

## The ETS1-LINC00278 negative feedback loop plays a role in COL4A1/COL4A2 regulation in laryngeal squamous cell carcinoma

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

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

\*Correspondence: hebwangbs@163.com

Received March 10, 2022 / Accepted April 19, 2022

The present study aimed to investigate LINC00278 expression in laryngeal squamous cell carcinoma (LSCC) and its involvement in the process of proliferation, migration, and invasion, providing a rationale for mining potential diagnostic and therapeutic targets of LSCC. Univariate and multivariate Cox regression analyses were performed to identify optimal prognostic lncRNAs. MTS, colony formation, wound healing, and Transwell invasion assays were used to determine the effects of LINC00278 overexpression on the proliferation, migration, and invasion of cancer cells. The expressions of signaling pathway-related proteins and epithelial-mesenchymal transition (EMT) marker proteins were detected using western blot. The chromatin immunoprecipitation (ChIP) and dual-luciferase reporter assays were performed to demonstrate the binding of ETS proto-oncogene 1, transcription factor (ETS1), and LINC00278 promoter region. The molecular targets of LINC00278 were identified by RNA sequencing analysis and co-expression analysis. Kaplan-Meier analysis and CIBERSORT algorithm were used to analyze survival and immune cell infiltration based on LINC00278, COL4A1, and COL4A2. Multivariate Cox regression was used to establish a six-gene prognostic model. LINC00278 expression was low in LSCC tissues, and it was significantly associated with the TNM (tumors/nodes/metastases) stage ( $p < 0.001$ ), lymphatic metastasis ( $p < 0.01$ ), and pathological differentiation ( $p < 0.01$ ). LINC00278 overexpression significantly reduced LSCC cell proliferation, migration, and invasion in TU686, TU177, and AMC-HN-8 cell lines. E-cadherin protein expression was increased, while N-cadherin, Vimentin, Zeb1, and Snail protein expression was decreased in the LINC00278 group, compared to the pcDNA3.1 group. Additionally, in AMC-HN-8 and FaDu cell lines, the LINC00278-treated group had significantly lower p-AKT and p-mTOR protein levels than the control group. ETS1 is a direct transcriptional regulator of the LINC00278 gene based on luciferase reporter assays and ChIP experiments. Western blot analysis demonstrated that high LINC00278 expression inhibited both ETS1 expression and phosphorylation. COL4A1/COL4A2 were identified as potential downstream targets of LINC00278. Meanwhile, the LINC00278/COL4A1/COL4A2-dominated low-risk group showed higher antigen-presenting activity and a higher immune score than the high-risk group. The findings indicated that ETS1 upregulated LINC00278 expression on the Y chromosome, which in turn inhibited LSCC growth *in vivo* and *in vitro* by inhibiting the AKT/mTOR signaling pathway via downregulation of COL4A1/COL4A2.

*Key words:* LSCC, LINC00278, transcription factor, ETS1, COL4A1, COL4A2

Laryngeal cancer is the second most common malignancy in otorhinolaryngology-head and neck cancers, with the squamous cell carcinoma subtype accounting for 95% of all larynx cancers [1]. The causes of laryngeal cancer are complex and multifaceted. Furthermore, the relevance of early clinical signs is unknown, and survival rates at later stages remain poor [2]. Over the last three years, the overall number of new cases of LSCC has slightly decreased due to the introduction of smoking restrictions. However, the incidence remains high, accounting for approximately 1% of all cancers [3]. As a result, it is critical to continue investigating novel treatment targets and useful prognostic indicators.

The cost of sequencing has dropped significantly as a result of technological advancements, and massive quantities of transcriptome data have been generated concurrently. Typically, these RNA data are derived from a variety of platforms, each of which has a unique representation and statistical properties. The simple intersection of multiple expression RNAs from various platforms may not be an efficient method for identifying an optimal tumor marker [4]. As a result, a computational method is required that allows a thorough analysis of this data as well as accurate integration of data produced across many platforms. To address this problem and obtain stable biomarkers, Kolde developed

a novel rank aggregation method referred to as robust rank aggregation (RRA), to identify overlapping genes across ranked gene lists [5]. Several reliable biomarkers in breast cancer and hepatocellular carcinoma were discovered using a combined analysis of multiple microarray datasets from the GEO database. In this study, we performed a comprehensive analysis of the LSCC transcriptome data in TCGA, given that TCGA database included complete gene expression profiles and clinical information of LSCC, and selected genes were subsequently validated using RRA techniques in external GEO cohorts.

Long non-coding RNAs (lncRNAs) are a class of non-coding RNAs >200 nucleotides in length and are transcribed by RNA polymerase II. LncRNAs play an important role in a variety of diseases, including cancer [6]. In our previous study, we examined the expression and functions of many lncRNAs in LSCC and found that they played important roles in the regulation of mRNA transcription and protein translation [7–9]. Given that nearly all LSCC patients are males, and the larynx is a marker of male secondary sexuality, it remains unclear if laryngeal cancer is a sex hormone-dependent tumor [10]. Meanwhile, there have been numerous reports in the literature that genes on the Y chromosome play significant roles in tumor progression [11–13]. Furthermore, a recent study identified a tumor suppressor gene LINC00278 on the Y chromosome for male esophageal squamous cell carcinoma, indicating that long noncoding RNA encoded by the Y chromosome may be implicated in male-dominant tumors [14]. However, considerably less research has been conducted to date on the role of Y-linked lncRNA in LSCC.

The phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K)/protein kinase B (Akt) signaling pathway, which regulates cell proliferation, survival, and migration, is one of the most commonly activated signaling pathways in a wide variety of human cancers. Additionally, this pathway may be activated by many cell membrane proteins, including the integrin receptor and G protein-coupled receptor. When PI3K is activated, a signal transduction cascade is initiated, with the kinases AKT and mTOR acting as key effectors [15]. Previous studies reported that the 5-hydroxytryptamine receptor promotes LSCC growth by activating the PI3K/AKT pathway [16]. Furthermore, FADS1 has also been shown to play an oncogenic role in the progression of laryngeal cancer through the AKT/mTOR signaling [17]. The PI3K/

Akt signaling pathway, which is active in LSCC, regulates the development and progression of LSCC.

The findings indicated that LINC00278 on the Y chromosome inhibits LSCC growth *in vivo* and *in vitro* by inhibiting the AKT/mTOR signaling pathway and downregulating COL4A1/COL4A2 in LSCC. Furthermore, LINC00278 and its downstream target genes were found to increase the upregulation of alternative splicing events, which regulate the infiltration levels of tumor-infiltrating immune cells and can therefore affect the antitumor immune response.

## Patients and methods

**Data acquisition from a public database and bioinformatics analysis.** TCGA provided FPKM data for 123 LSCC samples as well as pertinent clinical data. Genes with  $|\log_2\text{foldchange}(\log_2\text{FC})| > 1$  and adjusted  $p < 0.05$  were considered differentially expressed genes (DEGs). The gene matrix expression profiles of LSCC were obtained from GEO (GSE51985, GSE59102, GSE59652, GSE84957, GSE117005, GSE143224, See Table 1 for specific information). All genes in each data set were sorted by logFC and integrated using the RobustRankAggreg R package. Using hTFtarget database (<http://bioinfo.life.hust.edu.cn/hTFtarget#!/>) transcription factors binding to the LINC00278 promoter were predicted.

**Identification of prognosis-related genes.** The R-package ‘survival’ was used to perform univariate and multifactorial Cox regression analysis. A six-gene-based prognostic model was used to estimate each patient’s risk score using the equation: risk score = MYHAS $\times(-0.19)$ +MNX1-AS1 $\times 0.19$ +LINC02086 $\times 0.18$ +LSAMP-AS1 $\times 0.13$ +LINC00278 $\times(-0.28)$ +CASC20 $\times(-0.03)$

Based on the median risk score, the patients were subdivided into the high-risk and low-risk groups. The ‘survivalROC’ package in R was used to compute the area under the curve (AUC) and draw the receiver operating characteristic (ROC) curve. The AUC value showed predictive accuracy and was considered significant when it exceeded 0.60 [18]. Kaplan-Meier method was used to estimate the difference in survival time between low- and high-risk groups using the ‘Survminer’ package in R. Meanwhile, the ‘forestplot’ package in R was used to visualize the Cox results. Nomograms were constructed using the ‘rms’ package in R. The R package ‘estimate’ was used to determine the stromal score, immune

**Table 1. Details of the GEO LSCC data sets.**

Reference	GEO	Platform	Normal	Tumor	Biotype
Lian et al. (2013) [37]	GSE51985	GPL10558	3	10	mRNA
Wilson (2014) [38]	GSE59102	GPL6480	13	25	mRNA
Shen et al. (2014) [39]	GSE59652	GPL13825	7	6	lncRNA + mRNA
Feng et al. (2016) [40]	GSE84957	GPL17843	9	9	lncRNA + mRNA
Liu et al. (2020) [41]	GSE117005	GPL20115	5	4	mRNA
Nicolau et al. (2020) [42]	GSE143224	GPL5175	11	14	mRNA

score, and tumor purity. CIBERSORT (<https://cibersort.stanford.edu>) was used to estimate immune infiltration data.

**Patient samples.** A total of 72 pairs of LSCC tissues and noncancerous tissues were frozen at the Institute of Otorhinolaryngology-Head and Neck Surgery (Hebei, China). Clinical features and pathological diagnoses were extracted from corresponding data records. Before surgery, there was no chemotherapy and/or radiotherapy treatment received. This study was approved by the Ethical Committee of the Second Hospital of Hebei Medical University (Hebei, China), and all patients provided informed consent.

**Cell culture, plasmid transfections, and treatments.** The LSCC cell lines (TU686, TU177, and AMC-HN-8), hypopharyngeal cell lines (FaDu), and 293T cell lines were purchased from the Beijing Beina Chuanglian Institute of Biotechnology (Beijing, China). These cell lines were cultured as described in previous protocols [7]. The pcDNA3.1-LINC00278 expression vector was designed by GENEWIZ (Suzhou, China). The Lipofectamine3000 (Invitrogen, USA) was used to carry out the cell transfections following the manufacturer's instructions. Cells transfected with the corresponding plasmid for 48 h and/or treated with LY294002 (25  $\mu$ M, HY-10108, MCE, USA) and SC79 (4  $\mu$ g/ml, B5663, APEX-BIO, USA) for 24 h were subjected to cell function assays.

**Assays for cell biological behavior.** The experimental procedure for the cell proliferation, migration, and invasion assays was performed as previously described [7]. For the cell number assay, 2,000 cells in a 200  $\mu$ l of medium/well were seeded in 96-well plates and cultured in the complete medium. At the specified time points, 20  $\mu$ l MTS (Promega, USA) was added to each well and incubated for 2 h. A microplate reader (TECAN, Switzerland) was then used to measure the absorbance at OD 490 nm. The transfected LSCC cell lines were collected for the Transwell migration experiments by trypsinization, then re-suspended in a serum-free medium and cell counting. A total of  $1 \times 10^5$  cells were placed into the upper chamber (Corning, USA), and a 630  $\mu$ l complete culture medium was added to the lower chamber. The Transwell chamber was taken out after incubation at 37°C for 24 h, washed with PBS, fixed with 4% paraformaldehyde for 20 min, and stained with 1% crystal violet for 20 min. The stained cells were examined and manually counted. For the invasion assay, 50  $\mu$ l of Matrigel (Corning, USA) was applied to the upper chamber to form a matrix barrier; the same procedures were then used in the cell migration experiment. For the clone-forming assay, 2,000 cells were seeded in 6-well plates, incubated for 10–14 days, and stained after fixation. When the cells in 6-well plates reached 70–80% confluence for the wound-healing assay, they were scratched using a 200  $\mu$ l pipette tip. At 0 and 48 h, the scratched areas were photographed and measured.

**RNA-extraction and qPCR.** Total RNA extraction and the cDNA synthesis were performed as previously described [7]. GoTaq<sup>®</sup>RT-qPCR Master Mix (Promega, USA) was used for qPCR under standard thermocycling conditions.

GAPDH was utilized as an internal control in the qPCR results. Supplementary Table S1 presents a list of all of the primer sequences.

**Subcellular fractionation and subcellular localization.** The cytoplasmic and nuclear RNAs were extracted using a PARIS<sup>™</sup> Protein and RNA Isolation System (Invitrogen, USA) according to the kit instructions. GAPDH and U6 were used as the cytoplasmic and nuclear internal references, respectively.

**Construction of luciferase reporter vector and the dual-luciferase reporter assay** The LINC00278 promoter region and first exon were amplified and cloned into the pGL3-basic vector using the primers indicated in Supplementary Table S1. The dual-luciferase reporter assay was carried out according to the instructions in the protocol (Promega, USA). Briefly, 293T and AMC-HN-8 cells were co-transfected with ETS1, the above-constructed plasmids, and pRL-TK Renilla luciferase plasmid for 48 h. Whole-cell lysates were prepared and luciferase activity was detected using a dual-luciferase assay system.

**The chromatin immunoprecipitation (ChIP) assay.** The ChIP assay was carried out in accordance with the Magna ChIP<sup>™</sup> A/G (Millipore, Germany) protocol. Briefly, 48 h after ETS1 transfection, the cells were fixed, harvested, and sonicated (10 s sonication at 10 s interval) 18 times at 30% amplitude to obtain appropriate fragments. The prepared chromatin was precipitated overnight with ETS1 antibodies and 1% lysates as input. The binding complexes were then extensively washed, eluted, purified, and analyzed by qPCR. The promoter primers for LINC00278 are described in Supplementary Table S1.

