Check for updates

Identification of Novel Prognostic Markers Associated With Laryngeal Squamous Cell Carcinoma Using Comprehensive Analysis

Chao Huang[1], Jun He[1], Yi Dong[2,3,4], Li Huang[1], Yichao Chen[1], Anquan Peng[1] and Hao Huang[2,3,4,5]*

[1] Department of Otolaryngology-Head and Neck Surgery, Second Xiangya Hospital Central South University, Changsha, China, [2] Department of Nephrology, Xiangya Hospital Central South University, Changsha, China, [3] Department of Cell Biology, School of Life Sciences, Central South University, Changsha, China, [4] Hunan Key Laboratory of Organ Fibrosis, Central South University, Changsha, China, [5] National Clinical Research Center for Geriatric Disorders, Xiangya Hospital Central South University, Changsha, China

Background: Laryngeal squamous cell carcinoma (LSCC) is a leading malignant cancer of the head and neck. Patients with LSCC, in which the cancer has infiltrated and metastasized, have a poor prognosis. Therefore, there is an urgent need to identify more potential targets for drugs and biomarkers for early diagnosis.

Methods: RNA sequence data from LSCC and patients' clinical traits were obtained from the Gene Expression Omnibus (GEO) (GSE142083) and The Cancer Genome Atlas (TCGA) database. Differentially expressed gene (DEG) analysis and weighted gene co-expression network analysis (WGCNA) were performed to identify hub genes. Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, prognostic value analysis, receiver operating characteristic (ROC) curve analysis, gene mutation analysis, tumor-infiltrating immune cell abundance profile estimation, gene set variation analysis (GSVA), and gene set enrichment analysis (GSEA) were performed. Single-gene RNA sequencing data were obtained from the GSE150321 dataset. Cell proliferation and viability were confirmed by the CCK-8 assay and real-time PCR.

Results: A total of 701 DEGs, including 329 upregulated and 372 downregulated genes, were screened in the GSE142083 dataset. Using WGCNA, three modules were identified to be closely related to LSCC. After intersecting the DEGs and performing univariate and multivariate Cox analyses, a novel prognostic model based on three genes (*SLC35C1*, *HOXB7*, and *TEDC2*) for LSCC was established. Interfering *TEDC2* expression inhibited tumor cell proliferation and migration.

Conclusions: Our results show that *SLC35C1*, *HOXB7*, and *TEDC2* have the potential to become new therapeutic targets and prognostic biomarkers for LSCC.

Keywords: laryngeal squamous cell carcinoma, WGCNA, differentially expressed genes, prognostic gene, gene modules, single-cell analysis, TEDC2

INTRODUCTION

Laryngeal squamous cell carcinoma (LSCC) is a common type of head and neck squamous cell carcinoma (HNSCC), accounting for approximately 20% of all cancer patients and 2.4% of new malignancies worldwide each year (1–3). Patients with LSCC, in which the cancer has infiltrated and metastasized, have a poor prognosis with a 5-year survival rate of approximately 60% (4). Despite recent advances in comprehensive surgical, chemotherapeutic, and radiotherapeutic treatment strategies, the global mortality rate associated with LSCC has not decreased (5). There are almost no typical symptoms in the early stages of LSCC, and mild symptoms are almost always ignored by patients (6). Therefore, the identification of abnormally expressed genes in LSCC and early intervention are important strategies for prolonging the survival time of patients with LSCC.

In recent years, with the development of gene chip technologies, methods of cancer diagnosis have become increasingly more efficient and considerably simpler, allowing researchers to improve cancer diagnoses in a relatively short period of time, aiding the identification of improved treatment measures (7–9). Additionally, owing to these advantages, research into specific molecular markers of cancer has become a topic of increased interest in recent years (10–13). Weighted gene co-expression network analysis (WGCNA) is a powerful approach for identifying gene co-expression modules, exploring the correlation between the modules and phenotypes and discovering hub genes that regulate critical biological processes (14–16). However, only few reports in the literature describing the hub genes and biomarkers in LSCC patients were identified by WGCNA (17–20).
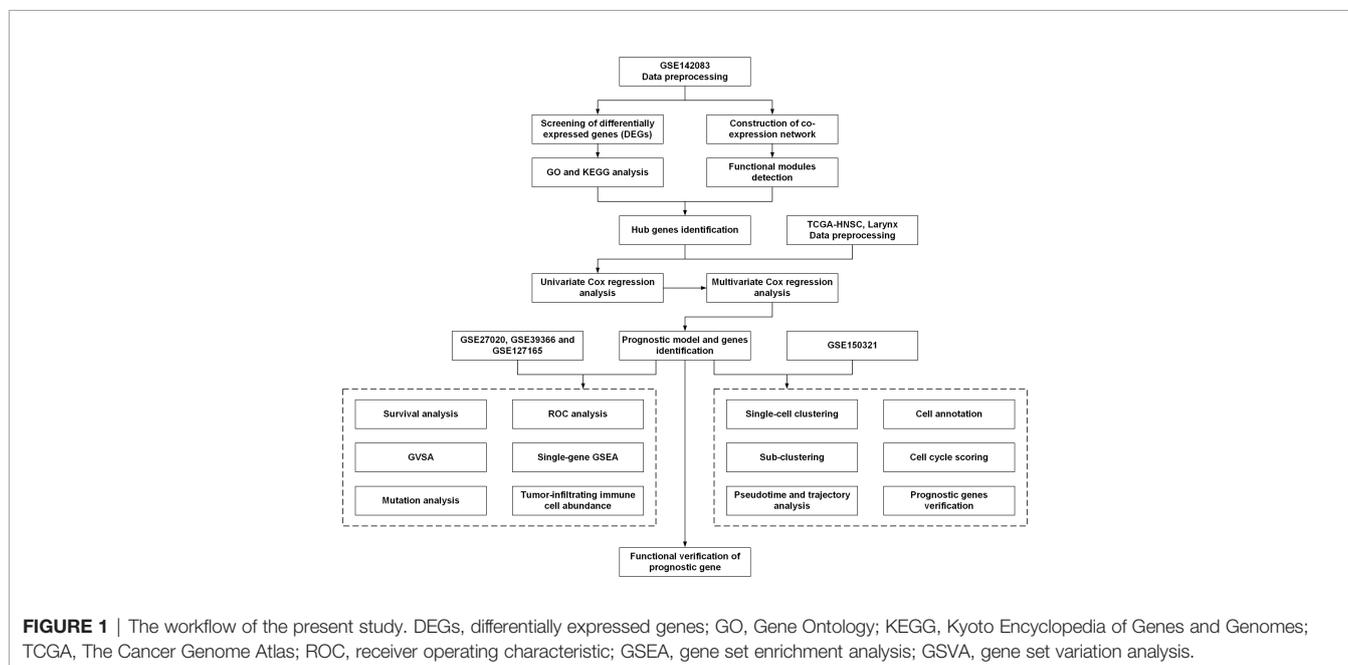
In this study, a transcriptome dataset of LSCC and adjacent normal tissues from patients was obtained from the Gene Expression Omnibus (GEO) and used to identify the differences in expression profiles and the differentially expressed genes (DEGs) between the LSCC and control groups using WGCNA. A prognostic model of LSCC was established using The Cancer Genome Atlas Head and Neck Squamous Cell Carcinoma (TCGA-HNSC) dataset. Single-cell RNA-sequencing (scRNA-Seq) data from GEO were also used to verify the prognostic model and the expression profile of prognostic genes in different cell types. The present study aimed to improve our understanding of the pathogenesis of LSCC by identifying the underlying specific molecular mechanisms and to provide insights into novel therapeutic drug targets for the treatment of LSCC.

METHODS

Data Collection and Single-Cell RNA-Sequencing Data Processing

The workflow used in this study is shown in **Figure 1**. The search words "laryngeal carcinoma" OR "LSCC" OR "laryngeal cancer" AND "human" AND "Expression profiling by array" were applied for dataset retrieval. Datasets that included non-tumor samples and the total sample number of over 100 could be selected. Therefore, GSE142083 was selected to perform the following bioinformatics analysis. The gene expression matrix of 106 LSCC samples (53 LSCC and paired adjacent normal mucosa tissues) was downloaded from the GSE142083 dataset obtained from the GEO database (21). The GEO expression matrix was annotated with gene symbols using information from the GPL20301 Illumina HiSeq 4000 (Homo sapiens) platform file (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL20301) as transcripts per million (TPM), and this

FIGURE 1 | The workflow of the present study. DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; TCGA, The Cancer Genome Atlas; ROC, receiver operating characteristic; GSEA, gene set enrichment analysis; GSVA, gene set variation analysis.

information was $\log_2$ transformed in R (version 4.0.4) and RStudio (version 1.2.5033) when necessary. Principal component analysis (PCA) using the "ggfortify" package was performed, and the outliers were excluded (GSM4219698 and GSM4219730) (**Supplementary Figure 1**). The remaining data from the 52 LSCC and 52 paired normal tissue samples were included in the subsequent analysis. GSE27020 (22), GSE39366 (23), and GSE127165 (24) were also downloaded from GEO for validation.

The gene expression RNA-Seq data and the corresponding clinical data from the TCGA-HNSC dataset were downloaded from the University of California, Santa Cruz (UCSC) Xena (http://xena.ucsc.edu/).

scRNA-Seq data from the GSE150321 dataset were obtained from the GEO database (25). The "Seurat" R package (version 4.0.2) was used to process the data (26). The scRNA-Seq data of GSM4546857 and GSM4546858 from two LSCC patients were included and processed as described previously. Cells with less than 200 genes detected genes were excluded. Data were normalized by the method "LogNormalize" with a scale factor equal to 10,000. After determining the statistically significant principal components, cell clusters were annotated using the information from previous literatures. Next, we implemented the t-distributed stochastic neighbor embedding (t-SNE) algorithm to explore and visualize the cluster classification across cell samples. The trajectory analysis and pseudotime analysis of tumor cells were performed using the "monocle" R package (version 2.18.0) (27).

