delta Ct}$ ) results and follow-up data, we carried out a survival analysis according to the median value (39.64). Patients with high MMP1 expression had poorer overall survival than patients with low MMP1 expression (Figure 5B). In order to verify the reliability of the model predictions, we used our independent datasets to verify the model's prediction capabilities. The validation set prediction correlation coefficient reached 0.847, and survival time was significantly different between groups ( $p = 0.0042$ ), indicating that the model had a good predictive ability (Figures 5C–5E).

**Downregulation of MMP1 inhibits cell proliferation, migration, and invasion *in vitro*.** Considering that the Oncomine and GEPIA databases presented only the expression profiles of HNSC, and hypopharyngeal squamous cell carcinoma (HSCC) was a head and neck malignant tumor with one of the worst prognoses. In addition to LSCC, we also evaluated HSCC cell lines. The mRNA expression levels of MMP1 in various laryngeal and hypopharyngeal squamous cell carcinoma cell lines were detected by qRT-PCR (Figure 6A). Consequently, the TU686 and FaDu cell lines were selected for the next functional experiments. The TU686 and FaDu cell lines were transfected with four well-designed small interfering RNAs (si-MMP1-1/si-MMP1-2/si-MMP1-3/si-NC). Besides, qRT-PCR confirmed the knock-

**Table 3. Clinical information in 40 cases of LSCC.**

Characteristics	No. cases (100%)
Gender	
male	40 (100.00)
female	0 (0.00)
Age	
<63	20 (50.00)
≥63	20 (50.00)
Smoking	
no	4 (10.00)
yes	36 (90.00)
Alcohol	
no	14 (35.00)
yes	26 (65.00)
Primary site	
supraglottic	17 (42.50)
glottic	12 (30.00)
infraglottic	7 (17.50)
transglottic	4 (10.00)
TNM stage	
I	9 (22.50)
II+III+IV	31 (77.50)
LN metastasis	
N0	28 (70.00)
N1+N2+N3	12 (30.00)
Pathological differentiation	
well	16 (40.00)
moderate + poor	24 (60.00)

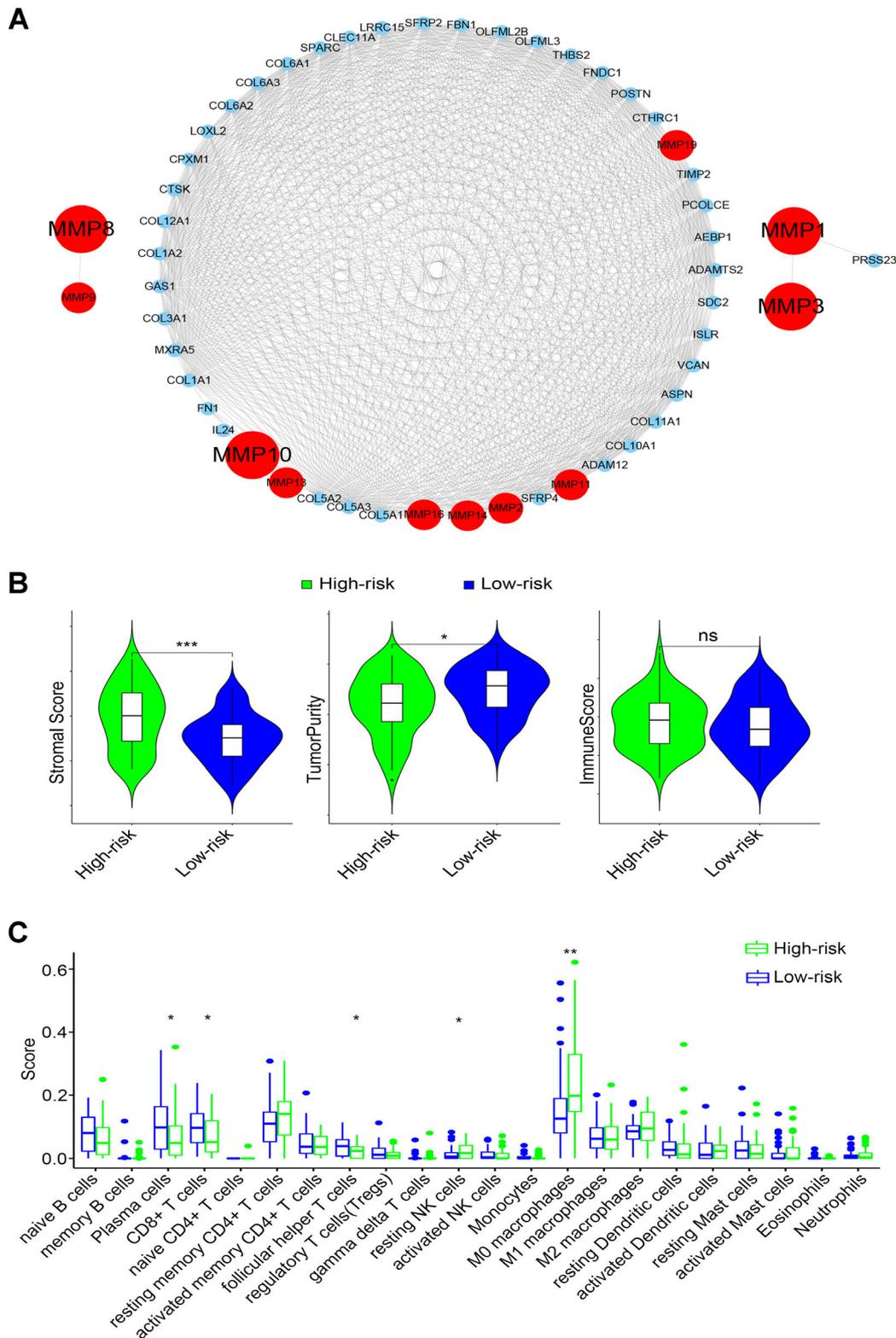


Figure 3. Co-expression network and tumor microenvironment analysis of MMPs family genes. A) Visualization of the co-expression between MMPs family genes and the risk-scores-based DEGs. The red nodes are MMPs family genes, and the bigger ones are the genes included in the 4-gene prediction model. The blue nodes are the co-expressed genes; B) Violin plot representing the differences in the stromal score, tumor purity, and immune score between high- and low-risk groups; C) The difference of immune cells infiltration between the low- and high-risk groups (\*p<0.05, \*\*p<0.01)