**Western blotting.** RIPA lysis buffer supplemented with proteinase and phosphatase inhibitors was used to prepare total proteins. Total protein concentration was determined using the BCA assay method. Proteins were separated on sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to polyvinylidene difluoride membranes. The membranes were blocked using 5% non-fat milk at room temperature for 2 h, incubated with primary antibodies at 4°C overnight, and with horseradish peroxidase-labeled secondary antibodies. Electrochemiluminescence (ECL) was used to expose the bands, which were then examined using ImageJ software. The following primary antibodies were used: COL4A1 (1:2000, AF0510, Affinity Biosciences, USA), N-cadherin (1:10000, 22018-1-AP, Proteintech, USA), E-cadherin (1:6000, 20874-1-AP, Proteintech, USA), Vimentin (1:10000, 10366-1-AP, Proteintech, USA),  $\beta$ -catenin (1:10000, 51067-2-AP, Proteintech, USA), SNAIL (1:2000, 26183-1-AP, Proteintech, USA), ZEB1 (1:2000, 21544-1-AP, Proteintech, USA), AKT (1:1000, ab8805, Abcam, USA), p-AKT (1:1000, ab81283, Abcam, USA), mTOR (1:1000, 2983, Cell Signaling Technology, USA), p-mTOR (1:1000, 5536, Cell Signaling Technology, USA), ETS1 (1:1000, 14069, Cell Signaling Technology, USA), and p-ETS1 (1:1000, #11658, Signalway Antibody, USA).

**Animal experiments.** Twenty BALB/c mice, aged 5 to 6 weeks, were purchased from Charles River Laboratories (Beijing, China). Each group consisted of 10 mice for the tumor growth studies. Mice in the experimental group (n=10) were injected with 100  $\mu$ l AMC-HN-8 cells transfected with LINC00278 lentivirus, while mice in the control group (n=10) received 100  $\mu$ l AMC-HN-8 cells transfected with control lentivirus. The tumor size was measured once every three days, and the tumor volume was calculated using the simplified formula of a rotational ellipsoid (length $\times$ width<sup>2</sup> $\times$ 0.50). The xenografts were removed from the mice 3–4 weeks after implantation and weighed. The experimental procedures were approved by the Ethics Review Committee of Hebei Medical University.

**Immunohistochemistry.** Paraffin-embedded tumor tissues from nude mice were cut into 4  $\mu$ m slices by a paraffin slicer, then baked at 95°C for 30 min. The baked slices were deparaffinized using a transparent agent, and hydrated with graded alcohols. After that, the sections were boiled in citrate buffer for 3 min for antigen retrieval. Next, the sections were incubated with a primary antibody (Ki-67, AF1738, 1:150, Beyotime, China) overnight at 4°C. After that, the slides were incubated with the second antibody and the third antibody complex at room temperature for 10 min. Coloration was achieved through staining with DAB for 3 min. After hematoxylin counterstaining and dehydration, the sections were sealed with coverslips.

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

## Results

**lncRNA DEGs in LSCC patients and construction of a prognostic model.** A total of 790 differentially expressed long noncoding RNAs were screened from TCGA database, 665 upregulated and 125 downregulated in LSCC tissues compared with adjacent non-tumor tissues. In TCGA dataset, a heatmap was shown to depict the relationship between clinical factors and DEGs lncRNAs expression levels (Supplementary Figure S1). Additionally, survival analysis was carried out independently using univariate Cox regression analysis and the Kaplan-Meier method with 30 survival-related lncRNAs overlapping. Next, we compared the expression of DEGs to the trends of corresponding survival curves, and 14 genes exhibiting opposing survival trends were excluded, leaving 16 candidates for further analysis (heatmap is presented in Supplementary Figure S2A). Then, based on the univariate Cox outcomes, multi-gene prediction models were constructed to assess the impact of the screened genes on patient overall survival. LINC00278, MYHAS, MNX1-AS1, LINC02086, LSAMP-

AS1, and CASC20 were selected to construct the model (lncRNA-related prognostic model), and risk scores for the patients were calculated using the previously described formulas (Supplementary Figure S2B).

The AUC for each ROC curve was evaluated to assess the performance of the lncRNA-related prognostic model. The AUC of the lncRNA-related prognostic model (area under red curve = 0.832) was significantly higher than the models based on TNM stages (area under yellow curve = 0.614), grade (area under green curve = 0.633), gender (area under orange = 0.485), alcohol (area under blue = 0.421), and cigarettes (area under pink = 0.511) (Supplementary Figure S2C). Furthermore, patients were segregated into the high- and low-risk groups based on the median risk score value, and the corresponding risk value was presented in ascending order (Supplementary Figure S2D). According to the survival curves, those with high-risk scores had a significantly lower survival rate (Supplementary Figure S2E). Patient survival status demonstrated that patients with high-risk scores had higher mortality than patients with low-risk scores (Supplementary Figure S2F). The gene expression patterns of the high-risk and low-risk groups were then shown on a heatmap. The expression of LINC00278 and MYHAS increased as the risk decreased, while MNX1-AS1, LINC02086, LSAMP-AS1, and CASC20 exhibited the reverse behavior (Supplementary Figure S2G).

**lncRNA-related prognostic model evaluated as an independent prognostic factor and construction of nomogram model.** Multivariate Cox regression analysis was used to estimate the lncRNA-related prognostic model, as well as age, gender, TNM stage, alcohol use, and grade. Consequently, the prognostic model (p<0.001) was significantly associated with the prognosis and was the most important independent prognostic factor for overall survival (Supplementary Figure S2H). Furthermore, the nomogram was constructed for predicting patient prognoses based on the prognostic model, age, gender, TNM stage, alcohol, and grade. The nomogram can be used to calculate the score of each component as well as the overall score of all factors. The total scores predicted the 1-, 3-, and 5-year survival rates (Supplementary Figure S2I). We verified the six prognostic model genes by integrating the two GEO microarray datasets GSE59652 and GSE84957 with the RRA package because the quantity of normal tissue in the TCGA database was significantly different from tumor tissue and there were some false-positive genes among these RNA-sequencing data from TCGA (Supplementary Figure S3A). Among the 6 lncRNAs, only LINC00278 was lowly expressed, while CASC20 was highly expressed in both TCGA and GEO databases. Furthermore, the male LSCC incidence rate was higher than that of females, and it is debatable if laryngeal cancer is a sex hormone-dependent tumor. qPCR results demonstrated that the androgen receptor (AR) levels were significantly higher in LSCC tissues compared to adjacent non-tumor tissues (Supplementary Figure S3B). LINC00278 gene (chromo-

some Yp11.2) was selected as the research target in this study, after comprehensive consideration.

**LINC00278 expression and the relationship between LINC00278 and clinical features.** GTEX data indicated that LINC00278 was differently expressed in various tissues (Supplementary Figure S3C), and the qPCR findings suggested that LINC00278 expression levels were significantly lower in LSCC tissues compared to adjacent nontumor tissues (Figure 1A). The relationship between LINC00278 expression and clinical parameters was also investigated in 72 paired LSCC patients. Patient detailed information is listed in Table 2. Statistical results revealed that LINC00278 expression was significantly associated with the TNM stage ( $p < 0.001$ ), lymphatic metastasis ( $p < 0.01$ ), and pathological differentiation ( $p < 0.01$ ). Furthermore, no correlation was found between age, smoking, alcohol use, and the site of the carcinoma (Figure 1B). We performed a survival analysis using the median value (0.146) of the qPCR ( $2^{-\Delta\Delta Ct}$ ) results and follow-up data. Patients with low LINC00278 expression had a poorer overall survival (OS) than patients with high LINC00278 expression (Figure 1C).

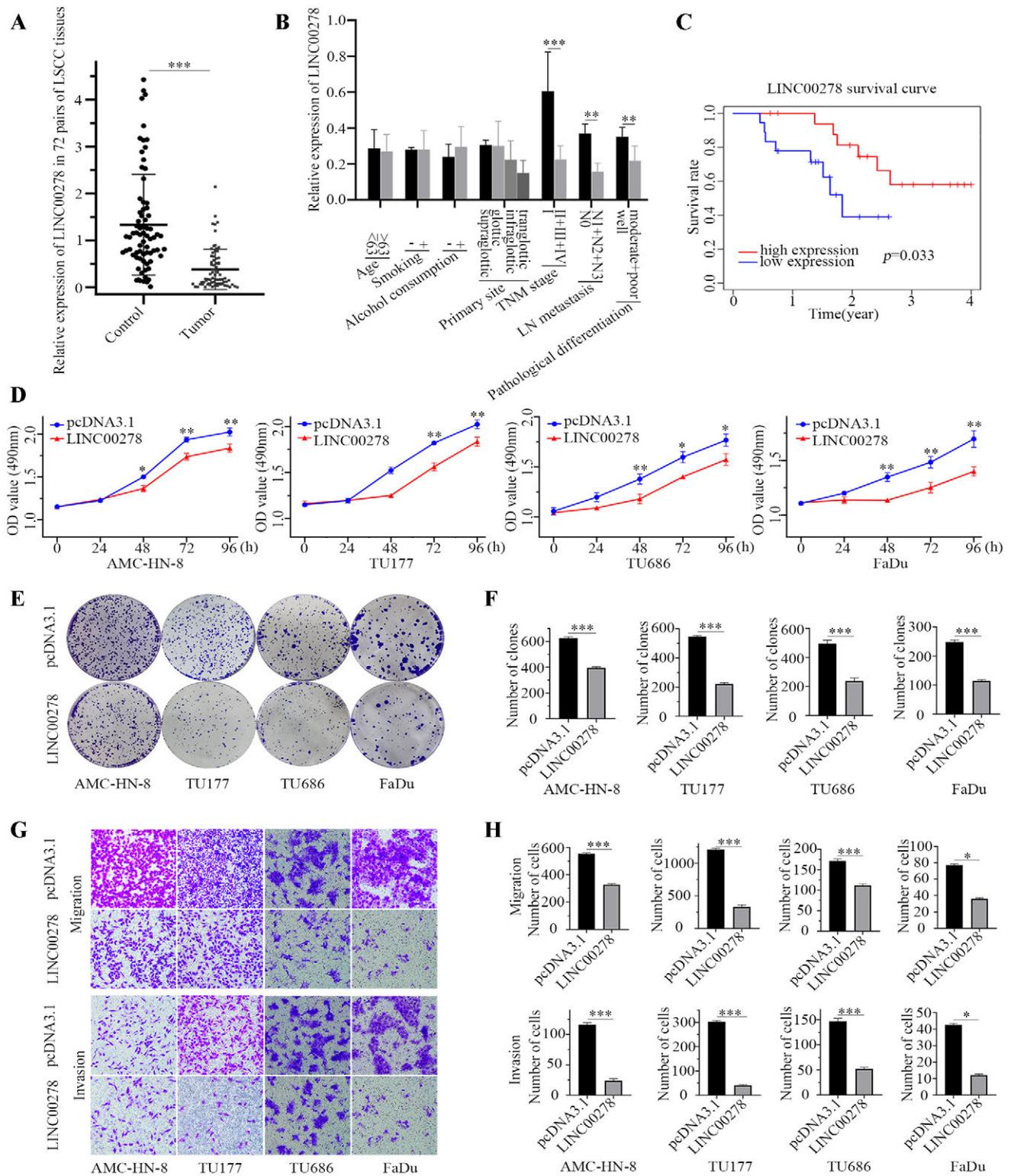
**LINC00278 overexpression inhibits cell proliferation, migration, and invasion *in vitro*.** We hypothesized that LINC00278 may play an essential role in regulating cell proliferation, migration, and invasion of LSCC cells based on their distinct expression. Considering that hypopharyngeal squamous cell carcinoma (HSCC) is a head and neck malignant tumor with one of the worst prognoses, LINC00278 was found to be lowly expressed in TCGA head and neck squamous cell carcinoma dataset (Supplementary Figure S3D). In addition to LSCC, we also investigated the effect of LINC00278 in HSCC cell lines. AMC-HN-8, TU177, TU686, and FaDu cell lines were transfected with well-designed pcDNA3.1-LINC00278 or pcDNA3.1. The efficiency of overexpression was then verified using qPCR (Supplementary Figure S3E). MTS assays revealed that LINC00278 overexpression significantly reduced the proliferative ability of the four cell lines (Figure 1D) while showing no effect on 293T cells (Supplementary Figure S4A). Meanwhile, the re-expression of LINC00278 significantly reduced the number of colonies produced by AMC-HN-8, TU177, TU686, and FaDu cell lines (Figures 1E, 1F). The effect of LINC00278 on the migration and invasive ability of squamous cells carcinoma was assessed using Transwell and wound-healing assays. Exogenous overexpression of LINC00278 significantly decreased the migratory and invasive capabilities of the four cell lines (Figures 1G, 1H). Furthermore, in the scratch wound-healing assay after 48 h, the LINC00278 overexpression attenuated the migration potential in the four cell lines (Figures 2A, 2B). Among the four cell lines, AMC-HN-8 and FaDu were the two defined male cell lines, thus mRNA and protein expression of EMT-related genes were detected in these cell lines. In the LINC00278 group, E-cadherin mRNA and protein expression were higher (E-cadherin expression was not detected in AMC-HN-8 cells), while N-cadherin, Vimentin,

Twist, Zeb1, and Snail mRNA and protein expression were lower than in the pcDNA3.1 group. When the expression of LINC00278 was intentionally altered in AMC-HN-8 cells, neither mRNA nor protein of  $\beta$ -catenin levels were altered but significantly reduced in FaDu cells (Figures 2C–2E). These findings suggested that re-expression of LINC00278 may inhibit LSCC cell invasion and migration.