DEG Identification
After removing the outliers, the expression data were standardized using R software. The "limma" (version 3.46.0) software package (http://www.bioconductor.org/packages/release/bioc/html/limma.html) (28) was used to perform the DEG analysis between the LSCC and adjacent normal control tissues from the GSE142083 dataset. Genes with an expression-adjusted p value < 0.05 and $\log_2$ (fold change) > 1.5 were regarded as significant DEGs. The volcano and heatmap plots were generated using "ComplexHeatmap" and "ggplot2" package. Extracellular matrix-associated genes were annotated using the Matrisome Project (http://matrisomeproject.mit.edu) (29).

GO Enrichment and KEGG Analysis of DEGs
GO and KEGG enrichment analyses of DEGs and hub modules were performed using the "clusterProfiler" (version 3.18.1) R package (http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) (30). The three main processes in GO analysis were as follows: biological process (BP), molecular function (MF), and cellular component (CC). Statistical significance was set at adjusted-p < 0.05. Plots were generated using the "GOplot" (version 1.0.2) package.

Weighted Gene Co-Expression Network Construction
The "WGCNA" package (version 1.70-3) in R (14) was used to construct co-expression networks of LSCC and adjacent normal

tissue samples. Genes with a mean fragments per kilobase per million mapped reads (FPKM) > 0.5 were selected. The adjacency matrices that stored information of the entire co-expression network were created based on Pearson's correlation matrices. A topological overlap matrix (TOM) was created from the adjacency matrix to estimate the connectivity of the network. Average linkage hierarchical clustering was used to construct a clustering dendrogram of the TOM matrix with a minimum module size of 30. Finally, similar gene modules were merged with a threshold of 0.25.

Identification and Verification of Prognostic Gene Signatures
The hub genes were identified as intersecting between the midnight-blue, blue, and green-yellow modules from the WGCNA and DEGs in the GSE142083 dataset. These hub genes were validated for their prognostic signature.

First, we identified the LSCC patient cohort through the clinical data from the TCGA-HNSC dataset (TCGA-HNSC-Larynx). Univariate Cox regression analysis of these hub genes was performed to screen the significant genes associated with overall survival (OS) in the training cohort. Genes with p < 0.1 were included in the subsequent analysis. Finally, multivariate Cox regression model analysis was performed to establish a prognostic model. The coefficients of multivariate Cox regression analysis were used to calculate the risk score (RS) of each sample using the following formula:

$$RS = \sum_{i=1}^{N}(coef_i \times expr_i)$$

According to the RS, samples were divided into a high-risk group and a low-risk group with a cutoff value of 50% in the TCGA-HNSC-Larynx cohort. Receiver operating characteristic (ROC) curve analysis and Kaplan–Meier analysis were conducted between the high-risk and low-risk groups. For ROC analysis, "survivalROC" package (version 1.0.3) was performed to draw 5-year OS ROC curves with the survival status (dead or alive) as outcome variable. Verification was also performed using the GSE27020, GSE39366, and GSE127165 datasets.

Mutation Analysis
The R package "maftools" (version 2.6.05) was used to analyze the mutation data of TCGA-HNSC-Larynx dataset. The tumor mutation burden (TMB) score for each sample from the high-risk and the low-risk groups was calculated.

Tumor-Infiltrating Immune Cell Abundance Analysis
The CIBERSORT (https://cibersort.stanford.edu/) algorithm was used to assess the proportions of 22 types of infiltrating immune cells based on the TCGA-HNSC-Larynx dataset and GSE142083 data following a previously reported procedure (31). A bar plot was drawn to show the differences in the composition of 22 kinds of tumor-infiltrating immune cells between the high-risk group and low-risk group.

Function Analysis of the Prognostic Genes by Single-Gene Gene Set Enrichment Analysis

KEGG analysis was conducted on the high- and low-risk groups *via* Gene Set Variation Analysis (GSVA). The reference information was downloaded from the Molecular Signature Database v7.4 (MSigDB v7.4, http://software.broadinstitute.org/gsea/msigdb/index.jsp) (32). Enriched pathways with *p*-value < 0.05 were considered to be statistically significant.

We conducted single-gene gene set enrichment analysis (GSEA) (33) of the prognostic genes using data from TCGA. All samples were divided into high-expression and low-expression groups based on the median of the prognostic genes, and GSEA was conducted to explore the enrichment of GO biological process (BP) and KEGG pathways in different groups. Statistical significance was set at *p* < 0.05.

Cell Culture and Transfection

The laryngeal cancer cell line AMC-HN-8 and the dysplastic oral keratinocyte cell line DOK were acquired from the Department of Otolaryngology-Head and Neck Surgery, Xiangya Hospital Central South University. AMC-HN-8 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) with high glucose (HyClone, Logan, UT, USA) and 10% fetal bovine serum (FBS) (Biological Industries, CT, USA). DOK cells were cultured in RPMI-1640 medium (HyClone, Logan, Utah, USA) and 10% FBS (Biological Industries, CT, USA). Cells were maintained at 37°C in a humidified incubator with 5% CO_2.

The *TEDC2* siRNA (siTEDC2 #1 and siTEDC2 #2) was produced by RiboBio Inc. (Guangzhou, China). SiRNA was transfected into cells using Lipofectamine 3000 Reagent (Invitrogen, USA) with Opti-MEM medium (Gibco, USA).

RNA Isolation and RT-PCR

The total RNA was extracted using TRIzol reagent (Solarbio, Beijing, China) and subjected to reverse transcription with random primers using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, USA). The expression level of targeted genes was measured with Maxima SYBR Green/ROX qPCR Mix (Thermo Scientific, USA) using real-time PCR system (Roche, Basel, Switzerland). The relative RNA expression levels were calculated using the $2^{(-\triangle\triangle CT)}$ method. The *18 s* rRNA was used as internal controls. The sequences of primers will be provided upon request.

CCK-8 Assay

For CCK-8 assay, 1,000 cells per well were seeded into 96-well plates. For each well, 10 μl CCK-8 (Solarbio, Beijing, China) was added into the medium. Then, incubated at 37°C for 2 h, the OD value at 450 nm was measured at different time.

Transwell Assay

For cell migration assay, 2.5×10^5 cells per well that were suspended in DMEM were seeded into a 24-well 8.0-μm transwell top chamber (Jet Biofil, Guangzhou, China). DMEM medium supplemented with 12% FBS was added to the bottom chambers. After incubating at 37°C for 12 h, cells at top chambers were fixed with 4% paraformaldehyde for 30 min, followed by permeabilization with methyl alcohol for 20 min. Then, cells were stained with 0.1% crystal violet (Solarbio, Beijing, China) for 15 min. Cells that did not migrate through the pores were removed by a cotton swab. Cells on the bottom of the chamber were counted using an inverted phase-contrast microscope at low magnifications (×5) (at least three randomly selected fields were quantified).

Statistical Analysis

Statistical analysis was conducted using R (version 4.0.4) or GraphPad Prism (version 8.0). Student-*t* test was used to compare two normal-distribution groups. NS: not statistically significant; *p < 0.05; **p < 0.01; ***p < 0.001.

RESULTS

DEG Screening

We analyzed the DEGs in the GSE142083 dataset using the limma package with a threshold of |log₂(fold-change)| > 1.2 and an adjusted-*p* < 0.05. A total of 701 DEGs, including 329 upregulated genes and 372 downregulated genes, were screened between the LSCC and adjacent normal samples. The heatmap and volcano plots show the expression patterns of these DEGs (**Figures 2A, B**).

The clusterProfiler package was then used to determine the role of DEGs in the pathogenesis of LSCC using a cutoff criterion of adjusted-*p* < 0.05. For the biological process group, skin development, epidermis development, cornification, keratinization, and keratinocyte differentiation were significantly enriched. The genes in the cellular component group were significantly enriched in cornified envelope, collagen-containing extracellular matrix (ECM), and ECM component. In the molecular function group, the DEGs were primarily enriched in ECM structural constituent and peptidase regulator activity. KEGG analysis revealed that the IL-17 signaling pathway, salivary secretion, ECM–receptor interaction, cell cycle, and protein digestion and absorption were significantly enriched (**Figure 2C**). We also noticed that ECM-associated genes were enriched in the DEGs, and most of them were highly expressed in the LSCC samples (**Supplementary Figure 2**). Moreover, hsa04970:Salivary secretion was one of the few terms that had decreasing Z-scores. This may indicate a significant decrease in salivary secretion in patients with LSCC.

