Open Access Full Text Article

ORIGINAL RESEARCH

Prognostic Risk Models Using Epithelial Cells Identify β-Sitosterol as a Potential Therapeutic Target Against Esophageal Squamous Cell Carcinoma

Zhenhu Zhang[1,*], Bin Shang[1,*], Xinyu Mao[1], Yamin Shi[2], Guodong Zhang [1], Dong Wang [1]

[1]Department of Thoracic Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, 250021, People's Republic of China; [2]School of Foreign Languages, Shandong University of Finance and Economics, Jinan, 250014, People's Republic of China

*These authors contributed equally to this work

Correspondence: Dong Wang, Department of Thoracic Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, No. 324 Jing Wu Road, Jinan, Shandong, 250021, People's Republic of China, Tel +8615168863605, Email sdphwd@163.com

Background: Esophageal squamous cell carcinoma (ESCC) is an aggressive and fatal malignancy that leads to epithelial cancer. The association between epithelial cell heterogeneity, prognosis, and immune response in this cancer remains uncertain. This study aimed to investigate epithelial cell heterogeneity in ESCC and develop a predictive risk model using the identified cell types.

Methods: Single-cell RNA sequencing (scRNA-seq) and differential ESCC gene data were accessed from the Gene Expression Omnibus. Functional enrichment analysis, inferCNV, cell development trajectories, and intercellular communication were analyzed following epithelial cell characterization. Differentially expressed ESCC (n = 773) and epithelial cell marker genes (n = 3407) were intersected to obtain core genes, and epithelial cell-related prognostic genes were identified. LASSO regression analysis was used to construct a prognostic model. The external dataset GSE53624 was used to further validate the stability of the model. Drug sensitivity predictions, and immune cell infiltration were analyzed. Molecular docking clarified the possible therapeutic role of β-sitosterol in ESCC. Finally, wound healing assay, cell colony, and transwell assay were constructed to detect the effects of the core gene PDLIM2 on ESCC cell proliferation, invasion, and migration.

Results: Eight cell clusters were identified, and epithelial cells were categorized into tumor and paratumor groups. The tumor group possessed more chromosomal variants than the paratumor group. Epithelial cells were associated with multiple cell types and significantly correlated with the Wnt, transforming growth factor, and epidermal growth factor signaling pathways. From 231 intersected genes, five core genes were screened for use in the risk model: CTSL, LAPTM4B, MYO10, NCF2, and PDLIM2. These genes may contribute to the cancerous transformation of normal esophageal epithelial cells and thereby act as biomarkers and potential therapeutic targets in patients with ESCC. β-Sitosterol furthermore displayed excellent docking potential with these genes. Meanwhile, further experiments demonstrated that the gene PDLIM2 plays a major role in the progression of oesophageal squamous carcinoma.

Conclusion: We successfully developed a risk model for the prognosis of ESCC based on epithelial cells that addresses the response of ESCC to immunotherapy and offers novel cancer treatment options.

Keywords: ESCC, epithelial cells, risk signature, β-sitosterol, TME

Introduction

Esophageal squamous cell carcinoma (ESCC) is a highly lethal cancer, with its highest incidence in Asia and Africa.[1] Despite considerable advances in esophageal cancer treatment over the past 10 years, the 5-year survival rate remains below 30%.[2] Surgery can treat such tumors successfully, but more than half of the patients with ESCC are in advanced stages and consequently ineligible for the procedure. Furthermore, the 5-year rates of metastasis and recurrence remain high (more than 40%), even after surgical treatment.[3] Chemotherapy can inhibit ESCC tumor growth and prevent distant

1193

metastasis but introduces unavoidable side effects such as toxicity and drug resistance.[4] Therefore, the search for safe and efficient alternative ESCC treatments remains crucial.