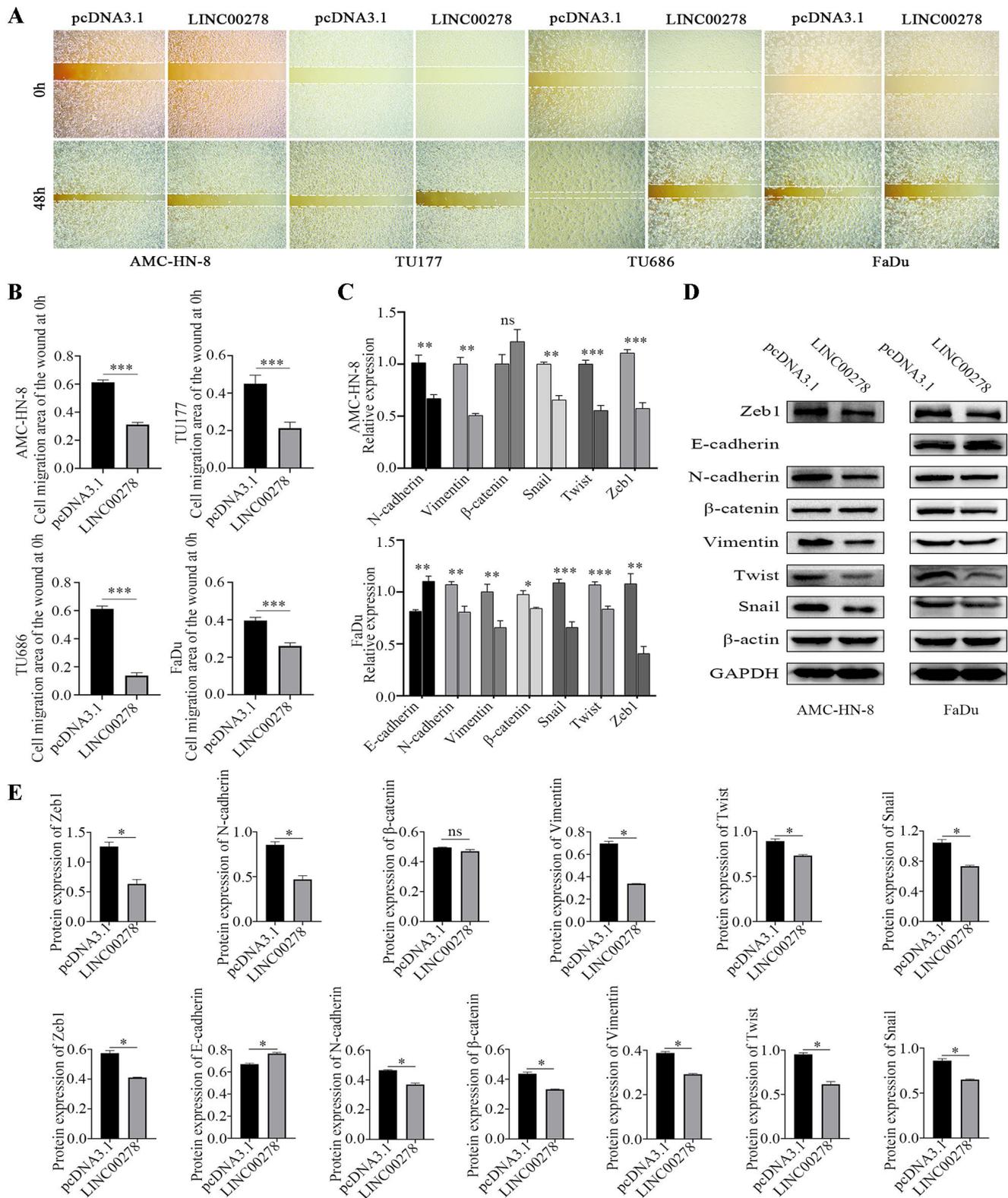
**ETS1 promotes the expression of LINC00278 by binding to the LINC00278 first exon region.** Based on the above findings, we found that the downregulation of LINC00278 was related to the development of LSCC, indicating the underlying molecular mechanism needs to be investigated. Although LINC00278 contained multiple m6A modification sites, these modifications were related to encoded polypeptides rather than regulating LINC00278 expression. According to the UCSC Genome Browser, the promoter region of LINC00278 does not contain a CpG island and no histone changes occur (data not shown). As a result, it is critical to further explore transcription factors to construct a comprehensive regulatory network. We screened out the potential transcription factors of LINC00278 using hTFtarget database and RNAInter database. Two genes: ETS1 and RUNX1 were eventually selected based on the number of human tissue sources. Because RUNX1 overexpression did not affect the expression of LINC00278, ETS1

**Table 2. Clinical information in 72 cases of LSCC.**

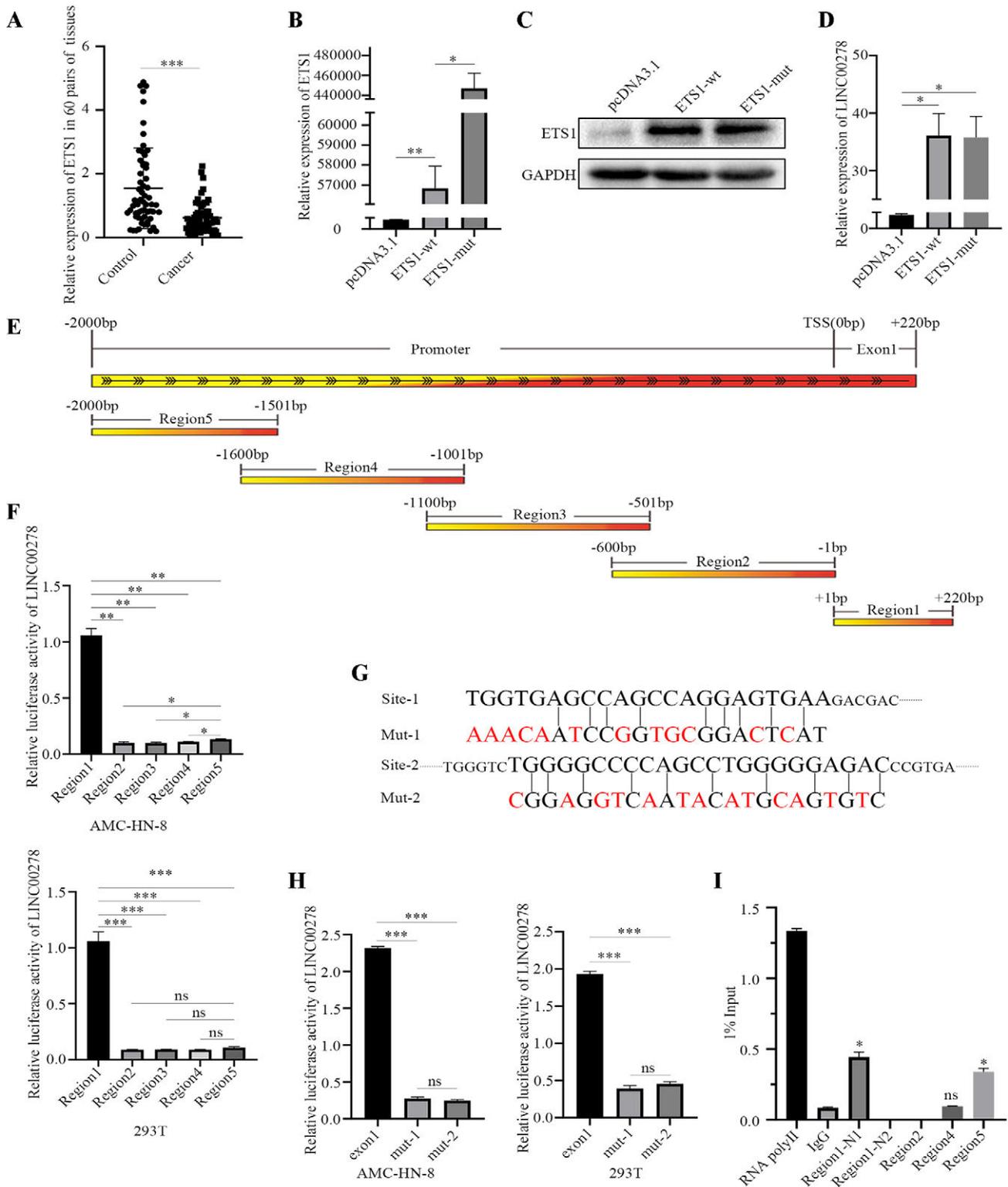
Characteristics	No. cases (100%)
Gender	
male	72 (100.00)
female	0 (0.00)
Age	
<63	34 (47.22)
≥63	38 (52.78)
Smoking	
no	7 (9.72)
yes	65 (90.28)
Alcohol	
no	26 (36.11)
yes	46 (63.89)
Primary site	
supraglottic	33 (45.83)
glottic	30 (41.67)
infraglottic	6 (8.33)
transglottic	3 (4.17)
TNM stage	
I	15 (20.83)
II+III+IV	57 (79.17)
LN metastasis	
N0	45 (62.50)
N1+N2+N3	27 (37.50)
Pathological differentiation	
well	42 (58.33)
moderate+poor	30 (41.67)



**Figure 1.** Overexpression of LINC00278 inhibits cell proliferation, migration, and invasion *in vitro*. **A)** Relative expression of LINC00278 in 72 paired LSCC tissues. **B)** The relation between the expression of LINC00278 and clinical parameters based on qPCR results. **C)** Kaplan-Meier estimated overall survival in patients with high or low LINC00278 expression. ( $*p<0.05$ ,  $**p<0.01$ ,  $***p<0.001$ ). **D)** The cell viability was assessed by MTS assay after being transfected with pcDNA3.1-LINC00278. **E, F)** The number of colonies was remarkably decreased after LINC00278 was upregulated. **G, H)** Upregulated the expression of LINC00278 attenuates the invasion and migration capability *in vitro*. ( $*p<0.05$ ,  $**p<0.01$ ,  $***p<0.001$ ).



**Figure 2.** Plasmid transfection of LINC00278 significantly downregulated EMT-related mRNAs and proteins in AMC-HN-8 and FaDu cell lines. A, B) The suppressed migration capability in cells treated with pcDNA3.1-LINC00278 was demonstrated by wound-healing assays. C-E) Expression changes of EMT-related mRNAs and proteins in AMC-HN-8 and FaDu cell lines. (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).



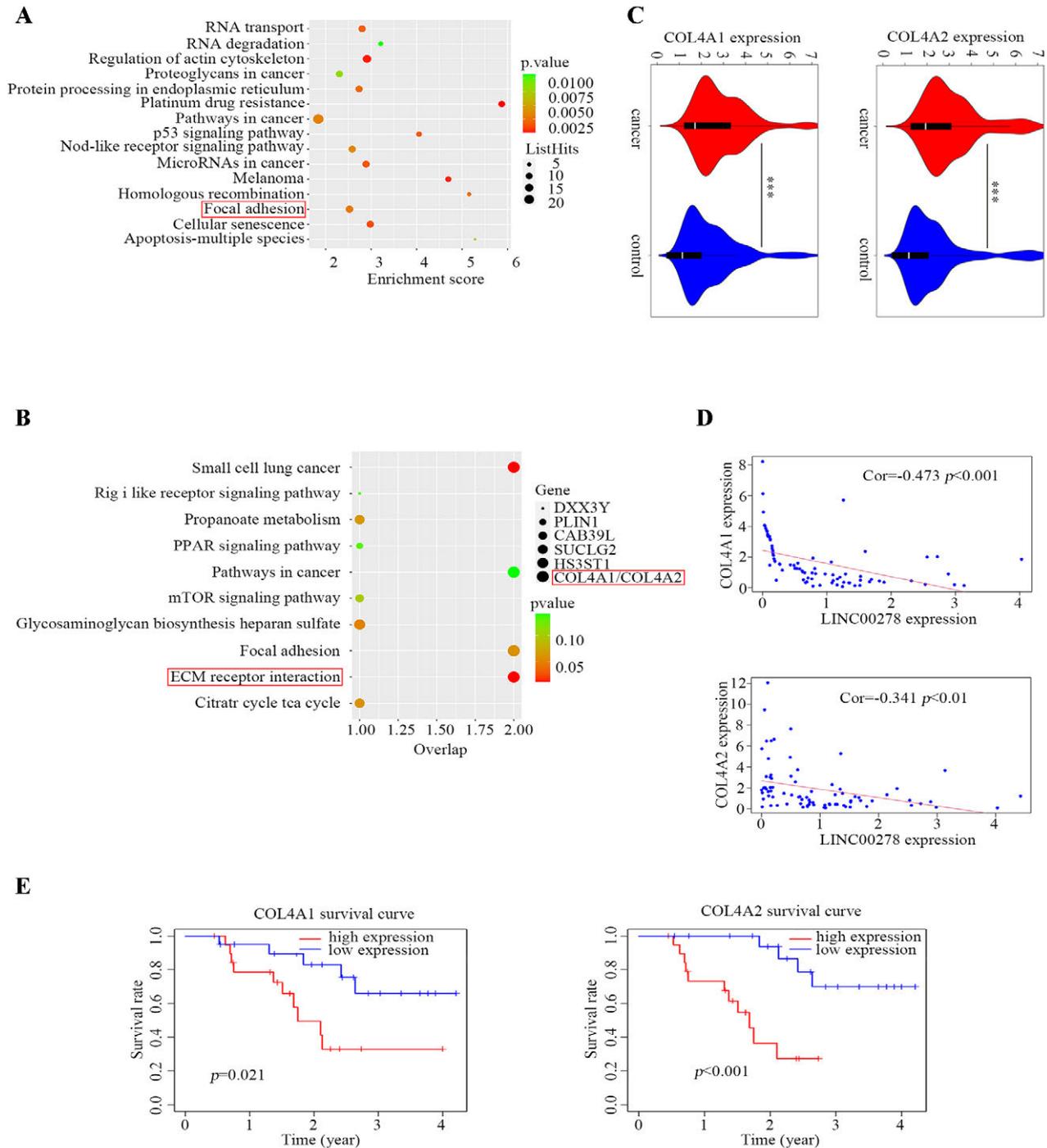
**Figure 3.** ETS1 promotes the expression of LINC00278 by binding to the LINC00278 first exon region. **A)** Relative expression of ETS1 in 60 paired LSCC tissues. **B, C)** mRNA expression levels and protein expression levels of ETS1. **D)** Both wild type and mutant ETS1 increased the expression of LINC00278. **E)** Schematic of promoter truncations of LINC00278. **F)** Relative luciferase activity of different LINC00278 promoter region truncations in AMC-HN-8 and 293T cells. **G)** Sequences of transcription factor binding site and mutation. **H)** Luciferase activity of LINC00278\_exon1 after site-directed mutagenesis of ETS1 transcription factor binding sites. **I)** ChIP-qPCR for ETS1 binding at the promoter regions of LINC00278. (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