Weighted Co-Expression Network Construction and Analysis

To detect the functional module in LSCC, we applied the WGCNA package based on the GSE142083 dataset to establish the gene co-expression networks. To ensure that the network was scale-free, an empirical analysis was conducted to choose an optimal parameter β. Both the scale-free topology model fit index (R^2) and mean connectivity reached a steady state when β = 9 (**Figures 3A, B**).
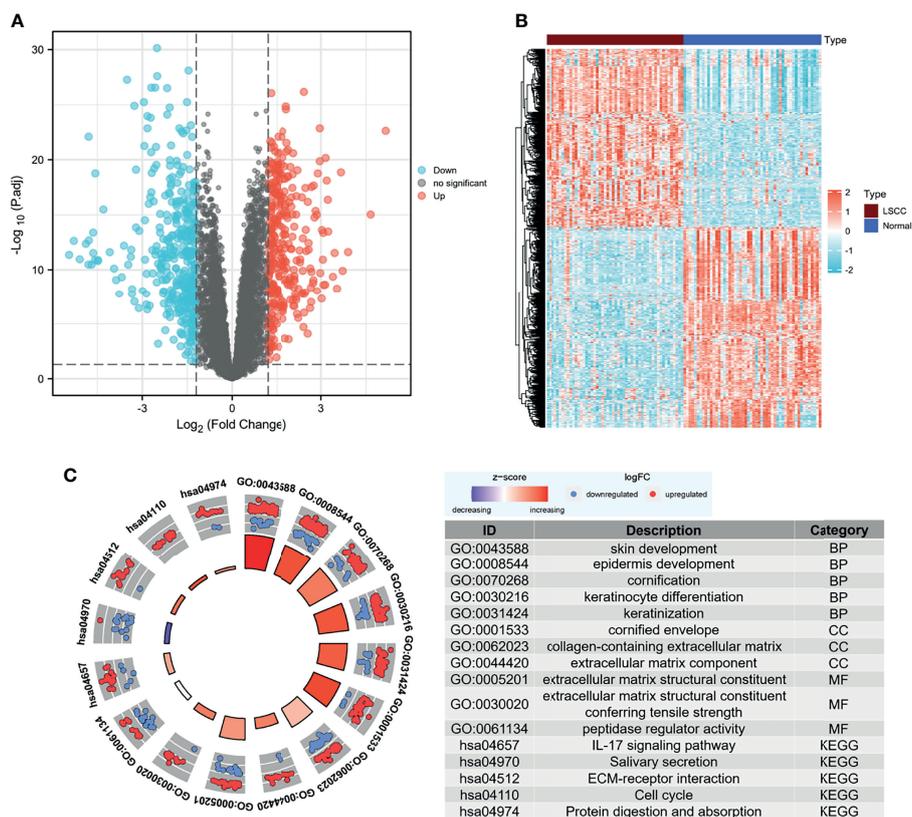
FIGURE 2 | Screening for DEGs. **(A)** Volcano map of DEGs between LSCC and normal samples in the GSE142083 dataset. The red plots in the volcano represent upregulated genes, and the blue points represent downregulated genes. **(B)** Heatmap of all the DEGs. The color in heatmaps from blue to yellow shows the progression from low expression to high expression, respectively. **(C)** GO and KEGG analysis of the DEGs. The outer circle shows the scatter plot of the assigned gene log$_2$fold change for all terms: red points show genes that exhibited increased expression, whereas the blue points represent genes that exhibited decreased expression. The inner circle indicates the Z-score value and the number of genes: red represents the higher z-score value, and purple represents a lower Z-score value. DEG, differentially expressed gene; BP, biological process; CC, cell component; MF, molecular function; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; LSCC, laryngeal squamous cell carcinoma.

A total of 28 modules were identified *via* average linkage hierarchical clustering, and each module is represented using a different color in **Figure 3C**. Among the modules, the modules midnight-blue, blue, and green-yellow exhibited a strong correlation with cancer traits (cor ≥ 0.8, *p*-value < 0.001) (**Figures 3D–G**). Therefore, these modules were selected as clinically significant modules for further analysis. A set of 400 selected genes was identified for the network heatmap (**Supplementary Figure 3**).

GO and KEGG analyses were used to assess the biological processes, molecular functions, cellular components, and KEGG pathways of genes in the modules. The midnight-blue module was primarily related to glycosylation events and glycosphingolipid biosynthesis, whereas the blue module was primarily associated with the DNA replication and the cell cycle pathway. The green-yellow module was associated with RNA splicing (**Supplementary Figure 4**).

Identification of the Key Genes
Hub genes were screened using a |module membership (MM) score| > 0.8 and |gene significance (GS) score| > 0.2 as the cutoff

criteria (34–36). Based on this criterion, we subsequently sorted the genes according to their connectivity to select candidate genes (**Figures 3E–G**). A total of 41 genes in the midnight-blue module, 120 genes in the blue module, and 99 genes in the green-yellow module were screened out.

To identify the "real" key genes, we compared these candidate genes obtained from WGCNA with the 701 DEGs. A total of 85 key genes from the midnight-blue module, blue module, and DEGs were identified (**Figure 3H**).

Validation of Key Genes *via* Survival Analysis
To validate and explore the prognostic values of these "real" key genes in LSCC, we used the TCGA-HNSC-Larynx dataset (*N* = 111) as training cohort. Univariate Cox regression analysis of the 85 key genes was conducted with regard to the OS of samples from the training cohort (**Supplementary Table 1**). Six genes with *p* < 0.1 were then included for multivariate Cox regression model analysis to establish a prognostic model. Finally, we identified three genes—*SLC35C1*, *HOXB7*, and *TEDC2*—as
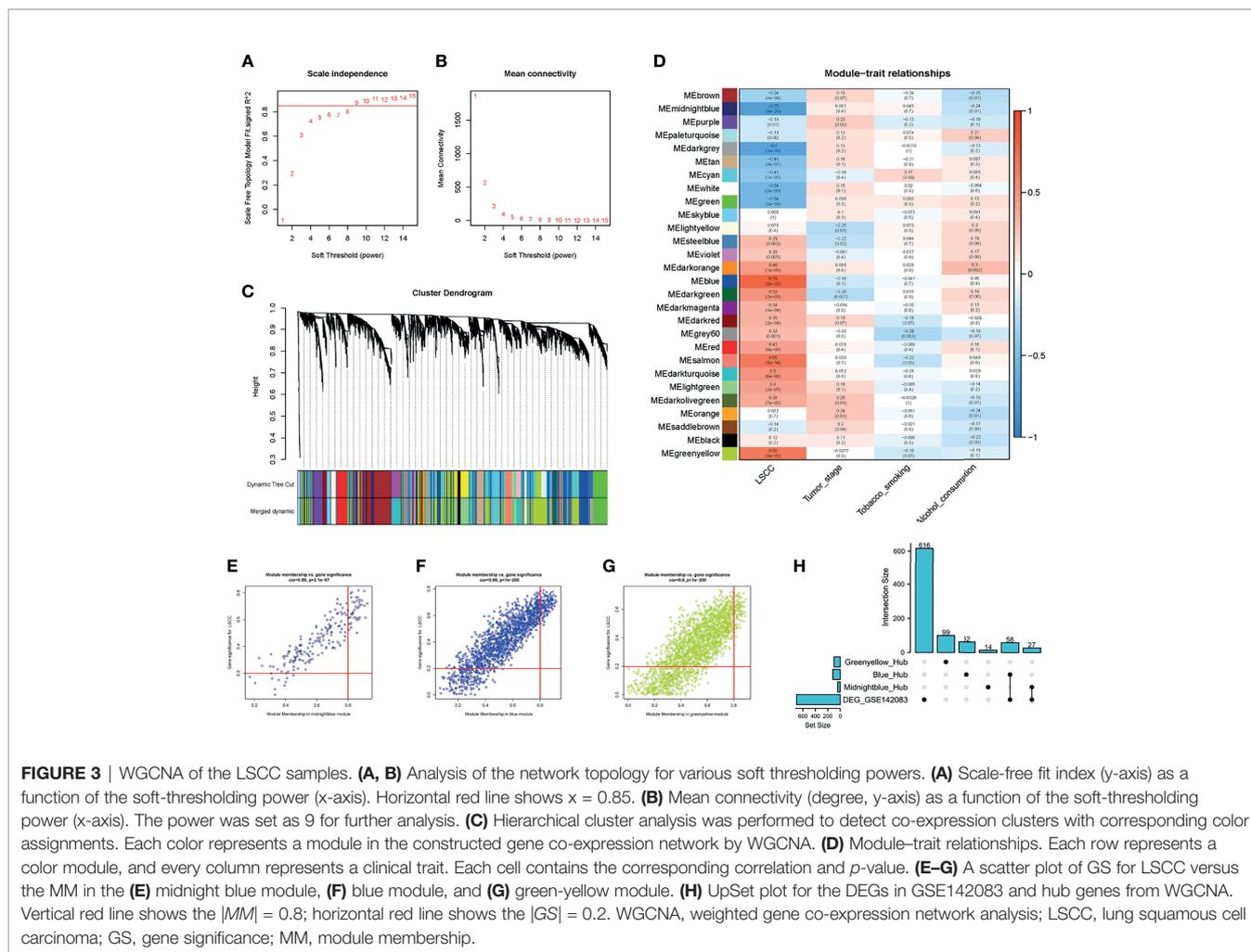
FIGURE 3 | WGCNA of the LSCC samples. **(A, B)** Analysis of the network topology for various soft thresholding powers. **(A)** Scale-free fit index (y-axis) as a function of the soft-thresholding power (x-axis). Horizontal red line shows x = 0.85. **(B)** Mean connectivity (degree, y-axis) as a function of the soft-thresholding power (x-axis). The power was set as 9 for further analysis. **(C)** Hierarchical cluster analysis was performed to detect co-expression clusters with corresponding color assignments. Each color represents a module in the constructed gene co-expression network by WGCNA. **(D)** Module–trait relationships. Each row represents a color module, and every column represents a clinical trait. Each cell contains the corresponding correlation and p-value. **(E–G)** A scatter plot of GS for LSCC versus the MM in the **(E)** midnight blue module, **(F)** blue module, and **(G)** green-yellow module. **(H)** UpSet plot for the DEGs in GSE142083 and hub genes from WGCNA. Vertical red line shows the |MM| = 0.8; horizontal red line shows the |GS| = 0.2. WGCNA, weighted gene co-expression network analysis; LSCC, lung squamous cell carcinoma; GS, gene significance; MM, module membership.

potential independent prognostic markers for LSCC (**Supplementary Figure 5**). The risk score for each patient was calculated using the following formula: risk score = 0.703437*SLC35C1* + 0.833199*HOXB7* + (-0.891338)*TEDC2*.