The tumor microenvironment (TME) is closely associated with angiogenesis, apoptosis inhibition, immune system suppression, and surveillance evasion.[5] Numerous studies have demonstrated that intraepithelial neoplasia in normal epithelial cells leads to invasive carcinoma, which, in turn, stimulates ESCC development.[6] However, our understanding of the molecular pathways underlying ESCC pathogenesis in normal epithelial cells is limited. Several previous studies have focused on other cell types in ESCC;[7] however, the relationship between epithelial cell heterogeneity, ESCC prognosis, and immunotherapy response remains to be fully understood. Elucidating the transformation of epithelial cells into precancerous or cancerous cells is vital for understanding the pathogenesis of ESCC to aid in its early detection and treatment.

Bulk RNA sequencing (RNA-seq) can yield substantial information about the expression of tumor immune genes but cannot identify immune cell types, cell–cell interactions, or relationships with cell-specific markers.[8] An emerging tool, single-cell RNA sequencing (scRNA-seq), is frequently used to evaluate the molecular and cellular features of cancer. It can also be used to assess distinct cellular subpopulations.[9] The mechanism of action of distinct cells inside the ESCC TME has been identified in prior research, which also proposed new treatment approaches for ESCC.[10] In this study, we used scRNA-seq technology with the aim of creating transcriptome profiles of various ESCC cell subtypes. We characterized the association between the TME and epithelial cells and the survival of patients with ESCC by thoroughly analyzing epithelial cell, RNA-seq, and clinical data. A new risk model was developed that combined epithelial risk profiles with clinicopathology. The model was validated with external datasets to have a more stable prediction ability. Meanwhile further experiments also proved that the core genes for constructing the model play an important role in the progression of ESCC. In addition, we analysed the potential therapeutic mechanism of β-sitosterol in ESCC, which provides fundamental structural evidence for ESCC drug development.

Materials and Methods
Data Download
ESCC expression profile datasets GSE161533,[11] GSE20347,[12] GSE53624, and GSE45670[13] were downloaded from the official Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/).[14] All four datasets were sourced from Homo sapiens samples. Two datasets, GSE161533 and GSE20347, were used to examine differential gene acquisition. Of these, the GSE161533 dataset comprised 56 ESCC cases and the associated normal tissue samples, whereas the GSE20347 dataset consisted of 34 ESCC cases and their accompanying normal tissue samples; these samples were combined to obtain 90 ESCC and paired non-tumor tissue samples for differential analysis. Gene differences among several populations were explored using the R program limma. To identify differentially expressed genes, we established the criteria |logFC| > 1 and adjusted p < 0.05 and visualized the differences using volcano plots. The GSE45670 and GSE53624 dataset was used for external validation.

The single-cell dataset GSE196756 was also extracted for analysis,[15] containing Homo sapiens samples of tumor tissues from three patients with ESCC. Gene sequences of 93 ESCC patients were downloaded from The Cancer Genome Atlas (TCGA) website (https://portal.gdc.cancer.gov/) and analyzed with the TCGAbiolinks R package, with the expression profiling and survival data (Table S1) being described as transcripts per million.[16]

Cell Type Identification and Differential Gene Expression Analysis
Single-cell sequencing data of the three patients with ESCC from the GSE196756 dataset were filtered using the Seurat R package[17] to obtain cells with >10% mitochondrial genes. Low-quality cells were purified according to quality control criteria (0 < nFeature_RNA < 3000). The default parameters and conventional Seurat software practices were applied to normalize the data, identify highly variable genes, and conduct dimensionality reduction clustering analysis. Data were integrated across samples with the Harmony package. Single-cell populations were named using signature genes from published literature. Epithelial cells were classified into various subtypes according to their marker genes using inferCNV analysis. Differential genes between cell subpopulations were calculated with the FindAllMarkers function (based on the

Wilcoxon test), with $p < 0.05$ and $|log2FC| > 0$ as screening criteria. The extracted marker genes for epithelial cells and differential genes for ESCC were plotted on a Wayne diagram.

Functional Enrichment Analysis

To explore the biological functions of the genes of interest, ClusterProfiler v4.6.0 was used to analyze the core genes using Gene Ontology (GO)[18] and Kyoto Encyclopedia of Genes and Genomes (KEGG).[19] A p-value < 0.05 was considered statistically significant and visualized.