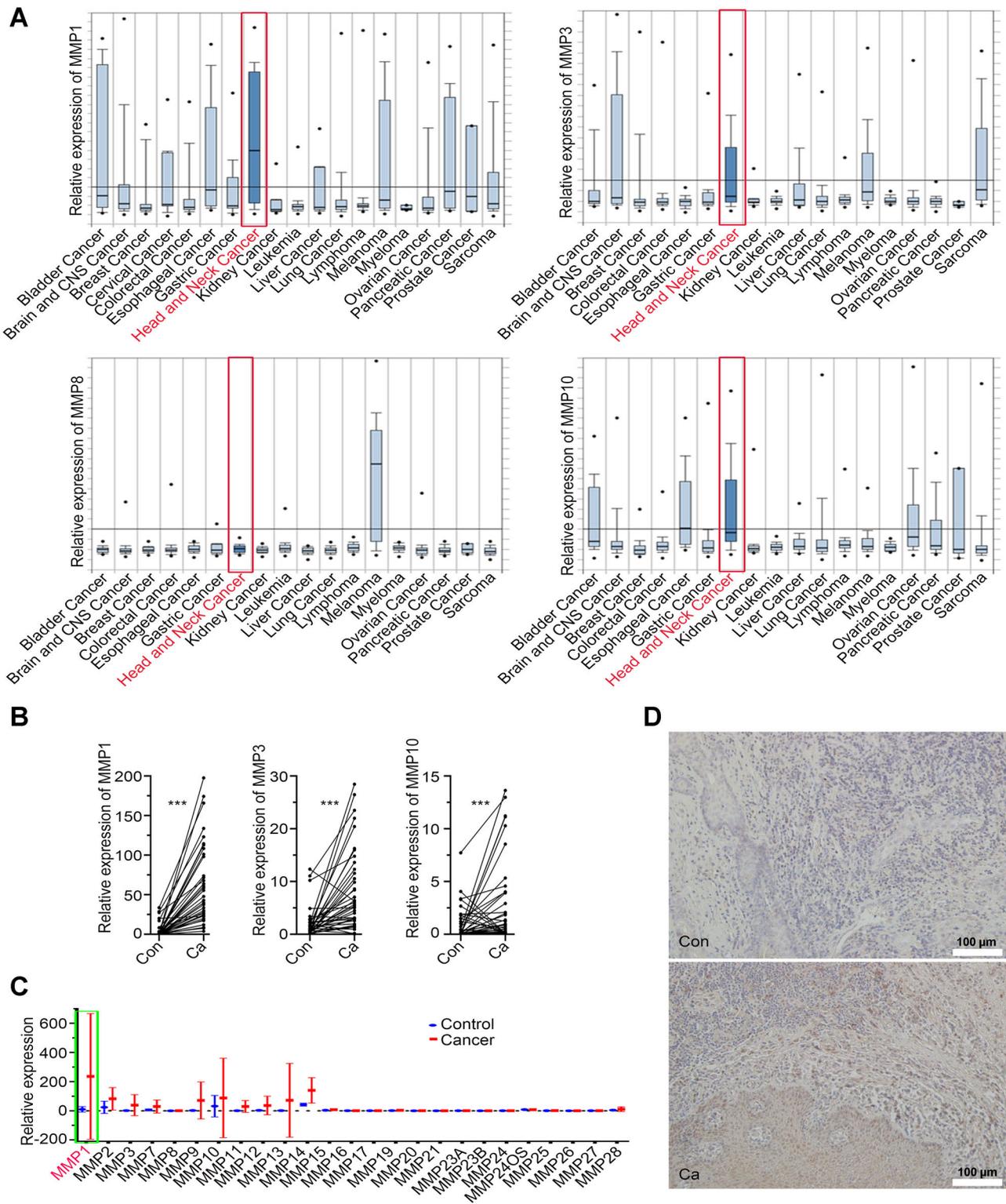


Figure 4. mRNA expression of MMP1, MMP3, MMP8, and MMP10 in various cancers and the expression verification result of individual genes. A) The expression of MMP1, MMP3, MMP8, and MMP10 in different types of cancers in the Oncomine database, and the expression data of HNSC was shown in the red box; B) qRT-PCR analysis of the expression of MMP1, MMP3, and MMP10 in 40 paired LSCC tissues; C) mRNA expression levels of MMP family members compared LSCC patients with normal people from the TCGA by boxplot; D) Expression of MMP1 in LSCC paired tissues by IHC