was finally chosen as the target. ETS1 was found to be poorly expressed in tumor tissues but highly expressed in normal tissues (Figure 3A). Meanwhile, ETS1 exists in both unphosphorylated and hyperphosphorylated forms *in vivo*. Previous studies indicate that phosphorylated ETS1 at threonine 38 plays important roles in the transcriptional regulation of downstream genes [19]. We constructed wild-type ETS1 plasmids and plasmids expressing non-phosphorylated ETS1 mutations plasmid by replacing the serine at 38 with alanine. LINC00278 expression was enhanced in both wild-type and mutant ETS1, indicating that the phosphorylated ETS1 at threonine 38 was not involved in the activation of LINC00278 transcription (Figures 3B–3D). To further elucidate the mechanism of LINC00278 transcription, five truncated variants of the LINC00278 promoter were created, and various truncated promoter regions were cloned into the pGL3-basic plasmid (Figure 3E). Luciferase assays were used to assess ETS1 transcriptional activity on various truncated LINC00278 promoter regions. The results were presented as firefly luciferase activity against Renilla luciferase activity, with promoter region 1 (the first exon) exhibiting the highest activation (Figure 3F). Several transcription factors binding sites were predicted in the first exon of the LINC00278 gene. Plasmids containing the mutations at each binding site were then constructed and used to perform the luciferase reporter assay (Figure 3G). The dual-luciferase reporter assay showed that the ETS1 overexpression significantly reduced the relative luciferase activity of LINC00278-mut, but not LINC00278-wt vectors (Figure 3H). This set of findings indicated a direct interaction between ETS1 and LINC00278, which was validated using a ChIP assay (Figure 3I). These findings indicate that ETS1 is a direct transcriptional target of LINC00278.

**Functional enrichment analysis of the LINC00278 target genes.** LncRNAs played an important role in the process of LSCC progression. To further elucidate LINC00278 growth inhibitory effects on LSCC, we investigated the molecular target of LINC00278. On the one hand, we performed RNA sequencing on AMC-HN-8 cells overexpressing LINC00278, while on the other hand, we performed co-expression analysis on the target genes using TCGA mRNA expression data. KEGG enrichment analysis of co-expression genes (Pearson correlation coefficient > 0.40) and the differential genes of RNA sequencing results ( $|\log_{2}FC| > 2$ ,  $p < 0.05$ ) were performed separately, as shown in Figures 4A and 4B. Given the strong correlation between *Focal adhesion* and *ECM receptor interaction signaling pathways*, we investigated the major signaling molecules involved in these two pathways. Moreover, genes COL4A1 and COL4A2 are mapped to the *ECM receptor interaction signaling pathways*, which are located upstream of the *Focal adhesion signaling pathway*. As a result, COL4A1/COL4A2 were identified as possible LINC00278 downstream targets. GTEx results indicated that COL4A1/COL4A2 were differentially expressed in distinct tissues (Supplementary Figures

S4B, S4C). qPCR was performed in 72 paired LSCC tissues to further verify the expression of COL4A1/COL4A2. The qPCR findings showed that COL4A1/COL4A2 expression levels were markedly upregulated in LSCC tissues as compared to adjacent nontumor tissues (Figure 4C). The GEO database confirmed that COL4A1/COL4A2 were highly expressed in LSCC tissues (Supplementary Figure S5A). Correlation analysis revealed a strong and negative correlation between COL4A1/COL4A2 mRNA and LINC00278 (Figure 4D). The relationship between COL4A1/COL4A2 expression and clinical parameters was also investigated. COL4A1/COL4A2 expression was found to be significantly associated with the TNM stage, lymphatic metastasis, and pathological differentiation (Supplementary Figures S5B, S5C). Additionally, overall survival analyses for COL4A1/COL4A2 were also performed. Patients with a high level of COL4A1/COL4A2 expression had a poorer OS than those with a low COL4A1/COL4A2 expression (Figure 4E). Then, we examined the COL4A1/COL4A2 expression levels in four human LSCC cell lines. The qPCR analysis revealed that the expression of COL4A1/COL4A2 mRNA was highest in AMC-HN-8 cells, followed by TU177 and TU686 cells, while FaDu cells had the lowest expression (Figures 5A, 5B). Therefore, AMC-HN-8 was selected for further investigations. qPCR was used to determine the expression of COL4A1/COL4A2 in AMC-HN-8 cells 48 h after transient transfection with LINC00278. The degree of the reduction was more pronounced for COL4A1 than for COL4A2 (Figures 5C, 5D). Additionally, since COL4A1 and COL4A2 are bidirectional gene pairs, they can share promoters or transcription factor binding sites, and their expression is highly correlated when expressed concurrently. COL4A1 antibodies were used to detect the protein level expression of COL4A1. Western blot results were coincident with that of qPCR analysis (Figures 5E, 5F). COL4A1 was selected for further analysis in AMC-HN-8 cells, to elucidate the molecular mechanism underlying LINC00278 modulation.

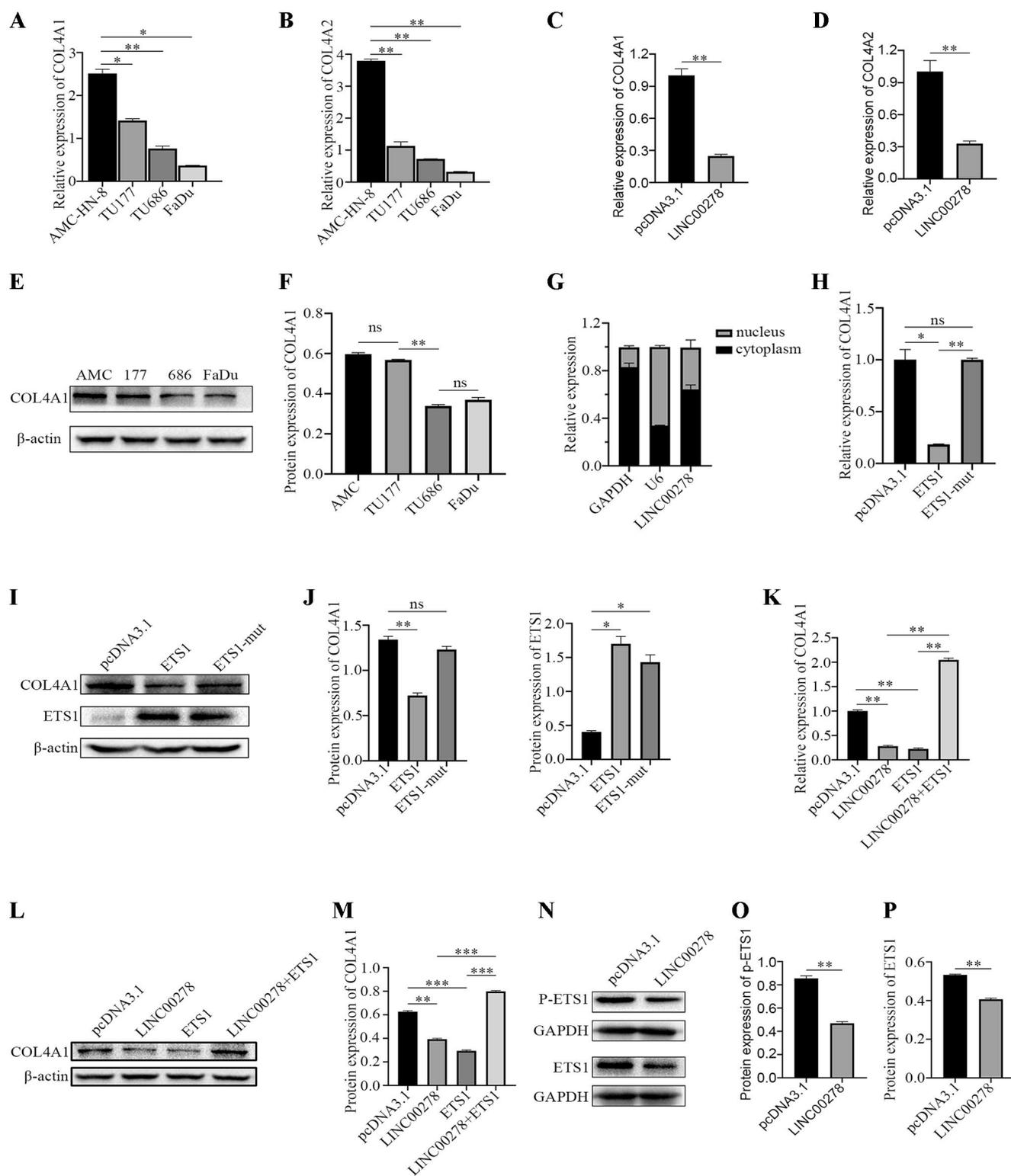
**LINC00278 inhibits COL4A1 mRNA and protein expression by inhibiting ETS1 phosphorylation.** To elucidate the mechanism by which LINC00278 regulates COL4A1 to inhibit the malignant progression of LSCC, we determined LINC00278 subcellular localization. The findings revealed that LINC00278 was primarily localized in the cytoplasm of AMC-HN-8 cells, which was consistent with the anticipated localization results obtained using the lncLocator online website (<http://www.csbio.sjtu.edu.cn/bioinf/lncLocator/>) (Figure 5G). Therefore, we hypothesized that overexpression of LINC00278 enhanced its accumulation in the cytoplasm, preventing the transcription factors of target genes from translocating to the nucleus and therefore significantly reducing COL4A1 expression. A similar method was used to predict the transcription factors that regulate COL4A1 expression, and COL4A1 expression was also regulated by ETS1. qPCR was used to determine the expression of COL4A1 in AMC-HN-8 cells 48 h following transient transfection of



**Figure 4.** Functional enrichment analysis of the LINC00278 target genes. A) KEGG enrichment analysis of differentially expressed genes after overexpressing LINC00278 in AMC-HN-8 cells. B) KEGG enrichment analysis of LINC00278 co-expressed genes in LSCC tissues from TCGA database. C) Relative expression of COL4A1/COL4A2 in 72 paired LSCC tissues. D) Pearson correlation analysis of LINC00278 and COL4A1/COL4A2 expression. E) Kaplan-Meier overall survival analysis of the association between COL4A1/COL4A2 expression level and LSCC patient survival.

ETS1-wt and ETS1-mut. The levels of COL4A1 mRNA and protein expression were substantially reduced when ETS1-wt was overexpressed, but not in the case of ETS1-mut overexpression (Figures 5H–5J). These results showed that ETS1

inhibits COL4A1 transcriptional activity, depending on the amount of phosphorylated ETS1 at threonine 38. However, the precise regulatory relationship between phosphorylated ETS1 and COL4A1 still needs further verification using



**Figure 5.** LINC00278 lowered the mRNA and protein expression of COL4A1 by decreasing phosphorylation of ETS1. A, B) mRNA levels of COL4A1/ COL4A2 in AMC-HN-8, TU177, TU686, and FaDu cell lines. C, D) LINC00278 lower the mRNA expression of COL4A1/ COL4A2 in AMC-HN-8 cells. E, F) Protein levels of COL4A1/ COL4A2 in AMC-HN-8, TU177, TU686, and FaDu cell lines. G) The nuclear and cytosolic localization of LINC00278. Nuclear controls, U6, cytosolic controls, GAPDH. H–J) COL4A1 mRNA and protein expression levels following ETS1 and ETS1-mut overexpression in AMC-HN-8 cells. K–M) COL4A1 mRNA and protein expression levels after LINC00278 or ETS1 were transfected alone or LINC00278 and ETS1 were co-transfected. N–P) LINC00278 could reduce both ETS1 expression and phosphorylation. (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

CHIP and luciferase reporter gene assay. When LINC00278 or ETS1 was transfected alone, there was a significant decrease in COL4A1 expression relative to the empty vector plasmid. Surprisingly, when LINC00278 and ETS1 were co-transfected, a substantial increase in COL4A1 expression at the mRNA and protein levels was seen (Figures 5K–5M). Then, how does LINC00278 influence ETS1 regulation during LSCC development? Western blot analysis demonstrated that a higher LINC00278 expression can reduce both ETS1 expression and phosphorylation (Figures 5N–5P). As previously demonstrated, ETS1 overexpression elevated LINC00278 expression in a phosphorylation-independent manner, while overexpression of LINC00278 decreased the expression levels of ETS1. Additionally, both LINC00278 and ETS1 were expressed at a low level in LSCC tissues. Taken together, these results indicate that LINC00278 and ETS1 form a negative feedback loop that eventually suppresses their expression, thus maintaining their low expression levels in LSCC.