According to the level of risk score, samples from the two cohorts were divided into a high-risk group and a low-risk group based on a cutoff value of 50% in both training and testing cohorts, as shown in **Figures 4A**–**G**. The risk score was significantly associated with the survival time of patients with LSCC in the training cohort (TCGA-HNSC-Larynx) (**Figure 4A**) and testing cohort (GSE27020) (**Figure 4F**). The ROC curve was then used to evaluate the accuracy of the survival analysis. In the training and testing cohorts, the area under the curve (AUC) was 0.726 and 0.645, respectively, indicating that the prediction effect was good (**Figures 4B, G**). Besides, patients with continuous risk scores harbored various clinical outcomes in different groups (**Figures 4C, D, H, I**). Additionally, all three genes (*SLC35C1*, *HOXB7*, and *TEDC2*) were significantly associated with poor prognosis and unhealthy living habits, in either the training cohort (TCGA-HNSC-Larynx), the testing cohort (GSE27020), or other LSCC datasets (GSE39366 and GSE127165) (**Figures 4E, J** and **Supplementary Figure 6**).

Gene Mutation in the Different Risk Groups

Analysis of differences in gene mutations between high-risk and low-risk groups showed that the proportion of patients with gene mutations was 92.73% (51 in 55) and 96.30% (52 in 54) in the high-risk and low-risk score groups, respectively (**Figures 5A, B**). The mutation frequency of *TP53* was observably higher in the high-risk group (82%) than in the low-risk group (69%) (**Figures 5A, B**).

Comparison of Tumor-Infiltrating Immune Cell Abundance Between Different Risk Groups

The distribution of tumor-infiltrating immune cells is an important indicator of a patient's lymph node status and prognosis. In order to explore the relationship of the prognostic model and the tumor immune microenvironment, we analyzed the tumor-infiltrating immune cell abundance in the TCGA-HNSC-Larynx dataset using the CIBERSORT algorithm (**Figure 5C**). The results showed that the infiltration levels of the following cells: "T cells CD4 memory activated", and "T cells
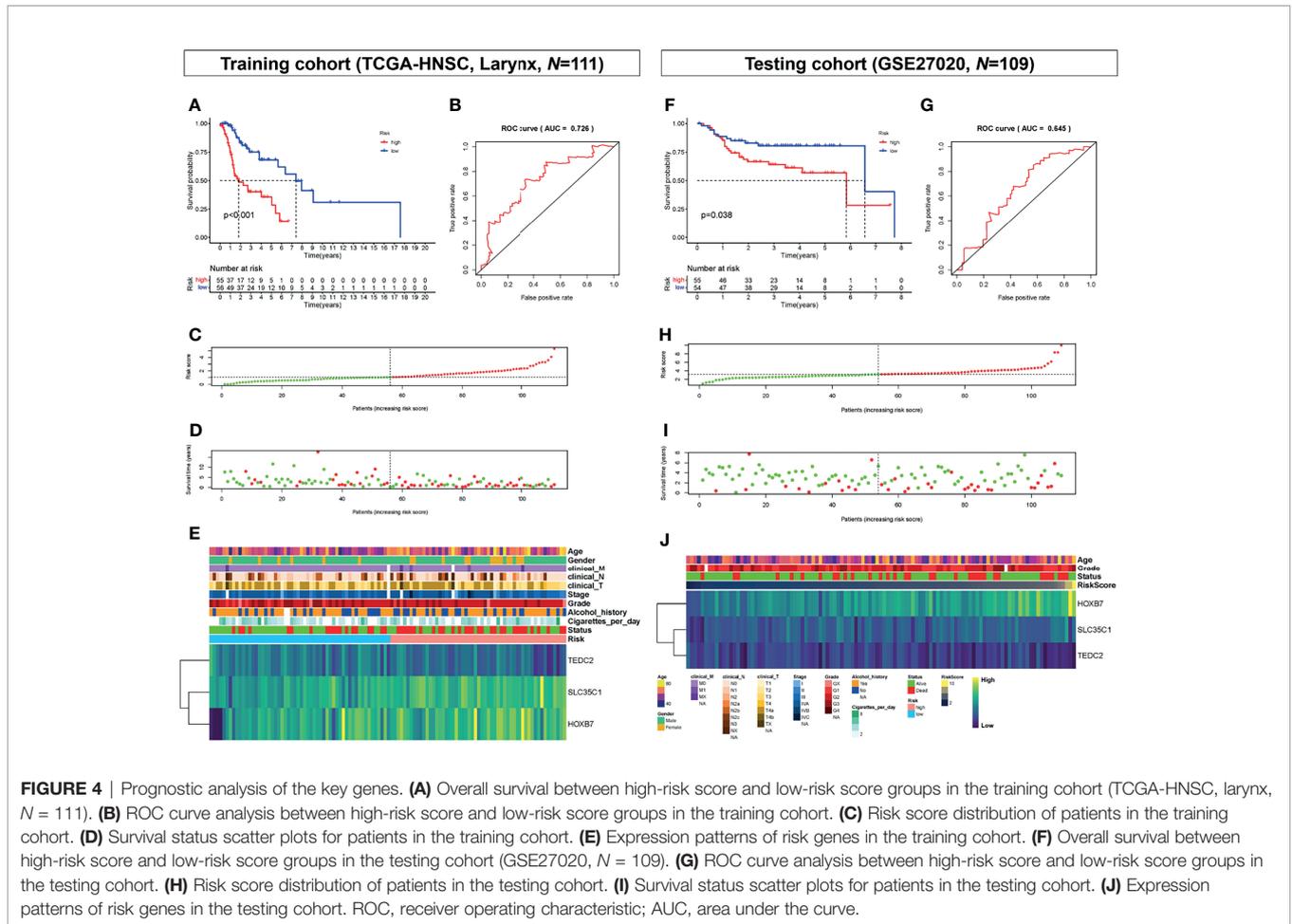
FIGURE 4 | Prognostic analysis of the key genes. **(A)** Overall survival between high-risk score and low-risk score groups in the training cohort (TCGA-HNSC, larynx, N = 111). **(B)** ROC curve analysis between high-risk score and low-risk score groups in the training cohort. **(C)** Risk score distribution of patients in the training cohort. **(D)** Survival status scatter plots for patients in the training cohort. **(E)** Expression patterns of risk genes in the training cohort. **(F)** Overall survival between high-risk score and low-risk score groups in the testing cohort (GSE27020, N = 109). **(G)** ROC curve analysis between high-risk score and low-risk score groups in the testing cohort. **(H)** Risk score distribution of patients in the testing cohort. **(I)** Survival status scatter plots for patients in the testing cohort. **(J)** Expression patterns of risk genes in the testing cohort. ROC, receiver operating characteristic; AUC, area under the curve.

regulatory" were significantly higher in the low-risk group than in the high-risk group. In contrast, the infiltration levels of "Macrophages M2" and "Mast cells resting" were significantly lower in the low-risk group. A similar abundance was also estimated in the GSE27020 dataset (**Supplementary Figure 7**).

Functional Enrichment of Prognostic Genes

GSVA and GSEA were conducted to search for GO and KEGG pathways in which the prognostic genes or risk scores were enriched in samples with high-risk levels from the TCGA-HNSC-Larynx dataset. Metabolism and cancer-associated pathways including "Glyoxylate and dicarboxylate metabolism", "Protein export", "Sphingolipid metabolism", "Lysosome", and "Bladder cancer" were significantly enriched with low expression of most prognostic genes and high-risk score (**Figure 6A**). Besides, GSEA based on DEGs between the high- and low-risk score groups also revealed that "Apical part of cell", "Anion transmembrane transporter activity", and "Tissue homeostasis" were significantly enriched with GO BP (**Figure 6B**) and "Metabolism of xenobiotics by cytochrome p450", "Cardiac muscle contraction", and "Systemic lupus erythematosus" were significantly enriched with KEGG analysis (**Figure 6C**).

Single-Cell Transcriptomic Context of the Prognostic Genes

To further verify the relationship between the prognostic genes and risk score in LSCC, we employed single-cell RNA-Seq data from the GSE150321 dataset (25), which comprised data from two LSCC samples. We identified five cell groups, namely tumor cells, immune cells, epithelial cells, mesenchymal cells and endothelial cells (**Figure 7A**). We then calculated the risk score for each cell and plotted them in a t-SNE plot and the violin plots (**Figure 7B** and **Supplementary Figure 8**). Tumor cells and some immune cells showed higher risk scores than other cell types.