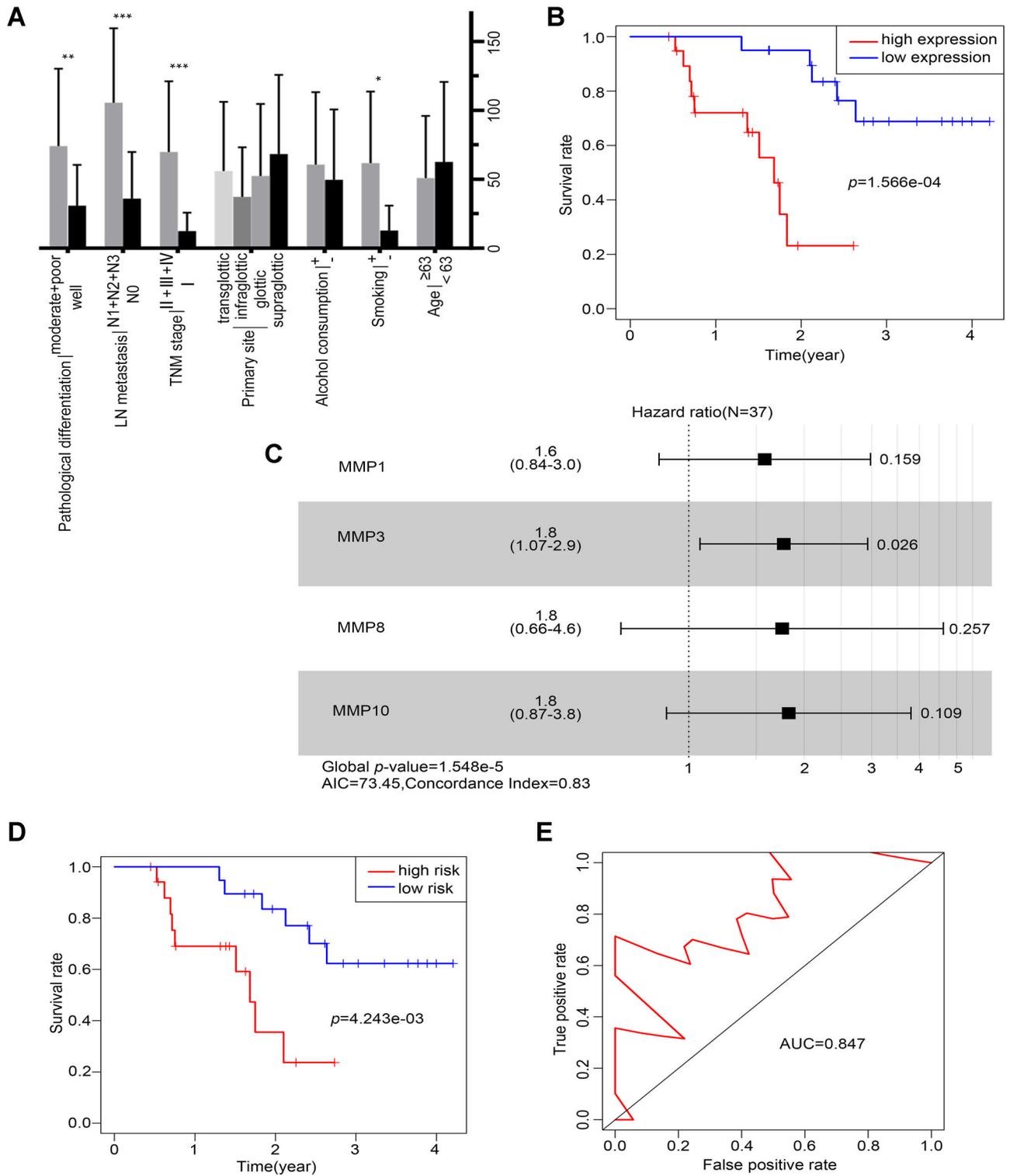


Figure 5. Model validation using independent datasets. A) The relation between the expression of MMP1 and clinical parameters based on qPCR results; B) Kaplan-Meier estimated overall survival in patients with high or low MMP1 expression, higher MMP1 expression with poorer overall survival; C) Validation set multivariate Cox regression analysis forest map; D) Survival curve of LSCC patients based on risk score model; E) ROC curve for the validation set