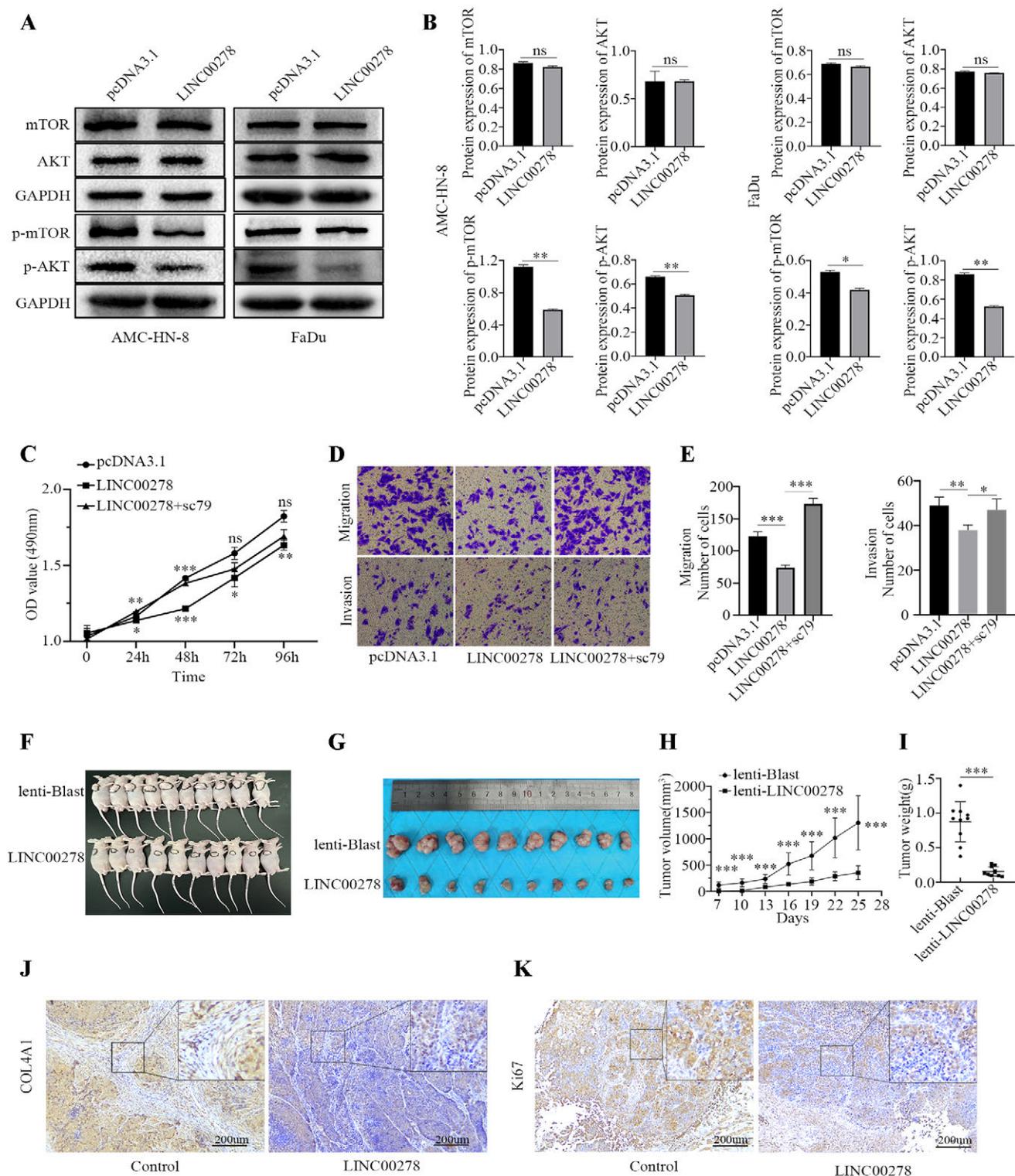
**The Akt/mTOR pathway is involved in the effect of LINC00278 re-expression on LSCC cell functions.** The COL4A1/COL4A2 genes are located upstream of the AKT/mTOR signaling pathway [20]. Therefore, we evaluated the relationship between LINC00278 and the PI3K-Akt signaling pathway in LSCC progression. First, AMC-HN-8 cells were treated with the SC79 (Akt agonist) and LY294002 (PI3K inhibitor) to determine their effects on cell proliferation, migration, and invasion. SC79 promoted cell proliferation, migration, and invasion, while LY294002 inhibited these effects (Supplementary Figures S5D–S5F). These results imply that the PI3K/Akt signaling pathway was activated to promote cell proliferation, migration, and invasion in LSCC. Next, we investigated the effects of LINC00278 on the PI3K/Akt signaling pathway in AMC-HN-8 and FaDu cells. Treatment with LINC00278 for 48 h significantly suppressed p-AKT and p-mTOR protein levels (Figures 6A, 6B), implying that re-expression of LINC00278 suppressed the AKT/mTOR signaling pathway in LSCC. Furthermore, the effects of LINC00278 on LSCC cell proliferation, migration, and invasion in the presence of SC79 were evaluated. These findings suggested that SC79 partially restored the LINC00278-induced inhibitory effects in AMC-HN-8 cells (Figures 6C–6E). In conclusion, LINC00278 suppresses LSCC cell proliferation, migration, and invasion through the PI3K/Akt/mTOR signaling.

**Overexpression of LINC00278 suppressed tumor growth *in vivo*.** To elucidate the *in vivo* roles of LINC00278 in tumorigenesis, we determined whether LINC00278 overexpression suppresses tumor growth using LSCC xenograft tumor models. AMC-HN-8 cells with stably overexpressed LINC00278 or empty vector, generated by lentiviral vector transduction, were inoculated into nude mice models. Then, the animals were sacrificed and imaged (Figures 6F, 6G). In agreement with *in vitro* results, mean tumor weights and volumes were significantly reduced in the LINC00278

overexpressed group compared to the empty vector group, indicating that LINC00278 suppresses LSCC xenograft tumor growth *in vivo* (Figures 6H, 6I). qPCR analysis showed that LINC00278 overexpression did not result in a steady decrease in expressions of either COL4A1 or COL4A2 in tissues (Supplementary Figures S5G–S5I). In addition, immunohistochemistry showed that the expressions of collagen IV and Ki-67 were suppressed following LINC00278 overexpression (Figures 10J, 10K).

Low LINC00278 and high COL4A1/COL4A2 expression correlated with immune responses in human LSCC

We have shown that LINC00278 negatively regulates COL4A1 and COL4A2, which are its downstream target genes. To determine the significance of LINC00278, COL4A1, and COL4A2 in LSCC progression, survival curves of the three co-expressed genes were plotted. Based on the expression levels of COL4A1, COL4A2, and LINC00278, the 110 TCGA LSCC cases were allocated into two groups using the Cox regression model. Kaplan-Meier analyses showed that LSCC patients with elevated transcriptional levels of COL4A1/COL4A2 and suppressed transcriptional levels of LINC00278 (high risk) were significantly associated with short OS outcomes, in contrast to low COL4A1/COL4A2 and high LINC00278 (low risk) (Figure 7A). Then the ESTIMATE algorithm was used to calculate the stromal scores, immune scores, and tumor purity of high- and low-risk groups. The low-risk group exhibited a higher immune score than the high-risk group. However, there were no significant differences in stromal scores and tumor purity between the two groups (Figure 7B). Consistent with the ESTIMATE algorithm, gene expression levels of human leukocyte antigens (HLA), programmed cell death ligand-1 (PDL-1), and cytotoxic T-lymphocyte-associated antigen 4 (CTLA4) were elevated in the low-risk group, relative to the high-risk group (Figures 7C, 7D). In addition, CIBERSORT was used to estimate differences in 22 kinds of infiltrating immune cells between low- and high-risk groups. Compared to the high-risk group, infiltration levels of M1 macrophages were significantly higher in the low-risk group (Figure 7E). These findings imply that in the tumor microenvironment, M1 macrophages might be involved in LINC00278-COL4A1/COL4A2 mediated prognostic outcomes in LSCC. A recent article in the *Cell* journal reported that, due to alternative splicing (AS), intron retention (IR) can efficiently activate antitumor immune responses *in vitro* [21]. Accordingly, we evaluated changes in AS events following LINC00278 overexpression in AMC-HN-8 cells. We found that LINC00278 overexpression significantly increased AS events, among which IR events increased from 2.00% to 2.72% (Figures 7F, 7G). Therefore, we postulated that LINC00278 and its downstream target genes upregulated AS events, thereby regulating the infiltration levels of tumor-infiltrating immune cells, which has an effect on antitumor immune responses. Thus, for the first time, we report on the function and mechanisms of LINC00278 in LSCC (Figure 8).



**Figure 6.** LINC00278 suppressed LSCC cells proliferation, migration, and invasion by downregulation PI3K/AKT/mTOR signaling pathway. **A, B)** LINC00278 inhibits the AKT/mTOR signaling pathway. **C–E)** SC79 reversed the inhibitory effects of LINC00278 on LSCC cells proliferation, migration, and invasion. **F)** Representative pictures of tumors formed in nude mice bearing AMC-HN-8 cells in LINC00278 and control group. **G)** Volumes of LINC00278 overexpression nude mice tumors were significantly smaller than those in control nude mice tumors. **H)** Volume and **I)** weight of xenograft tumors in nude mice. **J, K)** Immunohistochemical staining of COL4A1 and Ki-67 proteins in tumor tissues from LSCC xenografts in nude mice (magnification,  $\times 10$ ). (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).



## Discussion

In this study, we performed a computational analysis of high-throughput RNA-sequencing data obtained from TCGA database. Given that the predictive ability of a single or a limited number of genes was barely satisfactory, we identified six prognosis-associated signature genes (LINC00278, MYHAS, MNX1-AS1, LINC02086, LSAMP-AS1, and CASC20) to develop a prognostic model. To establish the role of the lncRNA-associated model in LSCC prognosis, overall survival analysis and ROC curves were performed. Kaplan-Meier survival curves showed that our lncRNA-associated model was excellent at discriminating patients among different risk groups with respect to death. ROC curves confirmed that the model had a satisfactory sensitivity and specificity. Interestingly, compared to the TNM stages, the predictive accuracy of the lncRNA-based model was significantly high. The TNM staging system remains the standard, while other methods are auxiliary when classifying infiltration ranges of tumors, and are widely used to predict cancer prognosis. Therefore, the lncRNA-based model might be useful as a supplement for tumor staging to better stratify patients for individualized treatments.

As the only male-determining chromosome, the Y chromosome serves significant role in the process of tumorigenesis [11–13]. However, the overall functions of the Y chromosome as well as those of its resident genes in LSCC have not been conclusively established. In this study, the crucial gene, LINC00278, on the Y chromosome was screened out in LSCC using TCGA database, and its expression levels were validated using multiple datasets from the GEO database. To verify the accuracy of bioinformatics analyses based on public databases, we performed qPCR experiments and cellular functional assay. In our series of 72 laryngeal cancer patients, LINC00278 mRNA expression levels in LSCC tissues were considerably low, when compared to normal tissues. We also observed that LINC00278 was not randomly expressed, but was significantly correlated with stage, lymphatic metastasis, and pathological differentiation. *In vitro*, overexpression of LINC00278 in LSCC cell lines inhibited cell proliferation, migration, and invasion. Meanwhile, expression levels of EMT-related marker proteins in each group confirmed the above results, and LINC00278 overexpression suppressed EMT. Bioinformatics analyses showed that LINC00278 was expressed at low levels in the HNSC tumor tissues, and comparable findings were

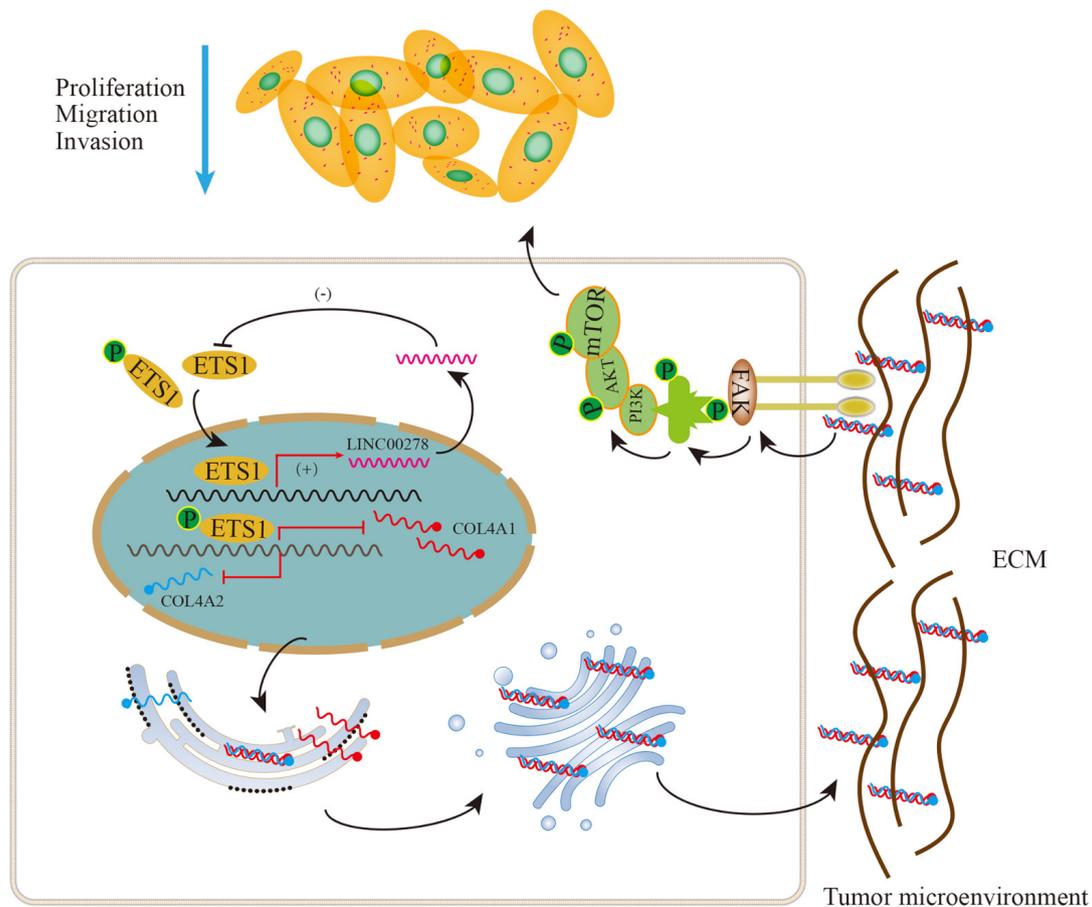


Figure 8. Diagram of the mechanism of action of LINC00278 against LSCC.

found in HSCC cells. These results suggest that LINC00278 is a potential biomarker for prognostic outcomes and therapeutic responses in LSCC and HSCC.