Next, we profiled the tumor cells and arranged them into five clusters (**Figure 7C** and **Supplementary Figures 9A, B**). Pseudotime and trajectory analyses of the tumor cells suggested a continuous cell fate, starting at tumor 3 and progressing toward tumor 1 and tumor 4, with tumor 2 and tumor 5 being a transitioning state (**Figures 7D, E**). Diffusion mapping placed tumor 4 and tumor 1 within populations of high-risk score cells, indicating a differentiation trajectory from low-risk tumor cells to high-risk tumor cells (**Figure 7F**). Together with previous results, these results further supported that the identified prognostic genes influence the LSCC progression.
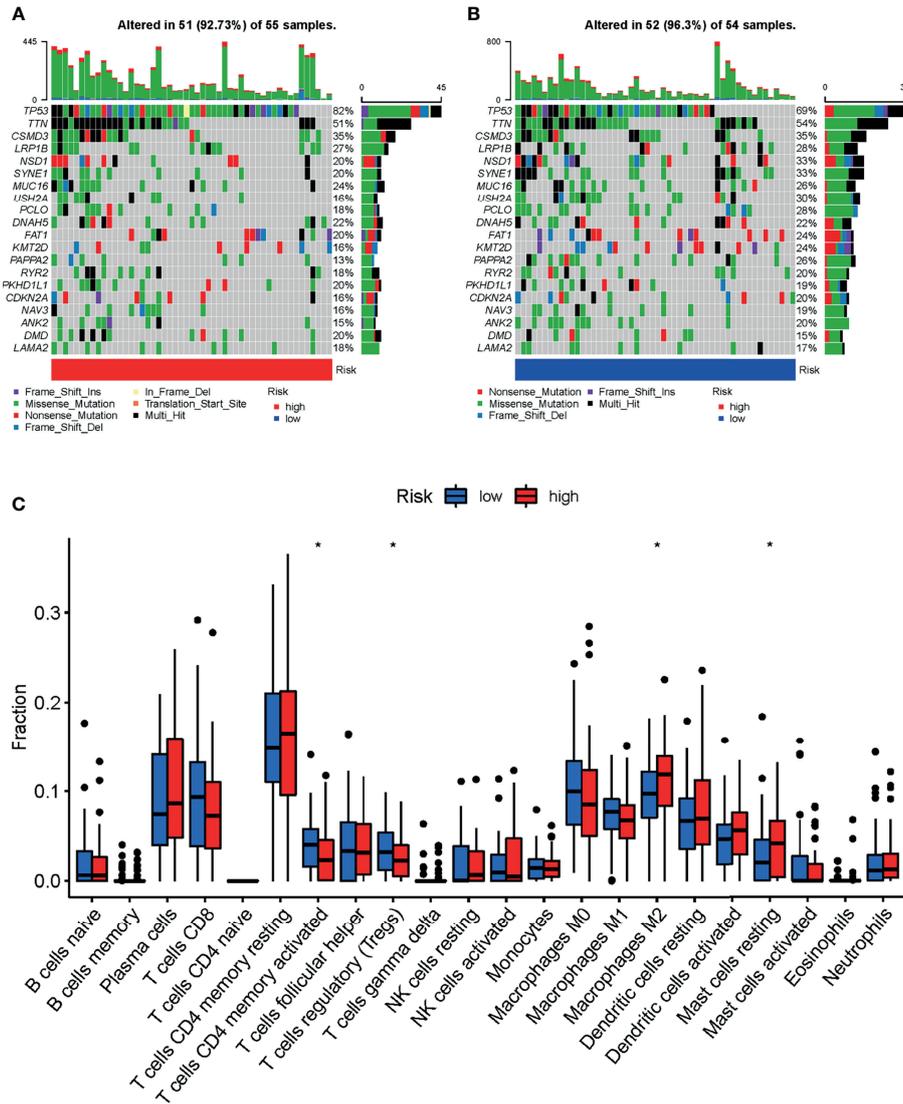
FIGURE 5 | Landscape of mutation profiles and tumor-infiltrating immune cells abundance between high and low-risk LSCC patients. Waterfall plots represent mutation information in each sample of the **(A)** high- and **(B)** low-risk group of TCGA-HNSC-Larynx dataset. **(C)** The difference of 22 tumor-infiltrating immune cells between the high- and low-risk group of TCGA-HNSC-Larynx dataset. *$p < 0.05$; LSCC, lung squamous cell carcinoma.

## TEDC2 Affects Tumor Cell Proliferation and Migration

Based on the results of the previous analysis, we selected *TEDC2*, one of the three prognostic genes with less research involving LSCC, for further analysis. We first verified the expression level of *TEDC2* in different datasets, and they both indicated that *TEDC2* was higher in LSCC than in paired normal tissues (**Figure 8A**). Similar results were observed in the LSCC cell line and a precancerous cell line (**Figure 8A**). We therefore used siRNA to inhibit the expression of *TEDC2*. Results from RT-PCR indicated that both of the two *TEDC2* siRNAs in AMC-HN-8 showed a good silenced efficiency (**Figure 8B**).

CCK-8 assays showed that silencing the expression of *TEDC2* promoted AMC-HN-8 cell proliferation (**Figure 8C**). Cell-cycle

genes, including *PCNA*, *CNNB1*, and *CNND1*, were both downregulated when *TEDC2* was knocked down (**Figure 8D**). Results of the transwell assays showed that knockdown of *TEDC2* promoted migration (**Figure 8E**). Based on these validations, we concluded that silenced *TEDC2*, to a certain extent, promoted LSCC cell proliferation and migration.

DISCUSSION

In the present study, 701 DEGs in the GSE142083 dataset were screened. We firstly constructed weighted co-expression networks using the WGCNA algorithm based on the
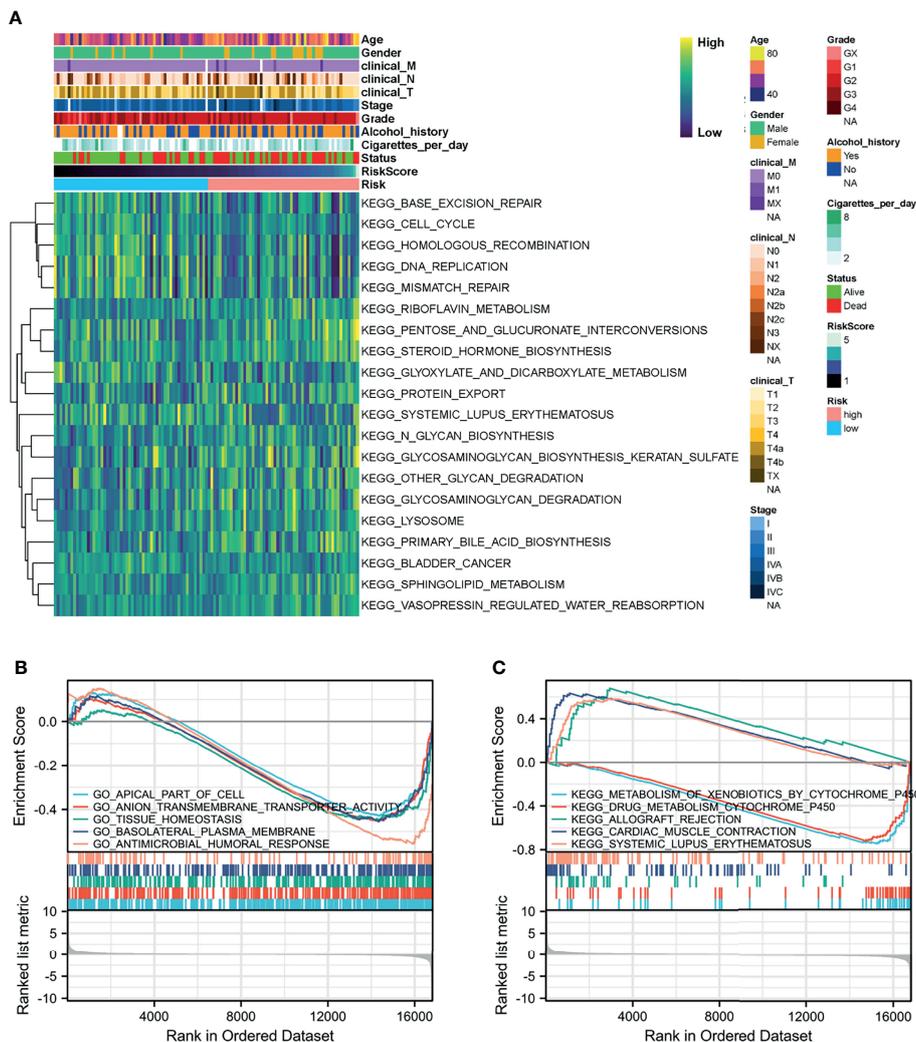
FIGURE 6 | Function analysis of the prognostic model. **(A)** Heatmap of KEGG analysis based on risk score in the TCGA-HNSC-Larynx dataset. **(B, C)** GSEA analysis for GO and KEGG enrichment in the TCGA-HNSC-Larynx dataset according to risk score. GSEA, gene set enrichment analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology.

GSE142083 dataset. Three modules in LSCC tissues were detected based on the co-expression network. After intersecting with the DEGs, 85 key genes were screened out. Following univariate and multivariate analyses, a novel prognostic model for LSCC based on three genes (*SLC35C1*, *HOXB7*, and *TEDC2*) was established.

LSCC is a leading malignant type of HNSC. There is an urgent need to identify potential targets for drugs and biomarkers to improve early diagnosis and outcomes (37–40). A growing body of evidence suggests that solid tumor was stiffer than normal tissue (41). Moreover, the increased deposition of ECM proteins is one of the main reasons for this. The ECM is a fundamental and important component of all tissue organs; it also interacts with tumor cells and regulates tumor growth, proliferation, differentiation, adhesion, and metastasis (42–44). Here, we confirmed that most of the core ECM genes, such as

collagens, proteoglycans, and glycoproteins, were highly expressed in the LSCC tissues than in the adjacent normal mucosal tissues (**Supplementary Figure 2**). Moreover, ECM-associated biological processes were enriched in the DEGs of the GSE142083 dataset (**Figure 2C**). These results further confirmed the importance of ECM accumulation in LSCC. Due to the excess deposition of ECM, tumors in this region may be regarded as difficult to treat (45–47). Therefore, exploring novel methods to reduce ECM deposition may be a strategy for LSCC treatment.