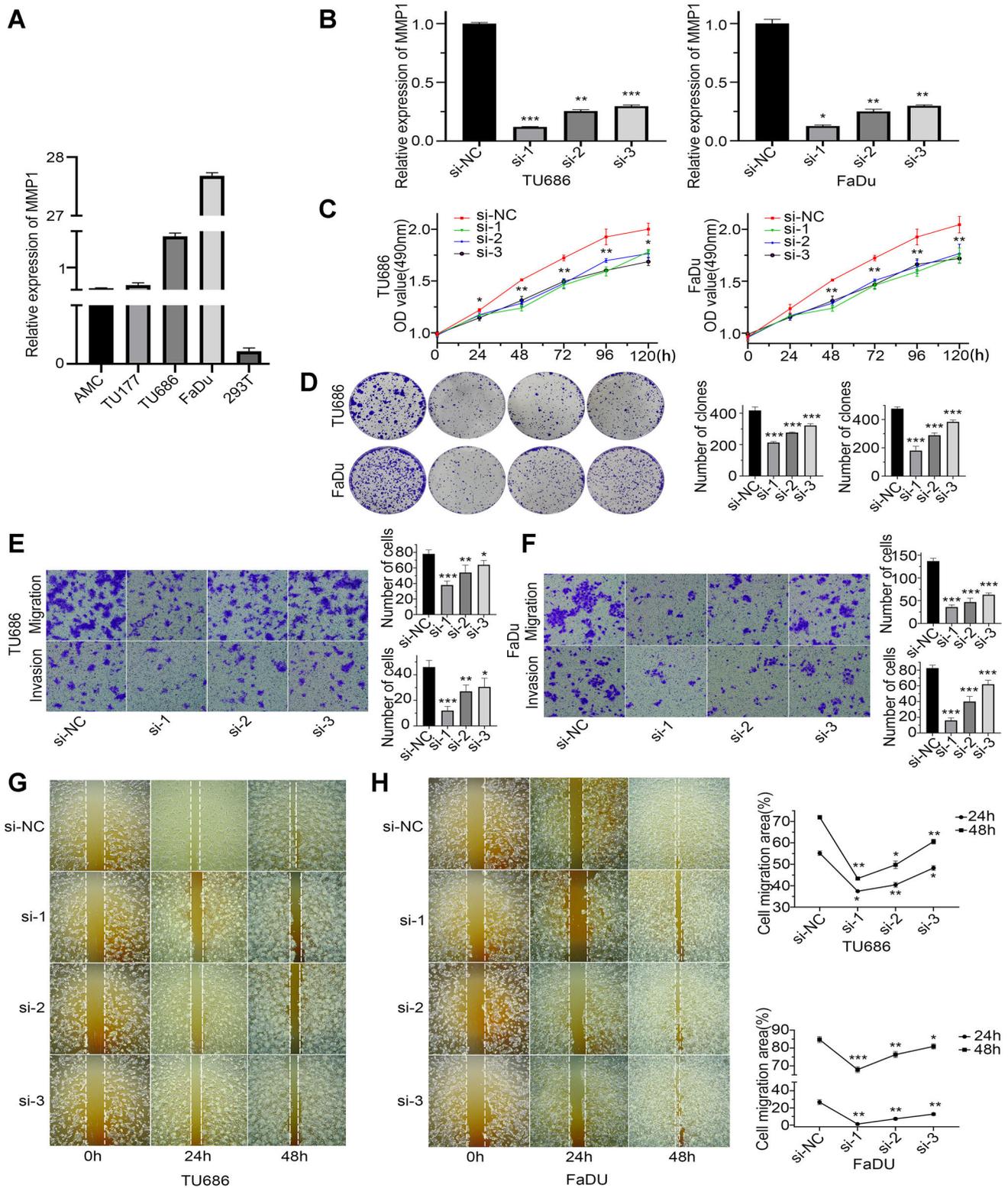


Figure 6. Downregulation of MMP1 inhibits cell proliferation, migration, and invasion *in vitro*. A) Relative expression of MMP1 in five cell lines; B) The expression of MMP1 significantly reduced after the cells were transfected with four small interfering RNAs; C) The cell viability was assessed by MTS assay after transfected with MMP1 siRNA; D) The number of colonies was remarkably decreased after MMP1 was knocked down; E, F) Decreased expression of MMP1 attenuates the invasion and migration capability *in vitro*; G, H) The suppressed migration capability in cells treated with si-MMP1 was demonstrated by wound-healing assays. Data were statistically analyzed presented by bar graphs (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ )

down efficiency, where the expression of MMP1 reduced after the cells were transfected with si-MMP1 (Figure 6B). MTS assays showed that reduction of MMP1 expression impaired the proliferative capability of the two cell lines (Figure 6C). Meanwhile, the number of colonies formed by TU686 and FaDu cell lines was remarkably decreased after the expression of MMP1 was knocked down (Figure 6D). Then, Transwell and wound healing assays were subsequently performed to evaluate the effect of MMP1 on the migration and invasive ability of squamous cells carcinoma. Knockdown of MMP1 significantly decreased the migratory and invasive capabilities of the two cell lines (Figures 6E, 6F). Moreover, in the scratch wound healing assay at 24 h and 48 h, silencing of MMP1 attenuated the migration potential in the two cell lines (Figures 6G, 6H).

## Discussion

MMPs are produced by tumor cells themselves or by surrounding stromal cells including fibroblasts, stimulated by the nearby tumor with a significant role in cell proliferation and migration [4]. Herein, we synthetically analyzed the MMPs in LSCC implicated in expression, prognostic value, co-expression, immune cell infiltration, and cellular functional assay.

MMP genes were highly expressed in LSCC tissues when compared to surrounding non-tumor tissues. Among the MMPs, seven common DEGs were discovered by integrating our microarray data and public databases. Given that the predictive ability of single or limited genes was barely satisfactory, we identified four prognosis-related signature genes (MMP1, MMP3, MMP8, and MMP10) to develop the prognostic model. And these four genes that affect independently tumor proliferation or prognosis were respectively described in the literature [9–12]. In order to investigate the role of the MMP-related model in the prognosis of LSCC, overall survival analysis and ROC curves were performed. Kaplan-Meier survival curves identified the excellent efficiency of our MMP-related model in discriminating patients of different risk groups with different risks of death. ROC curves confirmed satisfactory sensitivity and specificity of the MMP-based prognostic model. Moreover, this prognostic model was then independently validated with our clinical datasets. This suggests that the prediction model had good predictive ability. Interestingly, the predictive accuracy of the MMP gene-based model was significantly high compared to the TNM stages. The TNM staging system remains the standard, while other methods are recognized auxiliary methods classifying the actual infiltration range of tumors and widely used to predict cancer prognosis clinically. Therefore, the MMP-related model might be useful as a supplement to the tumor stage for better stratifying patients to more individualized treatments.

We then established that MMP1, MMP3, and MMP10 expression was highest in HNSC than in other cancers by

OncoPrint, and were tissue-specific. Moreover, MMP1 was expressed at the most significant higher levels in LSCC than in tumor-adjacent tissues in the TCGA database. Despite MMP1, MMP3, or MMP10 high expression in LSCC had been reported in several studies, most of the studies were based on individual genes [13–15]. Systematically analyzing the MMP gene family in LSCC was performed for the first time in this study.

To further elucidate the underlying mechanisms of the MMP-regulated downstream signals in LSCC, co-expression network and tumor immune cell infiltration were performed. As shown by the co-expression network, the four prognostic model genes (MMP1, MMP3, MMP8, and MMP10) were co-expressed with genes implicated in oncogenic function, i.e., PRSS23, FN1, and IL24. Moreover, MMP1 in the co-expression network tends to be highly interconnected with MMP3 and PRSS23, and these findings were verified by qRT-PCR to some degree (Supplementary Figure 5). Meantime, Han et al. suggested that PRSS23 downregulation impairs gastric tumorigenesis via the EIF2 pathway, which displayed a better prognosis [16]. An analysis of integrated three microarray datasets revealed that PRSS23 is a potentially critical factor in the lapatinib resistance [17]. Therefore, knockdown of MMP1 in TU686 and FaDu cells, diminished cell proliferation, migration, and invasion ability are potentially related to the down-expression of PRSS23.