Analysis of Intercellular Communication

We examined intercellular communication between different cell typologies in single-cell transcriptome sequencing data using the CellChat R package[20] Inferring communication networks between different cell typologies were constructed to show the volume, particularly exploring potential connections between epithelial cells and other cell types. Heatmaps were used to demonstrate essential ligands during intercellular signaling. Several pathways commonly associated with cancer—such as the transformation growth factor (TGF)-β, epidermal growth factor (EGF), and Wnt signaling pathways —were further analyzed.

AUCell Score Analysis

To determine how particular genomes were enriched in cells, we evaluated genome-specific activity in single-cell RNA sequencing data with the AUCell R package.[21] Three cancer-related pathways were validated using the Reactome pathway database (https://reactome.org).

Pseudo-Time Trajectories Analysis

Cellular subpopulations were analyzed using the Monocle v2.26.0 R package[22] in the proposed time–series analysis, and highly discrete genes were downscaled using the DDRTree method. Based on clustering characterization, we obtained a trajectory map of the time of differentiation at which the cell subpopulations were located in the single-cell dataset. The top 50 differential genes in ESCC were analyzed.

Prognostic Modeling

Using the data of 93 patients with ESCC from the TCGA–ESCC dataset, we conducted univariate Cox regression analysis to identify potential genes with p-values < 0.05. This allowed us to thoroughly study the effects of crucial epithelial cell genes on the prognosis of ESCC. Next, we used external datasets from the GSE45670 and Gene Expression Profiling Interactive Analysis (GEPIA) databases to validate these potential genes using LASSO regression analysis. A risk models was constructed from the screened genes, with risk scores being calculated using the following equation: −0.240* LAPTM4 −0.211* MYO10+ 0.247* CTSL+ 0.259* NCF2 −0.390* PDLIM2. Next, we calculated the area under the curve (AUC) of the subject operating characteristic curve (ROC) to measure the performance of the model and evaluated the clinical value of the model by decision curve analysis (DCA). Column plots, calibration curves, ROC curves, and DCA curves were plotted using R software (ROC was performed using the "survival" package. column plots and calibration plots were plotted using "rms"). DCA was performed using the function "stdca.R". p-values < 0.05 were considered statistically significant). Meanwhile, the external dataset GSE53624 was used for further validation.

Immunoreactivity and Immune Infiltration Analysis

We used the TCGA–ESCC dataset to evaluate the tumor immune response with the ESTIMATE package in R.[23] By examining the transcriptional patterns of cancer samples, ESTIMATE focused on estimating the composition of tumor cells and the presence of stromal and immune cells that have invaded the tumor. Specific immune cell infiltration results were calculated using the R package CIBERSORTx[24,25] and further validated via single-sample Gene Set Enrichment Analysis (ssGSEA). Immunotherapy plays a vital role in oncology, therefore we analyzed several joint immune checkpoints.

Drug Sensitivity Analysis

The Genomics of Drug Sensitivity in Cancer (GDSC) database[26] was used to explore therapeutic and mutational cancers. It includes details on the molecular, cellular, and pharmacological sensitivity and responses to various tumor cells. The pRRophetic package[27] —which downloads cell line mutation data and IC50 values for several anticancer medications— was used to analyze the correlations between patients with high- and low-risk scores and various anticancer drugs.

Molecular Docking of β-Sitosterol to Core Targets

A molecular docking approach was used to study the relationship between β-sitosterol and its main targets. Protein Data Bank (PDB; http://www.rcsb.org/pdb/home/home.do) was used to collect structural data of the key core genes and PubChem (https://pubchem.ncbi.nlm.nih.gov) was used to obtain chemical data of β-sitosterol. To simulate how β-sitosterol binds to the target protein, PyMol software was applied to dehydrate, hydrogenate, and charge. AutoDockTools-1.5.6 software was used to determine the location and size of the Grid box, the binding site of the disease target protein to the ligand. Autodock vina V1.1.2 software was used to complete the molecular docking and assess the affinity of the compounds to the target proteins. PyMOL software was used for visualization and numerical heat map analysis.