Then, we explored the upstream and downstream mechanisms of LINC00278 and found that at low levels, LINC00278 may be a potential diagnostic and prognostic molecular marker for LSCC. As an important downstream effector molecule for LINC00278, collagen IV (COL4A1/COL4A2) might be involved in the molecular pathogenesis of LSCC by regulating tumor cell proliferation, migration, invasion, and immune responses. Collagen proteins form the scaffold of the tumor microenvironment and are important for tumor infiltration, angiogenesis, and metastasis [22]. Collagen IV family contains 6 homologous  $\alpha$  chains,  $\alpha 1-6$ , which are encoded by COL4A1-6 genes, which are the most abundant basement constituents [23]. Liu *et al.* reported that COL4A1 and COL4A2 are significantly correlated with hepatocarcinogenesis [24]. Moreover, the COL4A1 and COL4A2 are potential markers for vascular endothelial cells' growth and migration [25]. Cao *et al.* showed that COL4A1 is significantly associated with bladder cancer recurrence [26]. Moreover, upregulated COL4A1 promotes the migration and proliferation of invasive ductal carcinomas [27].

Then, bioinformatics analyses were performed to investigate upstream transcription factors of LINC00278. We found that ETS1 was significantly downregulated in the tumor group. Currently, it has not been determined whether ETS1 can regulate LINC00278 transcription. ETS1 is a dichotomous transcription factor that is involved in oncogene or pro-apoptotic genes activation, cellular differentiation, tissue remodeling, angiogenesis, drug resistance, and tumorigenesis [28]. Elevated ETS1 levels are correlated with higher grading, poorer differentiation, and/or increased invasiveness in various tumor types, including LSCC [29, 30]. However, we found that ETS1 was downregulated, and it exhibited tumor suppressor roles. This finding is in accordance with a report that ectopic expressions of ETS1 significantly suppress tumor cell proliferation by enhancing the promoter activity of RYBP [31]. We postulate that it probably is tissue- and cell type-specific, and co-determines by its binding partners.

LINC00278 and ETS1 exhibit a negative feedback loop to maintain their respective low expression levels in LSCC. When ETS1 was overexpressed, expressions of target genes, COL4A1/COL4A2, were downregulated, and when LINC00278 was overexpressed, the expressions of COL4A1/COL4A2 were also downregulated. Surprisingly, expressions of COL4A1/COL4A2 were not suppressed, but they instead increased when we co-transfected the ETS1 transcription factor and the LINC00278 expression vector. These unconventional expression patterns of COL4A1/COL4A2 indirectly illustrate the intimate relationship between ETS1 and LINC00278. Studies should aim at elucidating the involved mechanisms. They may be correlated with low expression levels of ETS1 and LINC00278 status in LSCC cell lines.

To determine how LINC00278 regulates the physiological processes of LSCC cells, we evaluated the effects of overexpressed LINC00278 on the AKT/mTOR pathway. Western blotting analysis showed that p-AKT and p-mTOR were dysregulated by the upregulated LINC00278.

The tumor microenvironment (TME) is potentially associated with tumor progression and therapeutic outcomes [32]. The TME is comprised of immune cells, extracellular components, and cancer-associated fibroblasts (CAFs) [33]. In this study, LINC00278/COL4A1/COL4A2-dominated low-risk group showed a higher immune score, relative to the high-risk group (in any other two gene combinations, there were no significant differences in immune scores between the two clustered groups). In addition to the high immune score and high present antigen activity, higher expression levels of immune exhaustion genes in the low-risk group implied an immune activation status, which may be susceptible to the immune checkpoint blockade. Moreover, the presence of more M1 macrophage cells within the tumor of low-risk patients exhibited a favorable prognostic outcome in this study. M1 macrophages secrete pro-inflammatory cytokines, such as TNF- $\alpha$ , IL-12, and IL-13, as well as chemokines such as CCL-5, CXCL5, CXCL9, and CXCL10. Moreover, due to their cytotoxic activities, M1 macrophages are also involved in anti-tumor immunity [34]. We found that PD-L1 expression is negatively correlated with the number of M1 macrophages, which is in accordance with findings from other studies [35]. These results confirmed the accuracy of the CIBERSORT method in assessing immune cell compositions of complex tissues on the basis of gene expression profiles from bulk tumor samples [36]. Collectively, expression levels of LINC00278, COL4A1, and COL4A2 were significantly correlated with the tumor microenvironment, especially with tumor-infiltrating M1 macrophages. These findings will inform on rational drug designs or antitumor-immuno therapy.

We found that LINC00278 on the Y chromosome is activated by ETS1 to inhibit tumor cell growth. LINC00278 and ETS1 form a negative feedback loop to ultimately inhibit their respective expressions, which eventually maintains their respective low expression levels in LSCC. As an important downstream effector molecule of LINC00278, COL4A1/COL4A2 may be involved in the molecular pathogenesis of LSCC by regulating tumor cell proliferation, migration, invasion, and the TME. Moreover, treatment with LINC00278 downregulated p-AKT and p-mTOR protein levels, implying that re-expression of LINC00278 suppressed the AKT/mTOR signaling pathway in LSCC. Our findings suggest that LINC00278 suppression is important for LSCC progression. Therefore, LINC00278 may be a potential target for LSCC diagnosis and treatment.

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

Acknowledgments: We acknowledge the financial support from the National Natural Science Foundation of China (81972553) and the Project of Clinical Medical Talent Training and Basic Project Research Funded by 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] KHATRI P, SIROTA M, BUTTE AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol* 2012; 8: e1002375. <https://doi.org/10.1371/journal.pcbi.1002375>
- [5] KOLDE R, LAUR S, ADLER P, VILO J. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* 2012; 28: 573–580. <https://doi.org/10.1093/bioinformatics/btr709>
- [6] PENG Y. Non-coding RNAs in human cancer. *Semin Cancer Biol* 2021; 75: 1–2. <https://doi.org/10.1016/j.semcancer.2021.04.010>
- [7] CUI W, MENG W, ZHAO L, CAO H, CHI W et al. TGF-beta-induced long non-coding RNA MIR155HG promotes the progression and EMT of laryngeal squamous cell carcinoma by regulating the miR-155-5p/SOX10 axis. *Int J Oncol* 2019; 54: 2005–2018. <https://doi.org/10.3892/ijo.2019.4784>
- [8] LIU T, MENG W, CAO H, CHI W, ZHAO L et al. lncRNA RASSF8AS1 suppresses the progression of laryngeal squamous cell carcinoma via targeting the miR664b3p/TLE1 axis. *Oncol Rep* 2020; 44: 2031–2044. <https://doi.org/10.3892/or.2020.7771>
- [9] 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>
- [10] FEI M, ZHANG J, ZHOU J, XU Y, WANG J. Sex-related hormone receptor in laryngeal squamous cell carcinoma: correlation with androgen estrogen-a and prolactin receptor expression and influence of prognosis. *Acta Otolaryngol* 2018; 138: 66–72. <https://doi.org/10.1080/00016489.2017.1373851>
- [11] BANDAY AR, PAPPENBERG BW, PROKUNINA-OLSSON L. When the Smoke Clears m(6)A from a Y Chromosome-Linked lncRNA, Men Get an Increased Risk of Cancer. *Cancer Res* 2020; 80: 2718–2719. <https://doi.org/10.1158/0008-5472.CAN-20-0961>
- [12] BROWN DW, MACHIELA MJ. Why Y? Downregulation of Chromosome Y Genes Potentially Contributes to Elevated Cancer Risk. *J Natl Cancer Inst* 2020; 112: 871–872. <https://doi.org/10.1093/jnci/djz236>
- [13] PATEL R, KHALIFA AO, ISALI I, SHUKLA S. Prostate cancer susceptibility and growth linked to Y chromosome genes. *Front Biosci (Elite Ed)* 2018; 10: 423–436. <https://doi.org/10.2741/e830>
- [14] WU S, ZHANG L, DENG J, GUO B, LI F et al. A Novel Micropeptide Encoded by Y-Linked LINC00278 Links Cigarette Smoking and AR Signaling in Male Esophageal Squamous Cell Carcinoma. *Cancer Res* 2020; 80: 2790–2803. <https://doi.org/10.1158/0008-5472.CAN-19-3440>
- [15] HAN N, JIANG Y, GAI Y, LIU Q, YUAN L et al. C-Labeled Pictilisib (GDC-0941) as a Molecular Tracer Targeting Phosphatidylinositol 3-Kinase (PI3K) for Breast Cancer Imaging. *Contrast Media Mol Imaging* 2019; 2019: 1760184. <https://doi.org/10.1155/2019/1760184>
- [16] SHENG X, LIU W, LU Z, XU M, LI R et al. HTR7 promotes laryngeal cancer growth through PI3K/AKT pathway activation. *Ann Transl Med* 2021; 9: 840. <https://doi.org/10.21037/atm-21-1069>
- [17] ZHAO R, TIAN L, ZHAO B, SUN Y, CAO J et al. FADS1 promotes the progression of laryngeal squamous cell carcinoma through activating AKT/mTOR signaling. *Cell Death Dis* 2020; 11: 272. <https://doi.org/10.1038/s41419-020-2457-5>
- [18] 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>
- [19] SINH ND, ENDO K, MIYAZAWA K, SAITOH M. Ets1 and ESE1 reciprocally regulate expression of ZEB1/ZEB2, dependent on ERK1/2 activity, in breast cancer cells. *Cancer Sci* 2017; 108: 952–960. <https://doi.org/10.1111/cas.13214>
- [20] WANG T, JIN H, HU J, LI X, RUAN H et al. COL4A1 promotes the growth and metastasis of hepatocellular carcinoma cells by activating FAK-Src signaling. *J Exp Clin Cancer Res* 2020; 39: 148. <https://doi.org/10.1186/s13046-020-01650-7>
- [21] BOWLING EA, WANG JH, GONG F, WU W, NEILL NJ et al. Spliceosome-targeted therapies trigger an antiviral immune response in triple-negative breast cancer. *Cell* 2021; 184: 384–403.e21. <https://doi.org/10.1016/j.cell.2020.12.031>
- [22] FANG M, YUAN J, PENG C, LI Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol* 2014; 35: 2871–2882. <https://doi.org/10.1007/s13277-013-1511-7>
- [23] KUO DS, LABELLE-DUMAIS C, GOULD DB. COL4A1 and COL4A2 mutations and disease: insights into pathogenic mechanisms and potential therapeutic targets. *Hum Mol Genet* 2012; 21: R97–110. <https://doi.org/10.1093/hmg/dds346>