Tumor microenvironment (TME) is an increasingly popular topic and potentially related to tumor progression and therapeutic outcomes [18]. The tumor TME is comprised mainly of immune cells, extracellular components, and cancer-associated fibroblasts (CAFs) [19]. In this study, MMP-dominated high-risk group showed a higher stromal score than those in the low-risk group, while the immune score did not differ significantly. This illustrates that CAFs showed a higher presence in the high-risk group than the low-risk group. Meanwhile, because the production of MMPs was CAFs dependent and the presentation of CAFs was diverse [19]. Therefore, we speculated that MMP1, MMP3, MMP8, and MMP10 based prognostic model may indirectly reflect the invasion and metastasis function of CAFs in LSCC. Moreover, CD8+ T cells can fight tumors by producing cytokines and killing effects [18]. The presence of more activated CD8+ T cells and fewer M0 macrophage cells within the tumor of the low-risk patients exhibited a favorable prognosis in this study. These findings corroborate with previous findings on cervical squamous cell carcinoma and breast cancer [20, 21]. In addition, this phenomenon was consistent with the reported function of CAFs in inhibiting T cell proliferation and gathering macrophages [19]. Altogether, the expression of MMP1, MMP3, MMP8, and MMP10 significantly correlated with tumor-infiltrating immune cells, and this will provide insight into rational drug design or antitumor immunotherapy.

To validate the accuracy of bioinformatics analysis based on the sequencing data, we performed qRT-PCR experi-

ments and a cellular functional assay. In our series of 40 patients with laryngeal cancer, we found that the MMP1 mRNA expression in LSCC tissues was considerably higher than that in the normal tissues. We also observed that MMP1 was not randomly expressed but significantly correlated with stage, lymphatic metastasis, and pathological differentiation. This result corresponded with a study by Pietruszewska et al., which demonstrated that a high expression of MMP1 was related to low histological differentiation in head and neck squamous cell cancer [22]. In line with this, MMP1 has been reported to relate to lymphatic metastasis in oropharyngeal cancer [9]. *In vitro* data demonstrated that downregulation of MMP1 in LSCC cell lines inhibited cell proliferation, migration, and invasion, indicating that MMP1 activation exerts oncogenic activities towards laryngeal. Likewise, similar results were confirmed in HSCC cells. This suggests that MMP1 has essential roles in the progression of laryngeal cancer. Meanwhile, Liu et al. reported that MMP1 promoted esophageal squamous cell carcinoma proliferation, migration, and invasion [23]. Saito et al. reported that MMP1 correlated with high abilities of migration and invasion in lung adenocarcinoma [24]. Furthermore, Wang et al. revealed that MMP1 participated in the invasion of colorectal cancer via the PI3K/Akt/c-myc signaling pathway [25]. Moreover, MMP1 was shown to be implicated in breast cancer therapy resistance as well as poor prognosis [26].

However, bioinformatic analyses in this study were based predominantly on LSCC data, we also demonstrated the oncosuppressive function of MMP1 in FaDu cells. The followings are the reasons: 1. HSCC occurs at an anatomic location adjacent to LSCC, they are similar in histomorphology. 2. MMP1 is more highly upregulated in FaDu than TU686 cells, and HSCC has a worse prognosis compared to LSCC. These results suggest that MMP1 might be used as a biomarker of prognosis and therapeutic response in LSCC and even in HSCC.

Collectively, we discovered a signature of four genes (MMP1, MMP3, MMP8, and MMP10) from 25 MMP genes family using the TCGA expression profiles and prognosis information. These signature genes were predictive of clinical outcomes of LSCC and significantly associated with tumor immune response. Furthermore, the most significant gene, MMP1, was selected for subsequent wet-lab validation of analysis. Knockdown MMP1 significantly impaired the proliferation, migration, and invasion capacities of TU686 cells, thereby improving the survival of LSCC patients.

**Supplementary information** is available in the online version of the paper.

**Acknowledgments:** We acknowledge the financial support by the National Natural Science Foundation of China (81972553) and the Project of Clinical Medical Talent Training and Basic Project Research Funded by the Government [Hebei finance society (2019) No.139].