SLC35C1, as a member of the solute carrier (SLC) family, encoded a guanosine 5′-diphosphate (GDP)-fucose transporter 1 channel that critically regulated the fucosylation of glycans. Mutation of *SLC35C1* was found to cause leukocyte adhesion deficiency, which was a rare congenital disease due to the defect in the biosynthesis of selectin ligands on leukocytes (48).
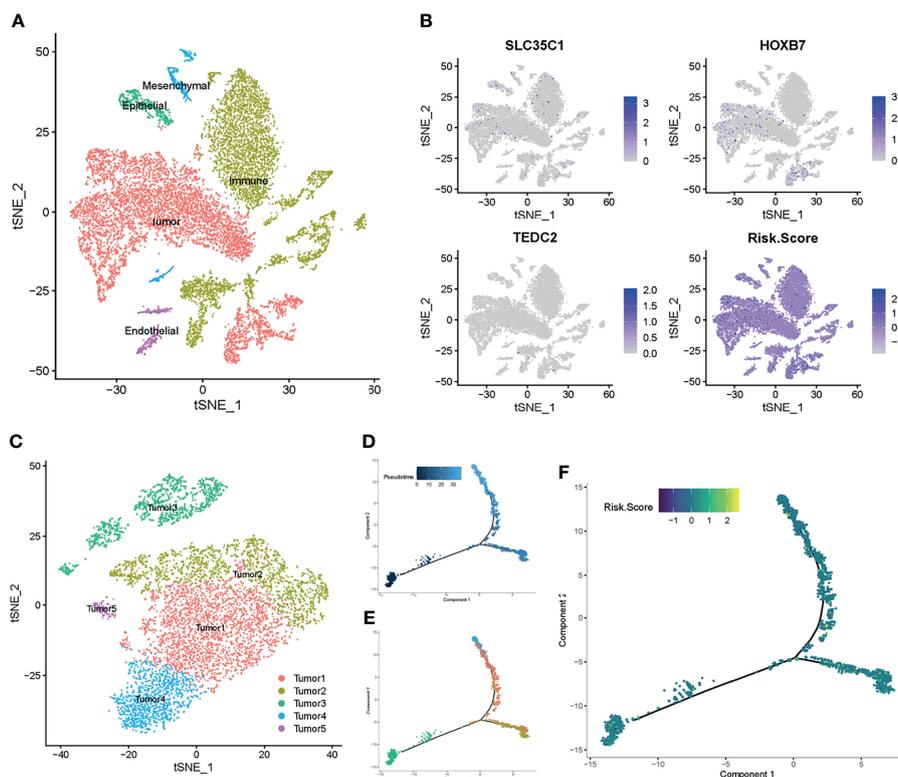
FIGURE 7 | Prognostic expression profile based on single-cell sequencing analysis. **(A)** Composition and distribution of single cells from GSE150321. **(B)** The distribution profile of *SLC35C1*, *HOXB7*, *TEDC2*, and risk score for each cell by the t-SNE plot. **(C)** t-SNE embedding of the tumor cells. **(D, E)** Pseudotime and trajectory analysis revealed the tendency curve from tumor 3 cluster to tumor 4 and tumor 1 clusters. Y-axis means the value of principal component 1 (the first principal direction of maximum sample change), and X-axis means the value of principal component 2 (the second principal direction of maximum sample change). **(F)** The expression profile of risk score annotated in pseudotime and trajectory plot. t-SNE, t-distributed stochastic neighbor embedding.

Fucosylation has been found to be an important type of posttranslational modification in many types of cancer (49). Moriwaki et al. have reported that *SLC35C1*, as a GDP-fucose transporter, was highly expressed and with an increased level of fucosylation in hepatocellular carcinoma (50). In colon cancer, however, Deng et al. showed that *SLC35C1* was reduced in all colon cancers and they further proved that loss of *SLC35C1* may promote colon cancer progression through the activation of the Wnt signaling pathway (51). In our analysis, we firstly identified that *SLC35C1* might be a potential independent prognostic marker for LSCC and the upregulation of *SLC35C1* may have an association with high risk score and poor prognosis of LSCC patient (**Figures 4E, J**). Nevertheless, future studies are warranted to clarify the underlying mechanisms of *SLC35C1* and fucosylation in LSCC.

Homeobox B7 (*HOXB7*) is a member of the Antp homeobox family and encodes a protein with a homeobox DNA-binding domain. As a sequence-specific transcription factor, HOXB7 has been shown to have specific functions in cellular proliferation, differentiation, and death (52). To date, *HOXB7* has been implicated to be aberrantly expressed in several types of cancers, including breast cancer (53), gliomas (54), gastric cancer (55), esophageal squamous cell carcinoma (56),

intrahepatic cholangiocarcinoma (57), and cervical cancer (58). Furthermore, study from Wu et al. (59) and study from Mo et al. (60) both proved that *HOXB7* may serve as an oncogene for HNSCC and LSCC. Our analysis of TCGA-HNSC-Larynx, GSE27020, GSE39366, and GSE127165 further confirmed the role of the potential prognostic marker of HOXB7 in LSCC and HNSCC. However, its detailed mechanism still requires further research.

In comparison with the other two prognostic genes, *TEDC2* has not been intensively investigated in cancer. *TEDC2* is also called as tubulin epsilon and delta complex 2 or *C16ORF59*. As the name implies, TEDC2 is a component of the cytoskeleton. A study from Hsu et al. showed that *TEDC2* was differentially expression between lung adenocarcinoma and paired normal tissues (61). Recently, Meng et al. used the monozygotic twin-pair database to detect the alteration of DNA methylation after alcohol drinking, and they found that a hypermethylation of cg07326074, located in *TEDC2*, was associated with alcohol consumption (62). This drinking-related methylation was also suspected to affect *TEDC2* gene function. Alcohol consumption is an established risk factor for HNSC (63). In TCGA-HNSC, we also confirmed that the expression of *TEDC2* was increased in the alcohol consumption cohort (**Supplementary Figure 10**), and knocking down *TEDC2*
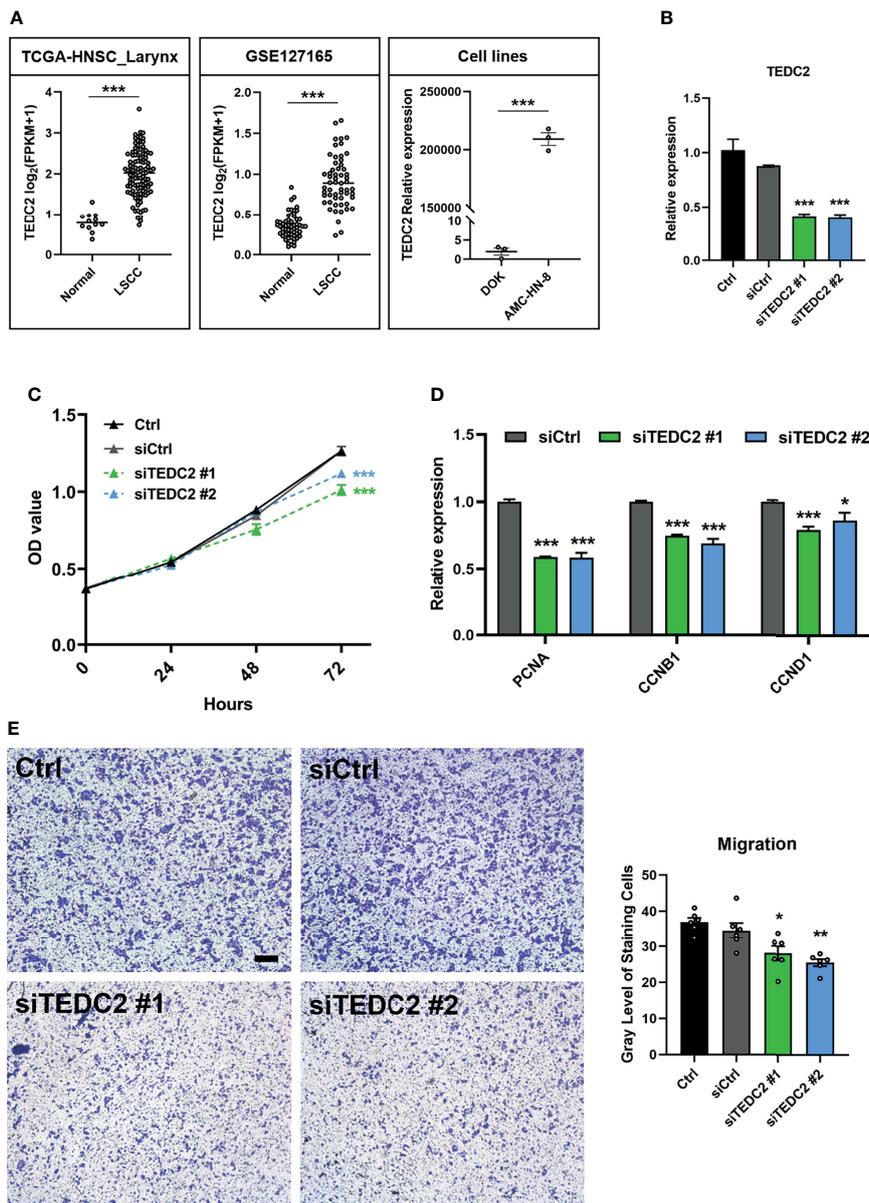
FIGURE 8 | *TEDC2* affects tumor cell proliferation and migration. **(A)** Expression profile of *TEDC2* between normal and LSCC samples in the TCGA-HNSC-Larynx dataset (left) and GSE127165 (middle). *TEDC2* expression between laryngeal cancer cell line AMC-HN-8 and dysplastic oral keratinocyte cell line DOK (right). **(B)** AMC-HN-8 cells were transfected with siRNAs targeting *TEDC2* (siTEDC2 #1 and siTEDC2 #2) or normal control siRNAs (siCtrl) for 48 h. The expression of *TEDC2* was examined by RT-PCR. **(C)** AMC-HN-8 cells were transfected with siRNAs targeting *TEDC2* (siTEDC2 #1 and siTEDC2 #2) or normal control siRNAs (siCtrl), then cell proliferation was determined by CCK-8 assay. **(D)** Expressions of *PCNA*, *CCNB1*, and *CCND1* were examined by RT-PCR. **(E)** The migration abilities of AMC-HN-8 cells after transfected with siRNAs targeting *TEDC2* (siTEDC2 #1 and siTEDC2 #2) or normal control siRNAs (siCtrl) were detected by transwell assays. Bar, 200 μm. Error bars represent SEM of at least three independent experiments; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

could inhibit the capacities of proliferation and migration in the LSCC cell line (**Figures 8A–E**). Therefore, as a tumor-promoting gene, *TEDC2* may increase in the condition of alcohol consumption and thus promote tumorigenesis and metastasis. In the future, detailed mechanism research still needed to elucidate the association of *TEDC2* and tumor.