- [24] LIU Y, ZHANG J, CHEN Y, SOHEL H, KE X et al. The correlation and role analysis of COL4A1 and COL4A2 in hepatocarcinogenesis. *Aging (Albany NY)* 2020; 12: 204–223. <https://doi.org/10.18632/aging.102610>
- [25] LOSCERTALES M, NICOLAOU F, JEANNE M, LONGONI M, GOULD DB et al. Type IV collagen drives alveolar epithelial-endothelial association and the morphogenetic movements of septation. *BMC Biol* 2016; 14: 59. <https://doi.org/10.1186/s12915-016-0281-2>
- [26] CAO H, CHENG L, YU J, ZHANG Z, LUO Z et al. Identifying the mRNAs associated with Bladder cancer recurrence. *Cancer Biomark* 2020; 28: 429–437. <https://doi.org/10.3233/CBM-190617>
- [27] JIN R, SHEN J, ZHANG T, LIU Q, LIAO C et al. The highly expressed COL4A1 genes contributes to the proliferation and migration of the invasive ductal carcinomas. *Oncotarget* 2017; 8: 58172–58183. <https://doi.org/http://dx.doi.org/10.18632/oncotarget.17345>
- [28] TESTONI M, CHUNG EY, PRIEBE V, BERTONI F. The transcription factor ETS1 in lymphomas: friend or foe? *Leuk Lymphoma* 2015; 56: 1975–1980. <https://doi.org/10.3109/10428194.2014.981670>
- [29] CALLI AO, SARI A, CAKALAGAOGLU F, ALTINBOGA AA, ONCEL S. ETS-1 proto-oncogene as a key newcomer molecule to predict invasiveness in laryngeal carcinoma. *Pathol Res Pract* 2011; 207: 628–633. <https://doi.org/10.1016/j.prp.2011.07.010>
- [30] DITTMER J. The role of the transcription factor Ets1 in carcinoma. *Semin Cancer Biol* 2015; 35: 20–38. <https://doi.org/10.1016/j.semcancer.2015.09.010>
- [31] ZHAO W, ZHANG S, WANG X, MA X, HUANG B et al. ETS1 targets RYBP transcription to inhibit tumor cell proliferation. *Biochem Biophys Res Commun* 2019; 509: 810–816. <https://doi.org/10.1016/j.bbrc.2019.01.006>
- [32] 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>
- [33] 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>
- [34] KARPATHIU G, DUMOLLARD JM, PEOC'H M. Laryngeal Tumor Microenvironment. *Adv Exp Med Biol* 2020; 1296: 79–101. [https://doi.org/10.1007/978-3-030-59038-3\\_5](https://doi.org/10.1007/978-3-030-59038-3_5)
- [35] YU D, CHENG J, XUE K, ZHAO X, WEN L et al. Expression of Programmed Death-Ligand 1 in Laryngeal Carcinoma and its Effects on Immune Cell Subgroup Infiltration. *Pathol Oncol Res* 2019; 25: 1437–1443. <https://doi.org/10.1007/s12253-018-0501-x>
- [36] CHEN B, KHODADOUST MS, LIU CL, NEWMAN AM, ALIZADEH AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol* 2018; 1711: 243–259. [https://doi.org/10.1007/978-1-4939-7493-1\\_12](https://doi.org/10.1007/978-1-4939-7493-1_12)
- [37] LIAN M, FANG J, HAN D, MA H, FENG L et al. Microarray gene expression analysis of tumorigenesis and regional lymph node metastasis in laryngeal squamous cell carcinoma. *PLoS One* 2013; 8: e84854. <https://doi.org/10.1371/journal.pone.0084854>
- [38] BARROS E LIMA BUENO R, RAMÃO A, PINHEIRO DG, ALVES CP, KANNEN V et al. HOX genes: potential candidates for the progression of laryngeal squamous cell carcinoma. *Tumor Biol* 2016; 37: 15087–15096. <https://doi.org/10.1007/s13277-016-5356-8>
- [39] SHEN Z, YUAN J, TONG Q, HAO W, DENG H et al. Long non-coding RNA AC023794.4-201 exerts a tumor-suppressive function in laryngeal squamous cell cancer and may serve as a potential prognostic biomarker. *Oncol Lett* 2020; 20: 774–784. <https://doi.org/10.3892/ol.2020.11595>
- [40] FENG L, WANG R, LIAN M, MA H, HE N et al. Integrated Analysis of Long Noncoding RNA and mRNA Expression Profile in Advanced Laryngeal Squamous Cell Carcinoma. *PLoS One* 2016; 11: e0169232. <https://doi.org/10.1371/journal.pone.0169232>
- [41] ZHAO R, LI FQ, TIAN LL, SHANG DS, GUO Y et al. Comprehensive analysis of the whole coding and non-coding RNA transcriptome expression profiles and construction of the circRNA-lncRNA co-regulated ceRNA network in laryngeal squamous cell carcinoma. *Funct Integr Genomics* 2019; 19: 109–121. <https://doi.org/10.1007/s10142-018-0631-y>
- [42] NICOLAU-NETO P, DE SOUZA-SANTOS PT, SEVERO RAMUNDO M, VALVERDE P, MARTINS I et al. Transcriptome Analysis Identifies ALCAM Overexpression as a Prognosis Biomarker in Laryngeal Squamous Cell Carcinoma. *Cancers (Basel)* 2020; 12: 470. <https://doi.org/10.3390/cancers12020470>

[https://doi.org/10.4149/neo\\_2022\\_220310N263](https://doi.org/10.4149/neo_2022_220310N263)

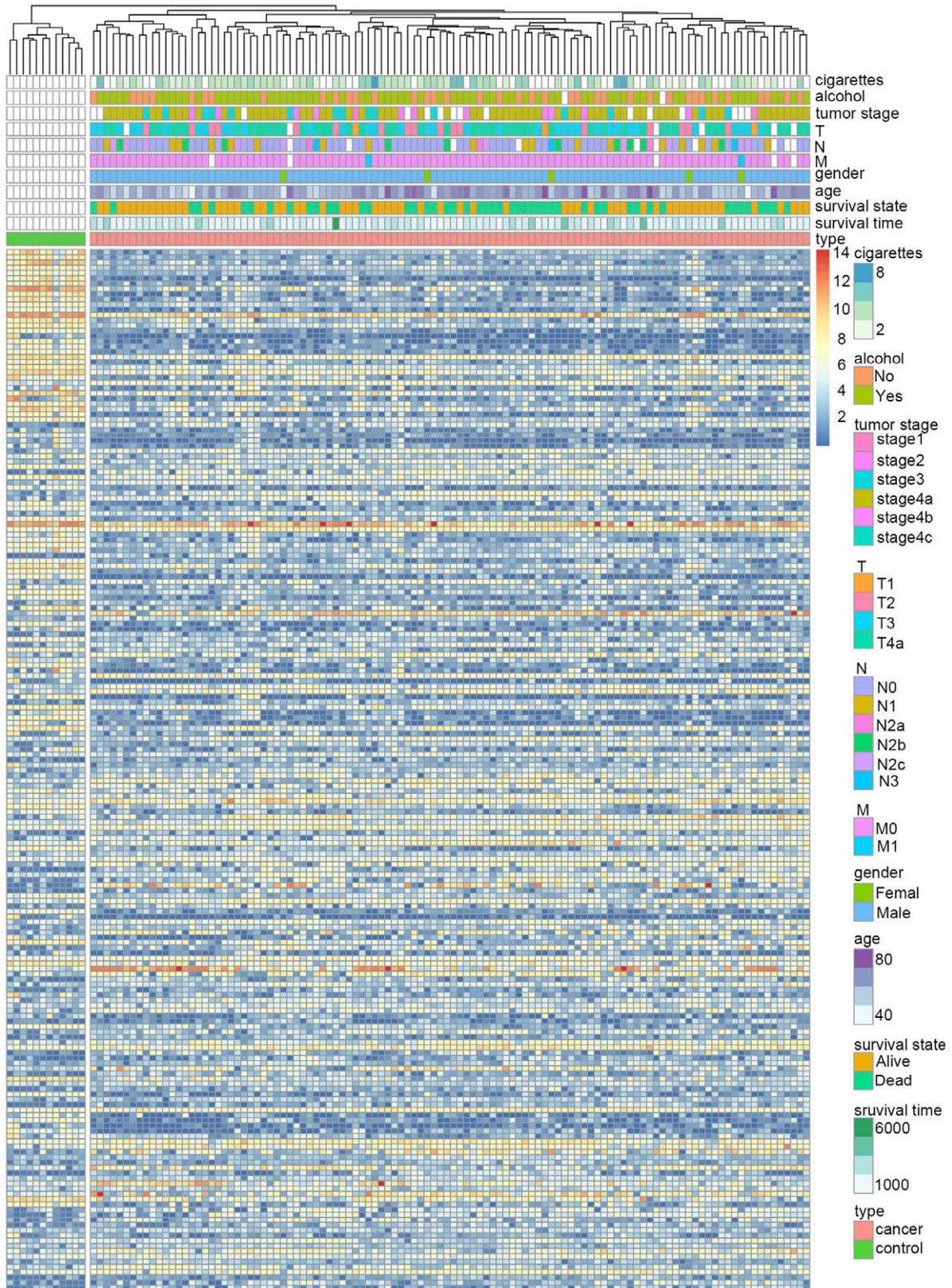
## The ETS1-LINC00278 negative feedback loop plays a role in COL4A1/COL4A2 regulation in laryngeal squamous cell carcinoma