## References

- [1] GALE N, POLJAK M, ZIDAR N. Update from the 4th Edition of the World Health Organization Classification of Head and Neck Tumours: What is New in the 2017 WHO Blue Book for Tumours of the Hypopharynx, Larynx, Trachea and Parapharyngeal Space. *Head Neck Pathol* 2017; 11: 23–32. <https://doi.org/10.1007/s12105-017-0788-z>
- [2] STEUER CE, EL-DEIRY M, PARKS JR, HIGGINS KA, SABA NF. An update on larynx cancer. *CA Cancer J Clin* 2017; 67: 31–50. <https://doi.org/10.3322/caac.21386>
- [3] SUNG H, FERLAY J, SIEGEL RL, LAVERSANNE M, SOERJOMATARAM I et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71: 209–249. <https://doi.org/10.3322/caac.21660>
- [4] VARGOVA V, PYTLIAK M, MECHIROVA V. Matrix metalloproteinases. *Exp Suppl* 2012; 103: 1–33. [https://doi.org/10.1007/978-3-0348-0364-9\\_1](https://doi.org/10.1007/978-3-0348-0364-9_1)
- [5] ROY R, MORAD G, JEDINAK A, MOSES MA. Metalloproteinases and their roles in human cancer. *Anat Rec (Hoboken)* 2020; 303: 1557–1572. <https://doi.org/10.1002/ar.24188>
- [6] LI J, WANG X, ZHENG K, LIU Y, LI J et al. The clinical significance of collagen family gene expression in esophageal squamous cell carcinoma. *PeerJ* 2019; 7: e7705. <https://doi.org/10.7717/peerj.7705>
- [7] LANGFELDER P, HORVATH S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; 9: 559. <https://doi.org/10.1186/1471-2105-9-559>
- [8] MENG W, CUI W, ZHAO L, CHI W, CAO H et al. Aberrant methylation and downregulation of ZNF667-AS1 and ZNF667 promote the malignant progression of laryngeal squamous cell carcinoma. *J Biomed Sci* 2019; 26: 13. <https://doi.org/10.1186/s12929-019-0506-0>
- [9] LASSIG AAD, JOSEPH AM, LINDGREN BR, YUEH B. Association of Oral Cavity and Oropharyngeal Cancer Biomarkers in Surgical Drain Fluid With Patient Outcomes. *JAMA Otolaryngol Head Neck Surg* 2017; 143: 670–678. <https://doi.org/10.1001/jamaoto.2016.3595>
- [10] TAHA EA, SOGAWA C, OKUSHA Y, KAWAI H, OO MW et al. Knockout of MMP3 Weakens Solid Tumor Organoids and Cancer Extracellular Vesicles. *Cancers (Basel)* 2020; 12: 1260. <https://doi.org/10.3390/cancers12051260>
- [11] BOCKELMAN C, BEILMANN-LEHTONEN I, KAPRIO T, KOSKENSALO S, TERVAHARTIALA T et al. Serum MMP-8 and TIMP-1 predict prognosis in colorectal cancer. *BMC Cancer* 2018; 18: 679. <https://doi.org/10.1186/s12885-018-4589-x>
- [12] PISKOR BM, PRZYLIPIAK A, DABROWSKA E, SIDORKIEWICZ I, NICZYPORUK M et al. Plasma Level of MMP-10 May Be a Prognostic Marker in Early Stages of Breast Cancer. *J Clin Med* 2020; 9: 4122. <https://doi.org/10.3390/jcm9124122>
- [13] FANG W, SHEN J. Identification of MMP1 and MMP2 by RNA-seq analysis in laryngeal squamous cell carcinoma. *Am J Otolaryngol* 2020; 41: 102391. <https://doi.org/10.1016/j.amjoto.2020.102391>

- [14] ZHOU X, QI Y. PLGF inhibition impairs metastasis of larynx carcinoma through MMP3 downregulation. *Tumour Biol* 2014; 35: 9381–9386. <http://dx.doi.org/10.1007/s13277-014-2232-2>
- [15] LAPA RML, BARROS-FILHO MC, MARCHI FA, DOMINGUES MAC, DE CARVALHO GB et al. Integrated miRNA and mRNA expression analysis uncovers drug targets in laryngeal squamous cell carcinoma patients. *Oral Oncol* 2019; 93: 76–84. <https://doi.org/10.1016/j.oraloncology.2019.04.018>
- [16] HAN B, YANG Y, CHEN J, HE X, LV N et al. PRSS23 knock-down inhibits gastric tumorigenesis through EIF2 signaling. *Pharmacol Res* 2019; 142: 50–57. <https://doi.org/10.1016/j.phrs.2019.02.008>
- [17] LEE YS, HWANG SG, KIM JK, PARK TH, KIM YR et al. Identification of novel therapeutic target genes in acquired lapatinib-resistant breast cancer by integrative meta-analysis. *Tumour Biol* 2016; 37: 2285–2297. <https://doi.org/10.1007/s13277-015-4033-7>
- [18] PITT JM, MARABELLE A, EGGERMONT A, SORIA JC, KROEMER G et al. Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. *Ann Oncol* 2016; 27: 1482–1492. <https://doi.org/10.1093/annonc/mdw168>
- [19] QIN Y, ZHENG X, GAO W, WANG B, WU Y. Tumor microenvironment and immune-related therapies of head and neck squamous cell carcinoma. *Mol Ther Oncolytics* 2021; 20: 342–351. <https://doi.org/10.1016/j.omto.2021.01.011>
- [20] ZHAO S, YU M. Identification of MMP1 as a Potential Prognostic Biomarker and Correlating with Immune Infiltrates in Cervical Squamous Cell Carcinoma. *DNA Cell Biol* 2020; 39: 255–272. <https://doi.org/10.1089/dna.2019.5129>
- [21] EIRO N, CID S, AGUADO N, FRAILE M, DE PABLO N et al. MMP1 and MMP11 Expression in Peripheral Blood Mononuclear Cells upon Their Interaction with Breast Cancer Cells and Fibroblasts. *Int J Mol Sci* 2020; 22: 371. <https://doi.org/10.3390/ijms22010371>
- [22] PIETRUSZEWSKA W, BOJANOWSKA-POZNIAK K, KOBOS J. Matrix metalloproteinases MMP1, MMP2, MMP9 and their tissue inhibitors TIMP1, TIMP2, TIMP3 in head and neck cancer: an immunohistochemical study. *Otolaryngol Pol* 2016; 70: 32–43. <https://doi.org/10.5604/00306657.1202546>
- [23] LIU M, HU Y, ZHANG MF, LUO KJ, XIE XY et al. MMP1 promotes tumor growth and metastasis in esophageal squamous cell carcinoma. *Cancer Lett* 2016; 377: 97–104. <https://doi.org/10.1016/j.canlet.2016.04.034>
- [24] SAITO R, MIKI Y, ISHIDA N, INOUE C, KOBAYASHI M et al. The Significance of MMP-1 in EGFR-TKI-Resistant Lung Adenocarcinoma: Potential for Therapeutic Targeting. *Int J Mol Sci* 2018; 19: 609. <https://doi.org/10.3390/ijms19020609>
- [25] WANG K, ZHENG J, YU J, WU Y, GUO J et al. Knockdown of MMP1 inhibits the progression of colorectal cancer by suppressing the PI3K/Akt/cmyc signaling pathway and EMT. *Oncol Rep* 2020; 43: 1103–1112. <https://doi.org/10.3892/or.2020.7490>
- [26] SHEN CJ, KUO YL, CHEN CC, CHEN MJ, CHENG, YM. MMP1 expression is activated by Slug and enhances multidrug resistance (MDR) in breast cancer. *PLoS One* 2017; 12: e0174487. <https://doi.org/10.1371/journal.pone.0174487>