CONCLUSION

Following the screening of DEGs, and WGCNA of LSCC and adjacent normal tissues, a novel prognostic model based on three genes (*SLC35C1*, *HOXB7*, and *TEDC2*) for LSCC was identified. These prognostic genes may play a significant role in improving

the prognostic prediction of patients with LSCC and may also serve as therapeutic targets and/or biomarkers for LSCC.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

HH and AP conceived and directed the project. CH and YD collected the data and information. CH, LH, JH, and YC analyzed and interpreted the data. CH and HH wrote the manuscript with the help of all the other authors. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.779153/full#supplementary-material

REFERENCES

1. Liang J, Zhu X, Zeng W, Yu T, Fang F, Zhao Y. Which Risk Factors Are Associated With Stomal Recurrence After Total Laryngectomy for Laryngeal Cancer? A Meta-Analysis of the Last 30 Years. *Braz J Otorhinolaryngol* (2020) 86:502–12. doi: 10.1016/j.bjorl.2020.03.002

2. Lin HW, Bhattacharyya N. Staging and Survival Analysis for Nonsquamous Cell Carcinomas of the Larynx. *Laryngoscope* (2008) 118:1003–13. doi: 10.1097/MLG.0b013e3181671b3d

3. Chen X, Gao L, Sturgis EM, Liang Z, Zhu Y, Xia X, et al. HPV16 DNA and Integration in Normal and Malignant Epithelium: Implications for the Etiology of Laryngeal Squamous Cell Carcinoma. *Ann Oncol* (2017) 28:1105–10. doi: 10.1093/annonc/mdx027

4. Marioni G, Marchese-Ragona R, Cartei G, Marchese F, Staffieri A. Current Opinion in Diagnosis and Treatment of Laryngeal Carcinoma. *Cancer Treat Rev* (2006) 32:504–15. doi: 10.1016/j.ctrv.2006.07.002

5. Haddad RI, Posner M, Hitt R, Cohen E, Schulten J, Lefebvre JL, et al. Induction Chemotherapy in Locally Advanced Squamous Cell Carcinoma of the Head and Neck: Role, Controversy, and Future Directions. *Ann Oncol* (2018) 29:1130–40. doi: 10.1093/annonc/mdy102

6. Gao C, Hu S. Mir-506 is a YAP1-Dependent Tumor Suppressor in Laryngeal Squamous Cell Carcinoma. *Cancer Biol Ther* (2019) 20:826–36. doi: 10.1080/15384047.2018.1564569

7. Grosselin K, Durand A, Marsolier J, Poitou A, Marangoni E, Nemati F, et al. High-Throughput Single-Cell Chip-Seq Identifies Heterogeneity of Chromatin States in Breast Cancer. *Nat Genet* (2019) 51:1060–6. doi: 10.1038/s41588-019-0424-9

8. Zhang P, Xia JH, Zhu J, Gao P, Tian YJ, Du M, et al. High-Throughput Screening of Prostate Cancer Risk Loci by Single Nucleotide Polymorphisms Sequencing. *Nat Commun* (2018) 9:2022. doi: 10.1038/s41467-018-04451-x

9. Braun S, Enculescu M, Setty ST, Cortes-Lopez M, de Almeida BP, Sutandy F, et al. Decoding a Cancer-Relevant Splicing Decision in the RON Proto-Oncogene Using High-Throughput Mutagenesis. *Nat Commun* (2018) 9:3315. doi: 10.1038/s41467-018-05748-7

10. Teufel M, Seidel H, Kochert K, Meinhardt G, Finn RS, Llovet JM, et al. Biomarkers Associated With Response to Regorafenib in Patients With Hepatocellular Carcinoma. *Gastroenterology* (2019) 156:1731–41. doi: 10.1053/j.gastro.2019.01.261

11. Guo B, Wu S, Zhu X, Zhang L, Deng J, Li F, et al. Micropeptide CIP2A-BP Encoded by LINC00665 Inhibits Triple-Negative Breast Cancer Progression. *EMBO J* (2020) 39:e102190. doi: 10.15252/embj.2019102190

12. Barros-Filho MC, Dos RM, Beltrami CM, de Mello J, Marchi FA, Kuasne H, et al. DNA Methylation-Based Method to Differentiate Malignant From Benign Thyroid Lesions. *Thyroid* (2019) 29:1244–54. doi: 10.1089/thy.2018.0458

13. Jin Y, Yang Y. Identification and Analysis of Genes Associated With Head and Neck Squamous Cell Carcinoma by Integrated Bioinformatics Methods. *Mol Genet Genomic Med* (2019) 7:e857. doi: 10.1002/mgg3.857

14. Langfelder P, Horvath S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559

15. Ding M, Li F, Wang B, Chi G, Liu H. A Comprehensive Analysis of WGCNA and Serum Metabolomics Manifests the Lung Cancer-Associated Disordered Glucose Metabolism. *J Cell Biochem* (2019) 120:10855–63. doi: 10.1002/jcb.28377

16. Liu J, Zhou S, Li S, Jiang Y, Wan Y, Ma X, et al. Eleven Genes Associated With Progression and Prognosis of Endometrial Cancer (EC) Identified by Comprehensive Bioinformatics Analysis. *Cancer Cell Int* (2019) 19:136. doi: 10.1186/s12935-019-0859-1

17. Zhang H, Zhao X, Wang M, Ji W. Key Modules and Hub Genes Identified by Coexpression Network Analysis for Revealing Novel Biomarkers for Larynx Squamous Cell Carcinoma. *J Cell Biochem* (2019) 120:19832–40. doi: 10.1002/jcb.29288

18. Li XT. Identification of Key Genes for Laryngeal Squamous Cell Carcinoma Using Weighted Co-Expression Network Analysis. *Oncol Lett* (2016) 11:3327–31. doi: 10.3892/ol.2016.4378

19. Yang CW, Wang SF, Yang XL, Wang L, Niu L, Liu JX. Identification of Gene Expression Models for Laryngeal Squamous Cell Carcinoma Using Co-Expression Network Analysis. *Med (Baltimore)* (2018) 97:e9738. doi: 10.1097/MD.0000000000009738

20. Li CY, Cai JH, Tsai J, Wang C. Identification of Hub Genes Associated With Development of Head and Neck Squamous Cell Carcinoma by Integrated Bioinformatics Analysis. *Front Oncol* (2020) 10:681. doi: 10.3389/fonc.2020.00681

21. Gao W, Zhang Y, Luo H, Niu M, Zheng X, Hu W, et al. Targeting SKA3 Suppresses the Proliferation and Chemoresistance of Laryngeal Squamous Cell Carcinoma *via* Impairing PLK1-AKT Axis-Mediated Glycolysis. *Cell Death Dis* (2020) 11:919. doi: 10.1038/s41419-020-03104-6

22. Fountzilas E, Kotoula V, Angouridakis N, Karasmanis I, Wirtz RM, Eleftheraki AG, et al. Identification and Validation of a Multigene Predictor of Recurrence in Primary Laryngeal Cancer. *PloS One* (2013) 8:e70429. doi: 10.1371/journal.pone.0070429

23. Walter V, Yin X, Wilkerson MD, Cabanski CR, Zhao N, Du Y, et al. Molecular Subtypes in Head and Neck Cancer Exhibit Distinct Patterns of Chromosomal Gain and Loss of Canonical Cancer Genes. *PloS One* (2013) 8:e56823. doi: 10.1371/journal.pone.0056823

24. Wu Y, Zhang Y, Zheng X, Dai F, Lu Y, Dai L, et al. Circular RNA Circcoro1c Promotes Laryngeal Squamous Cell Carcinoma Progression by Modulating the Let-7c-5p/PBX3 Axis. *Mol Cancer* (2020) 19:99. doi: 10.1186/s12943-020-01215-4

25. Song L, Zhang S, Yu S, Ma F, Wang B, Zhang C, et al. Cellular Heterogeneity Landscape in Laryngeal Squamous Cell Carcinoma. *Int J Cancer* (2020) 147:2879–90. doi: 10.1002/ijc.33192

26. Hao Y, Hao S, Andersen-Nissen E, Mauck WR, Zheng S, Butler A, et al. Integrated Analysis of Multimodal Single-Cell Data. *Cell* (2021) 184:3573–87. doi: 10.1016/j.cell.2021.04.048

27. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-Cell Mrna Quantification and Differential Analysis With Census. *Nat Methods* (2017) 14:309–15. doi: 10.1038/nmeth.4150

28. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007

29. Naba A, Clauser KR, Ding H, Whittaker CA, Carr SA, Hynes RO. The Extracellular Matrix: Tools and Insights for the "Omics" Era. *Matrix Biol* (2016) 49:10–24. doi: 10.1016/j.matbio.2015.06.003

30. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *Omics* (2012) 16:284–7. doi: 10.1089/omi.2011.0118

31. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust Enumeration of Cell Subsets From Tissue Expression Profiles. *Nat Methods* (2015) 12:453–7. doi: 10.1038/nmeth.3337

32. Hanzelmann S, Castelo R, Guinney J. GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

33. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc Natl Acad Sci USA* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102

34. Li B, Xiang W, Qin J, Xu Q, Feng S, Wang Y, et al. Co-Expression Network of Long non-Coding RNA and Mrna Reveals Molecular Phenotype Changes in Kidney Development of Prenatal Chlorpyrifos Exposure in a Mouse Model. *Ann Transl Med* (2021) 9:653. doi: 10.21037/atm-20-6632

35. Yi M, Li T, Qin S, Yu S, Chu Q, Li A, et al. Identifying Tumorigenesis and Prognosis-Related Genes of Lung Adenocarcinoma: Based on Weighted Gene Coexpression Network Analysis. *BioMed Res Int* (2020) 2020:4169691. doi: 10.1155/2020/4169691

36. Liang W, Sun F, Zhao Y, Shan L, Lou H. Identification of Susceptibility Modules and Genes for Cardiovascular Disease in Diabetic Patients Using WGCNA Analysis. *J Diabetes Res* (2020) 2020:4178639. doi: 10.1155/2020/4178639

37. Hui L, Zhang J, Ding X, Guo X, Jiang X. Matrix Stiffness Regulates the Proliferation, Stemness and Chemoresistance of Laryngeal Squamous Cancer Cells. *Int J Oncol* (2017) 50:1439–47. doi: 10.3892/ijo.2017.3877

38. Ma J, Hu X, Dai B, Wang Q, Wang H. Bioinformatics Analysis of Laryngeal Squamous Cell Carcinoma: Seeking Key Candidate Genes and Pathways. *Peerj* (2021) 9:e11259. doi: 10.7717/peerj.11259

39. Wang P, Li QY, Sun YN, Wang JT, Liu M. Long Noncoding RNA NEAT1: A Potential Biomarker in the Progression of Laryngeal Squamous Cell Carcinoma. *ORL J Otorhinolaryngol Relat Spec* (2021) 83:464–70. doi: 10.1159/000515228

40. Yang G, Fang J, Shi Q, Wang R, Lian M, Ma H, et al. Screening of Molecular Markers of Induced Chemotherapy in Supraglottic Laryngeal Squamouscell Carcinoma. *World J Otorhinolaryngol Head Neck Surg* (2020) 6:34–40. doi: 10.1016/j.wjorl.2019.05.001

41. Chen X, Wanggou S, Bodalia A, Zhu M, Dong W, Fan JJ, et al. A Feedforward Mechanism Mediated by Mechanosensitive Ion Channel PIEZO1 and Tissue Mechanics Promotes Glioma Aggression. *Neuron* (2018) 100:799–815. doi: 10.1016/j.neuron.2018.09.046

42. Yang X, Wu Q, Wu F, Zhong Y. Differential Expression of COL4A3 and Collagen in Upward and Downward Progressing Types of Nasopharyngeal Carcinoma. *Oncol Lett* (2021) 21:223. doi: 10.3892/ol.2021.12484

43. He Y, Liu R, Yang M, Bi W, Zhou L, Zhang S, et al. Identification of VWF as a Novel Biomarker in Lung Adenocarcinoma by Comprehensive Analysis. *Front Oncol* (2021) 11:639600. doi: 10.3389/fonc.2021.639600

44. Liot S, Balas J, Aubert A, Prigent L, Mercier-Gouy P, Verrier B, et al. Stroma Involvement in Pancreatic Ductal Adenocarcinoma: An Overview Focusing on Extracellular Matrix Proteins. *Front Immunol* (2021) 12:612271. doi: 10.3389/fimmu.2021.612271

45. Holle AW, Young JL, Spatz JP. *In Vitro* Cancer Cell-ECM Interactions Inform. *Vivo Cancer Treat Adv Drug Delivery Rev* (2016) 97:270–9. doi: 10.1016/j.addr.2015.10.007

46. Rahbari NN, Kedrin D, Incio J, Liu H, Ho WW, Nia HT, et al. Anti-VEGF Therapy Induces ECM Remodeling and Mechanical Barriers to Therapy in Colorectal Cancer Liver Metastases. *Sci Transl Med* (2016) 8:135r–360r. doi: 10.1126/scitranslmed.aaf5219

47. Hosein AN, Brekken RA, Maitra A. Pancreatic Cancer Stroma: An Update on Therapeutic Targeting Strategies. *Nat Rev Gastroenterol Hepatol* (2020) 17:487–505. doi: 10.1038/s41575-020-0300-1

48. Helmus Y, Denecke J, Yakubenia S, Robinson P, Luhn K, Watson DL, et al. Leukocyte Adhesion Deficiency II Patients With a Dual Defect of the GDP-Fucose Transporter. *Blood* (2006) 107:3959–66. doi: 10.1182/blood-2005-08-3334

49. Liao C, An J, Yi S, Tan Z, Wang H, Li H, et al. FUT8 and Protein Core Fucosylation in Tumours: From Diagnosis to Treatment. *J Cancer* (2021) 12:4109–20. doi: 10.7150/jca.58268

50. Moriwaki K, Noda K, Nakagawa T, Asahi M, Yoshihara H, Taniguchi N, et al. A High Expression of GDP-Fucose Transporter in Hepatocellular Carcinoma is a Key Factor for Increases in Fucosylation. *Glycobiology* (2007) 17:1311–20. doi: 10.1093/glycob/cwm094

51. Deng M, Chen Z, Tan J, Liu H. Down-Regulation of SLC35C1 Induces Colon Cancer Through Over-Activating Wnt Pathway. *J Cell Mol Med* (2020) 24:3079–90. doi: 10.1111/jcmm.14969

52. Errico MC, Jin K, Sukumar S, Care A. The Widening Sphere of Influence of HOXB7 in Solid Tumors. *Cancer Res* (2016) 76:2857–62. doi: 10.1158/0008-5472.CAN-15-3444

53. de Bessa GS, Araujo M, Pereira T, Freitas R. HOXB7 Overexpression Leads Triple-Negative Breast Cancer Cells to a Less Aggressive Phenotype. *Biomedicines* (2021) 9:515. doi: 10.3390/biomedicines9050515

54. Zhou X, Liang T, Deng J, Ng K, Li M, Lv C, et al. Differential and Prognostic Significance of HOXB7 in Gliomas. *Front Cell Dev Biol* (2021) 9:697086. doi: 10.3389/fcell.2021.697086

55. Wu J, Long Z, Cai H, Yu S, Liu X. Homeobox B7 Accelerates the Cancer Progression of Gastric Carcinoma Cells by Promoting Epithelial-Mesenchymal Transition (EMT) and Activating Src-FAK Pathway. *Onco Targets Ther* (2019) 12:3743–51. doi: 10.2147/OTT.S198115

56. Shen LY, Zhou T, Du YB, Shi Q, Chen KN. Targeting HOX/PBX Dimer Formation as a Potential Therapeutic Option in Esophageal Squamous Cell Carcinoma. *Cancer Sci* (2019) 110:1735–45. doi: 10.1111/cas.13993

57. Dai L, Hu W, Yang Z, Chen D, He B, Chen Y, et al. Upregulated Expression of HOXB7 in Intrahepatic Cholangiocarcinoma is Associated With Tumor Cell Metastasis and Poor Prognosis. *Lab Invest* (2019) 99:736–48. doi: 10.1038/s41374-018-0150-4

58. Zhao X, Dong W, Luo G, Xie J, Liu J, Yu F. Silencing of Hsa_Circ_0009035 Suppresses Cervical Cancer Progression and Enhances Radiosensitivity Through Microrna 889-3p-Dependent Regulation of HOXB7. *Mol Cell Biol* (2021) 41:e63120. doi: 10.1128/MCB.00631-20

59. Wu X, Li J, Yan T, Ke X, Li X, Zhu Y, et al. HOXB7 Acts as an Oncogenic Biomarker in Head and Neck Squamous Cell Carcinoma. *Cancer Cell Int* (2021) 21:393. doi: 10.1186/s12935-021-02093-6

60. Mo BY, Li GS, Huang SN, He WY, Xie LY, Wei ZX, et al. The Underlying Molecular Mechanism and Identification of Transcription Factor Markers for Laryngeal Squamous Cell Carcinoma. *Bioengineered* (2021) 12:208–24. doi: 10.1080/21655979.2020.1862527

61. Hsu MK, Wu IC, Cheng CC, Su JL, Hsieh CH, Lin YS, et al. Triple-Layer Dissection of the Lung Adenocarcinoma Transcriptome: Regulation at the Gene, Transcript, and Exon Levels. *Oncotarget* (2015) 6:28755–73. doi: 10.18632/oncotarget.4810

62. Lu M, Xueying Q, Hexiang P, Wenjing G, Hagg S, Weihua C, et al. Genome-Wide Associations Between Alcohol Consumption and Blood DNA Methylation: Evidence From Twin Study. *Epigenomics-Uk* (2021) 13:939–51. doi: 10.2217/epi-2021-0039

63. Kawakita D, Matsuo K. Alcohol and Head and Neck Cancer. *Cancer Metastasis Rev* (2017) 36:425–34. doi: 10.1007/s10555-017-9690-0

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.