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

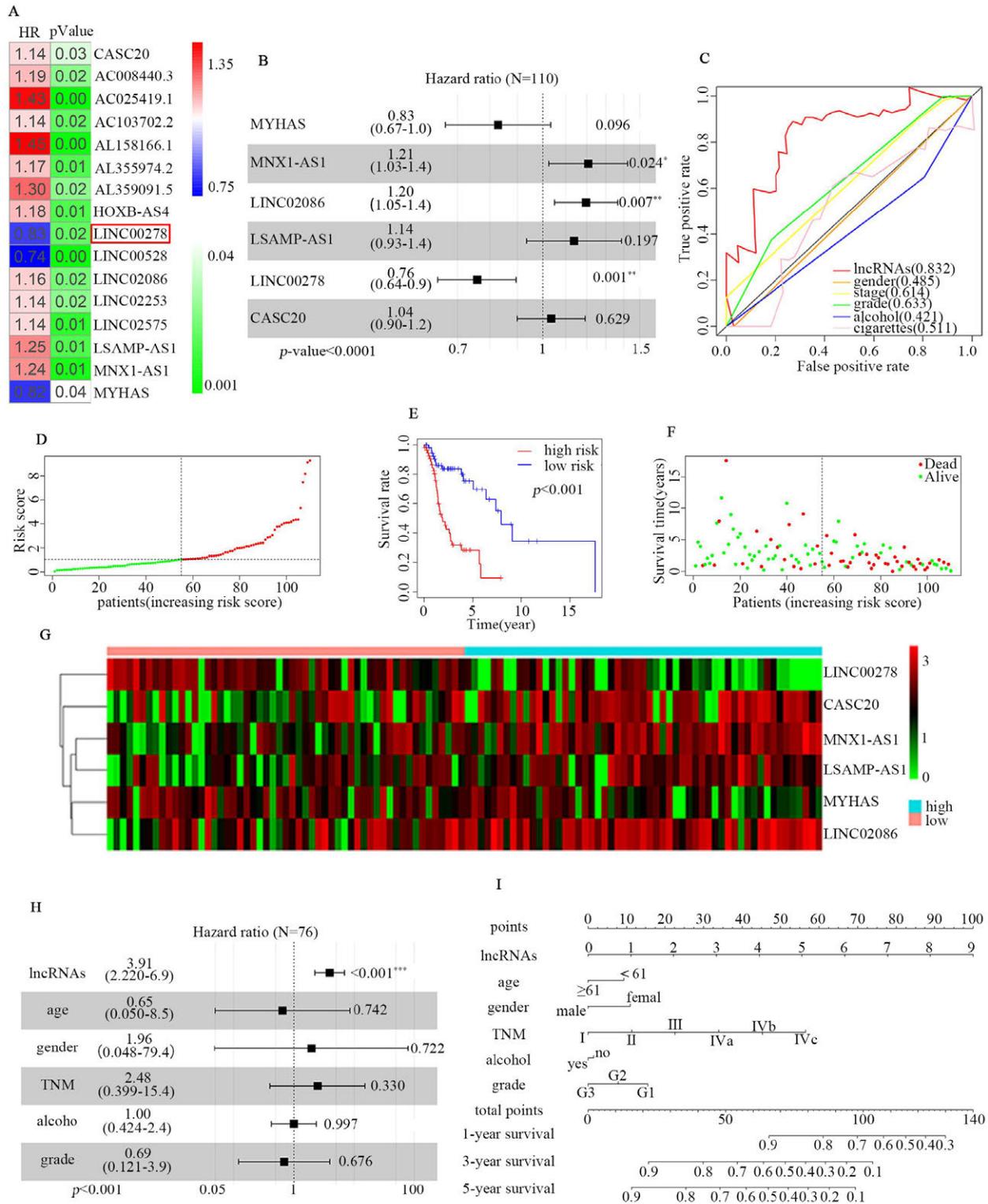
### Supplementary Information

**Supplementary Table S1. Primer sequences used in this study.**

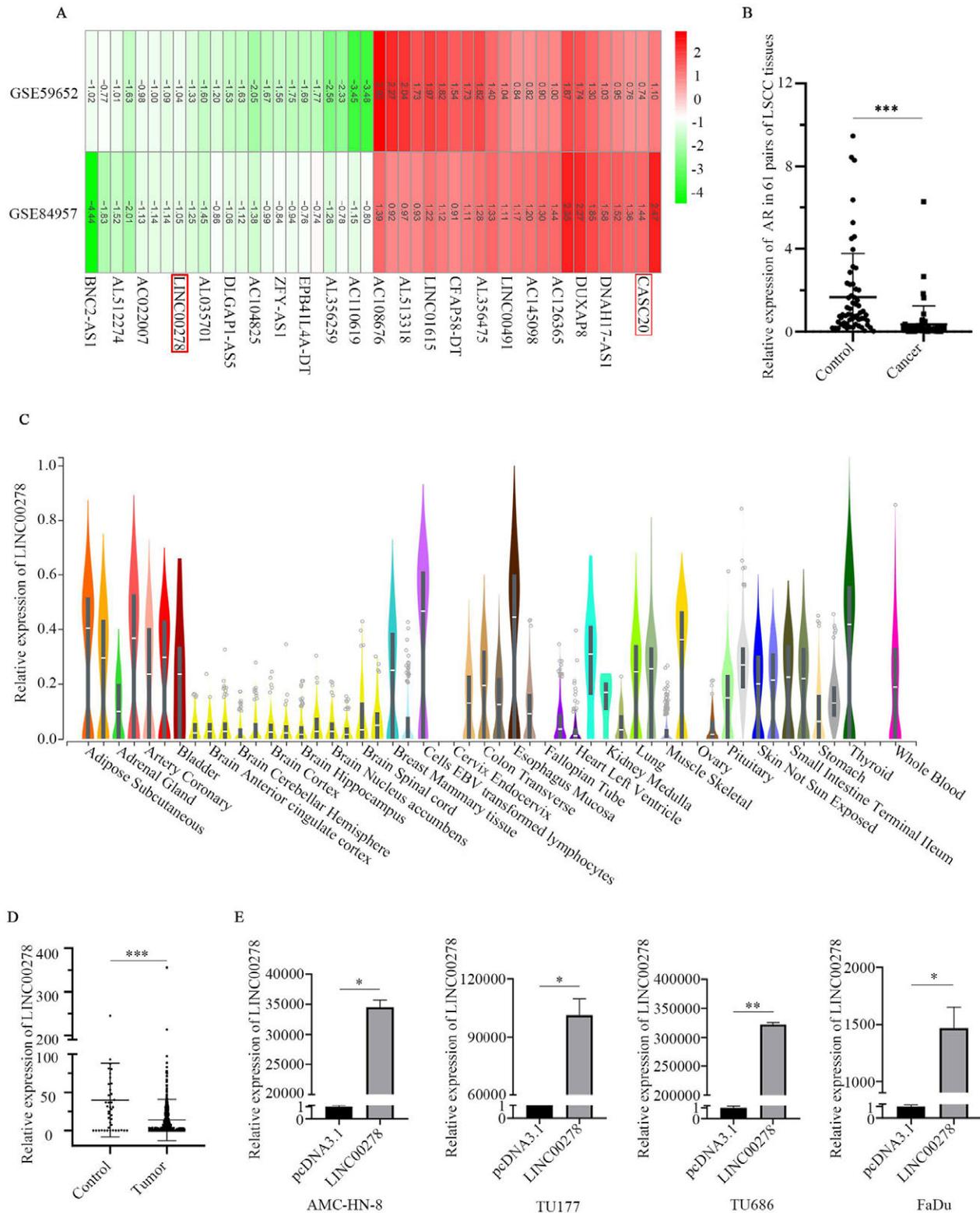
Gene	Sequence (5' - 3')	Gene	Sequence (5' - 3')
LINC00278	F: ACAGGACCTCGCAAGAAGAC R: TTGCAGACCTTGAGTGGTTG	U6	F: GCTTCGGCAGCACATATACTAAAAT R: CGCTTCACGAATTTGCGTGTTCAT
AR	F: GGCAGACAGAGGAAAAAGGG R: CCTTGCTTCCTCCGAGTCTT'	<b>ChIP assay</b>	
COL4A1	F: GGCAGATTCGGACCACTAGG R: GCGTCTGTGGCAATACTAGC	Region 5	F: CAGAGTTTCTCCATGTTGGTG R: TGGCTTAAACCTGTAATCCC
COL4A2	F: GGACAGACGAGACAACAGCA R: GAGCTGGCATAACATTGGCG	Region 4	F: CTTTTAAGAGCAAATAAGTGTT R: GCTTTCAATATTTGGGGTA
ETS1	F: AGATGTAGCGATGTAAGTGTCG R: TGTGCCAGCATCAGCTACTA	Region 3	F: TTAGCCTCACTTAGGGTTT R: ACTTGT'TTTCTTATATTTCC
GAPDH	F: AGGTGAAGTTCGGAGTCAACG R: AGGGGTCATTGATGGCAACA	Region 2	F: AAAGCTTGTGTTTCTTGCTC R: CTTGCAGGTAATGTTATCACAG
E-cadherin	F: CGAGAGCTACACGTTACGG R: GGCCTTTTGACTGTAATCACACC	Region 1-N1	F: TCTCCTTTGCCTGTGGGTCT R: GGTCTGTAGCACACTGTTCTT
N-cadherin	F: CAACTTGCCAGAAAACCTCCAGG R: ATGAAACCGGGCTATCTGCTC	Region 1-N2	F: ACAACTTCAACCCGAGGGGA R: GGTCTGTAGCACACTGTTCTT
Vimentin	F: CGCCTGCAGGATGAGATTGAG R: TCAGGGAGGAAAAGTTGGAAGA	<b>Luciferase assay</b>	
Snail	F: ACGAGGTGTGACTAACTAT R: CGACAAGTGACAGCCATT	Region 5	F: GAAGATCTCCACGACCAGCTAATT R: CCCAAGCTTCCCAACACTATTTGCTC
Twist	F: ACCATCCTCACACCTCTG R: GATTGGCACGACCTCTTG	Region 4	F: GAAGATCTCTCAACATCATTGTCC R: CCCAAGCTTAGAACTTTATCTTGTA AAAACA
Zeb1	F: TCATCGCTACTCCTACTGT R: TCACTGTCTTATCCTCTTTC	Region 3	F: CGGGGTACCGACAGTATGTAGATGATGG R: GAAGATCTCCTGGTCAGGGAAGGA
$\beta$ -catenin	F: ATGGCTTGGAATGAGAC R: AACTGGATAGTCAGCACC	Region 2	F: CGGGGTACCACTCTTCTACTGTGAAG R: GAAGATCTAAATGTGGCTGGAATC
		Region 1	F: GAAGATCTGGTGAGCCAGCCAGG R: CCCAAGCTTCGAGGTCTGTAGCACACT



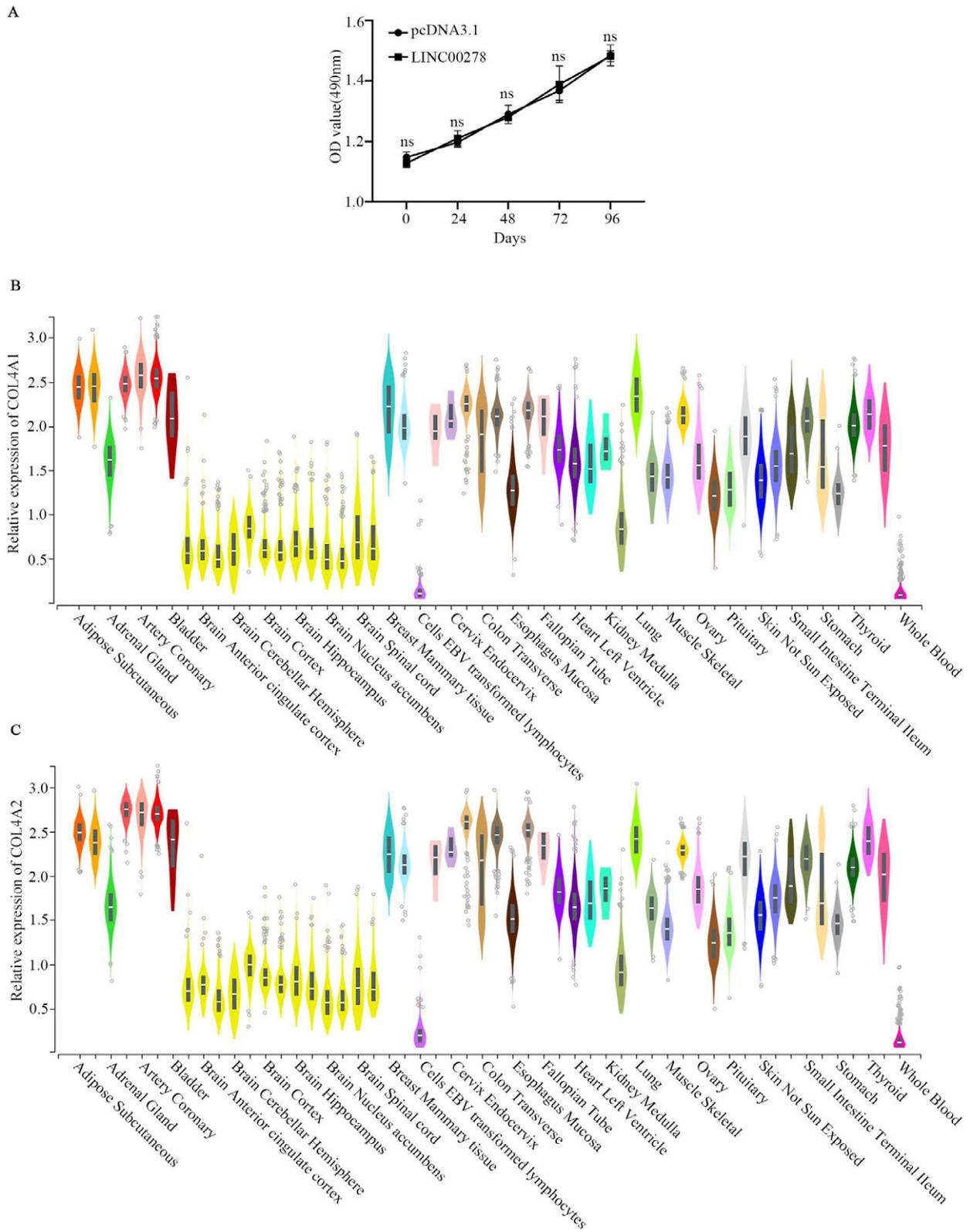
Supplementary Figure S1. A heatmap was shown for the relationship between clinical factors and some differential lncRNAs expression in the TCGA dataset.



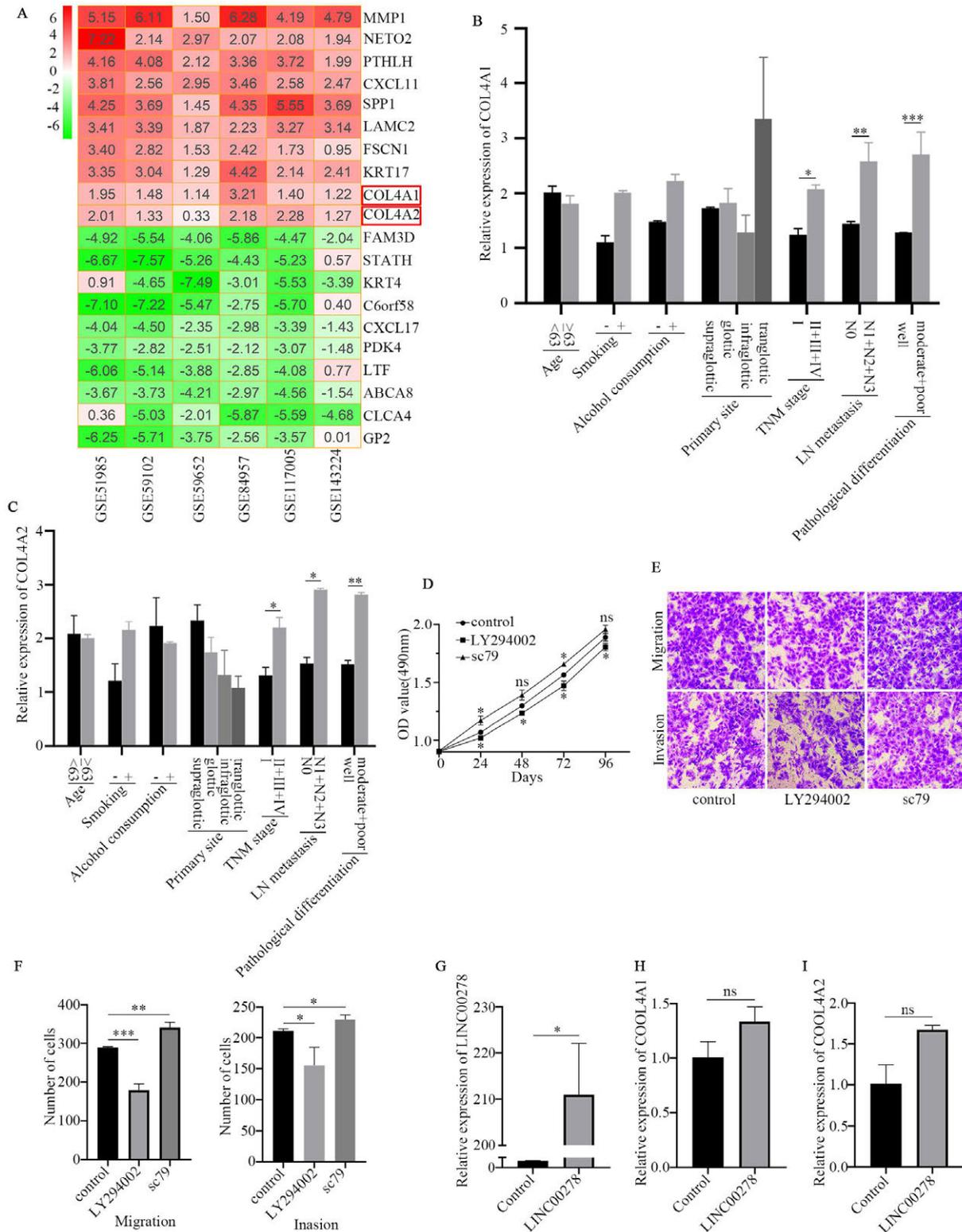
Supplementary Figure S2. The multivariate Cox regression analysis. A) The Univariate Cox regression of the 16 differential lncRNAs in LSCC patients. B) Forest plot of the multivariate Cox regression model showing six prognosis-related signature genes. C) The ROC curves and AUC of lncRNAs (red curve), alcohol (blue curve), cigarettes (pink curve), gender (orange curve), grade (green curve) and TNM stage (yellow curve). The risk score distribution (D), Kaplan-Meier curve (E), patient survival status (F) and heatmap of the six lncRNAs of the LSCC patients in high- and low-risk groups (G). H) Forest plot presenting the multivariate risk factors of LSCC patients. I) Nomogram for the 1, 3, 5-year OS prediction based on lncRNAs, age, gender, grade, alcohol, and TNM stage.



Supplementary Figure S3. LINC00278 was selected as marker gene of LSCC. A) Heatmap showed some differential lncRNAs in different studies by RRA methods from 2 different datasets. B) Relative expression of AR in 61 paired LSCC tissues. C) GTEx data show that LINC00278 were expressed divergently in different tissues. D) Relative expressions of LINC00278 in HNSC tumor tissues and non-tumor tissues from TCGA database. E) The expression of LINC00278 significantly upregulated after the cells were transfected with pcDNA3.1-LINC00278.



Supplementary Figure S4. The cell viability was assessed by MTS assay after transfected with pcDNA3.1-LINC00278 (A), COL4A1 (B) and COL4A2 (C) expression in 54 tissues from GTEx RNA-seq.



**Supplementary Figure S5. The correlation between COL4A1/COL4A2 expression and clinical features of LSCC. A) Heatmap showed some differential mRNAs in different studies by RRA methods from 6 different datasets. B) The relation between the expression of COL4A1 and clinical parameters based on qPCR results. C) The relation between the expression of COL4A2 and clinical parameters based on qPCR results. D-F) The cell viability, migration and invasion capability after transfected with SC79 and LY294002. G-I) LINC00278, COL4A1 and COL4A2 mRNA expression in nude mice tissues.**