[https://doi.org/10.4149/neo\\_2021\\_210511N643](https://doi.org/10.4149/neo_2021_210511N643)

# The matrix metalloproteinase gene family: a significant prognostic gene lineage correlated with immune infiltrates in laryngeal squamous cell carcinoma

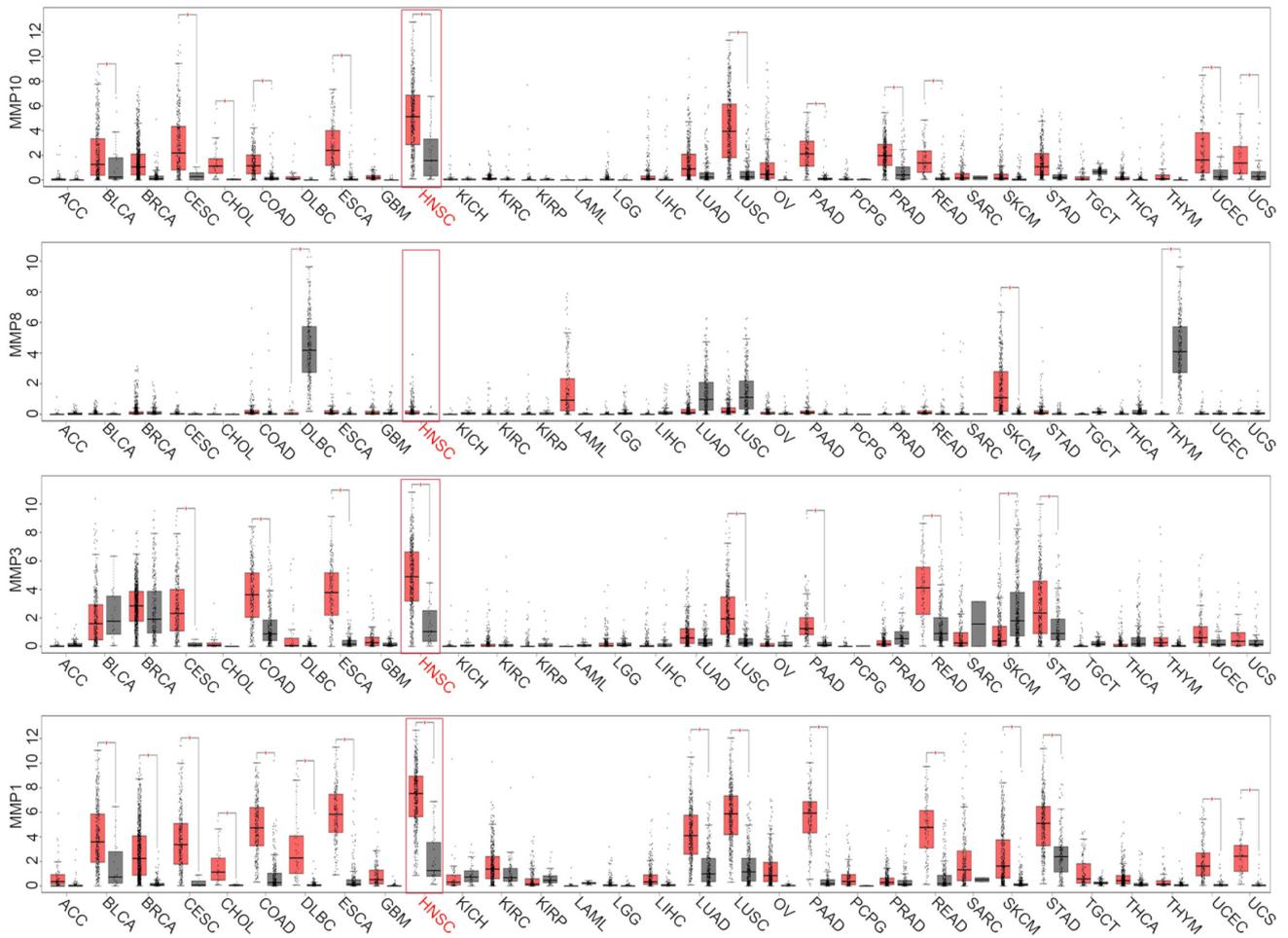
Chuan YANG, Huan CAO, Jian-Wang YANG, Tao LIU, Jian-Tao WANG, Bao-Shan WANG\*

Analysis Type by Cancer	Cancer vs. Normal																																					
	1	2	3	7	8	9	10	11	12	13	14	15	16	17	19	20	21	23A	23B	24	25	26	27	28														
Bladder Cancer	5	2	3	1	3	2	2	3	3	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
Brain and CNS Cancer	2	1	1	4	1	4	2	10	4	1	2	2	1	4	1	1	1	9	2	3	8	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Breast Cancer	13	3	5	10	2	6	1	25	8	3	1	10	15	1	10	7	1	1	3	2	1	4	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Cervical Cancer	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Colorectal Cancer	21	3	1	24	24	5	2	8	15	19	22	6	3	2	1	2	7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Esophageal Cancer	7	5	5	6	6	5	2	6	5	5	4	3	2	1	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gastric Cancer	5	2	8	10	2	5	2	5	2	8	4	1	7	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Head and Neck Cancer	14	2	11	12	13	11	11	5	11	13	1	12	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Kidney Cancer	5	1	5	4	2	5	4	1	4	3	1	2	5	5	3	1	4	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Leukemia	1	3	2	1	1	3	2	13	2	17	1	4	1	3	2	1	5	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Liver Cancer	1	4	3	2	3	2	4	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Lung Cancer	12	3	3	8	1	17	1	9	15	14	6	3	1	3	2	2	1	3	2	4	8	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Lymphoma	14	18	1	3	5	24	2	1	15	14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Melanoma	2	2	1	1	3	1	2	2	2	2	1	2	2	1	1	3	2	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Myeloma	7	5	2	4	1	2	3	14	5	7	2	11	4	2	4	1	2	4	1	2	5	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Other Cancer	3	1	2	7	1	3	1	2	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ovarian Cancer	6	5	1	6	1	5	1	6	1	6	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Pancreatic Cancer	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Prostate Cancer	4	9	2	4	2	1	4	1	1	3	3	1	2	11	2	3	6	6	1	1	3	7	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Sarcoma	4	9	2	4	2	1	4	1	1	3	3	1	2	11	2	3	6	6	1	1	3	7	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Significant Unique Analyses	126	6	74	28	80	5	92	35	15	22	152	24	75	116	11	110	5	53	9	53	18	17	20	39	20	15	19	33	27	20	1	6	3	3	2	11	8	56
Total Unique Analyses	338	356	304	359	306	363	332	345	308	348	355	361	340	349	347	277	145	35	116	319	78	240	280	256														

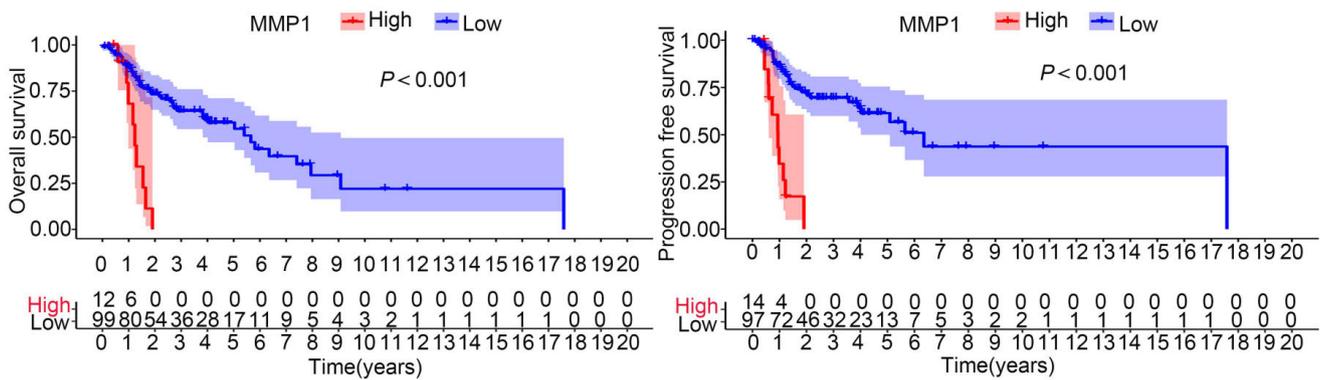
Supplementary Figure S1. The expression heatmap of MMP family genes in Oncomine database

HR	Pvalue	MMP Gene
0.05	0.28	MMP23B
0.59	0.10	MMP25
0.72	0.20	MMP17
0.88	0.86	MMP23A
0.89	0.55	MMP15
0.98	0.89	MMP19
1.00	1.00	MMP27
1.02	0.84	MMP7
1.03	0.70	MMP11
1.03	0.70	MMP28
1.06	0.29	MMP13
1.11	0.24	MMP9
1.11	0.08	MMP21
1.14	0.74	MMP24
1.15	0.03	MMP10
1.16	0.12	MMP12
1.16	0.20	MMP2
1.17	0.04	MMP3
1.20	0.40	MMP24OS
1.23	0.01	MMP1
1.23	0.22	MMP14
1.35	0.46	MMP16
1.72	0.16	MMP20
1.79	0.89	MMP26
3.07	0.01	MMP8

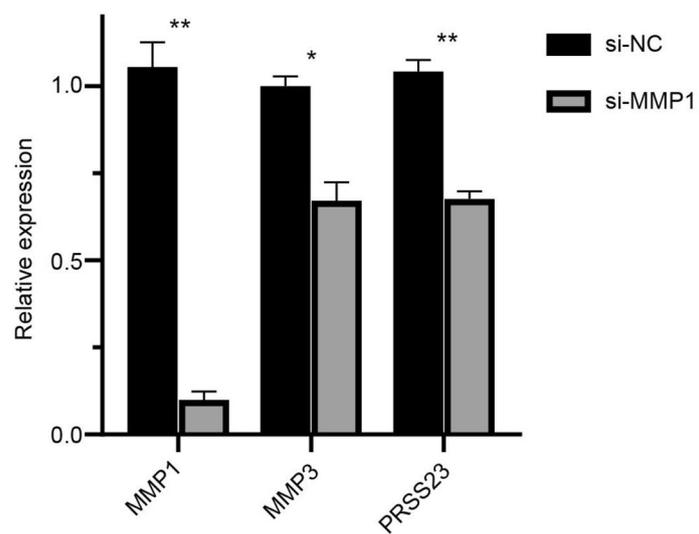
Supplementary Figure S2. The univariate Cox regression of MMP family genes in LSCC patients



Supplementary Figure S3. The expression of MMP1, MMP3, MMP8, and MMP10 in 31 various cancers in GEPIA



Supplementary Figure S4. Kaplan-Meier survival curves for the OS and PFS of MMP1 in LSCC



Supplementary Figure S5. The mRNA expression level of MMP1, MMP3, and PRSS23 in the TU686 cell line transfected with si-MMP1