Cell Culture

The esophageal squamous cell carcinoma cell lines KYSE150, KYSE30, KYSE510, Eca9706, and the normal esophageal epithelial cell line (HEEC) were procured from the Cell Bank of the Shanghai Institutes for Biological Sciences. The cell lines KYSE150, KYSE30, KYSE510, and Eca9706 were cultured in RPMI-1640 medium (Kibbutz Beit HaEmek BI) supplemented with 10% fetal bovine serum (FBS) at 37°C in a humidified incubator with 5% CO2. The HEEC cells were maintained under the same conditions in DMEM medium (Gibco) also supplemented with 10% FBS. All cell lines were passaged no more than 15 times. This study was approved by the Ethics Committee of the Affiliated Provincial Hospital of Shandong First Medical University (Ethics approval number: NSFC: NO.2024–619).

Lentiviral Transfection

PDLIM2 overexpression lentiviral vectors and negative control lentiviral vectors were purchased from GenePharma. The lentiviral vectors were transduced into KYSE150 and Eca9706 cells, and cells were selected with a medium containing 5 µg/mL puromycin to obtain stable transfectants.

RNA Extraction and RT-qPCR

Total RNA was extracted from cells using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Actin was used as the endogenous reference gene. The PCR primer sequences used in this study were as follows: PDLIM2: CGTTGACGGTGGATGTGG (F) and GCTCTGGCGGATCTTGCT (R). The relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method.

Cell Invasion and Migration Assays

Cell invasion was assessed using matrigel-coated polycarbonate membrane inserts (Corning), and cell migration was evaluated using uncoated inserts. Cells transduced with lentiviral vectors were seeded into the upper chamber of the transwell, and medium containing 20% FBS was added to the lower chamber. Tumor cells were seeded at a density of 2 × 10^5 cells per chamber. Filters were wiped with cotton swabs, and cells adhered to the bottom were fixed with 4% paraformaldehyde phosphate (Beyotime, Shanghai, China). Cells were stained with crystal violet (Beyotime). Experiments were independently repeated three times.

Wound Healing Assay

KYSE150 and Eca9706 cells were seeded into six-well culture plates, and a wound was carefully created using the tip of a 200 µL pipette. Cells were observed at 0 hours and 24 hours post-wounding using an inverted microscope. The wound area at the same location was measured using ImageJ software.

Colony Formation Assay

500 cells per well of six-well plates were used to harvest and seed the cells, which were subsequently incubated for ten days at 37 °C with 5% CO2. After 30 minutes of staining with 0.1% crystal violet, colonies were fixed with 4% paraformaldehyde.

Western Blot Analysis

Proteins were extracted, quantified, and separated from KYSE-150 and Eca9706 cells, then transferred to a PVDF membrane. The membrane was blocked with skim milk for 2 hours, incubated with the primary antibody against PDIM2 (1:2000) at 4 °C overnight, developed with ECL substrate, and photographed using a gel imaging system. The relative expression of proteins in each group was analyzed.

Statistical Analysis

R software v4.2.0 was used for data analysis. The Wilcoxon rank-sum test was used to examine differences between variables with non-normal distributions. The R package timeROC v0.4 was used to plot ROC curves and calculate area under the ROC curves (AUCs) to evaluate the accuracy of the risk scores in determining prognosis. Univariate and multivariate Cox analyses were used to identify independent prognostic factors, with $p < 0.05$ being considered statistically significant.

Results

Single-Cell Data Analysis

Single-cell clustering was visualized via t-distributed stochastic neighbor embedding (t-SNE), as depicted in Figure 1A. By querying the literature used to annotate these subclusters (Figure 1B), eight cell types were identified, namely: T cells (cluster 0,1,2,9), B cells (cluster 3,11,13,15), myeloid cells (cluster 4), fibroblasts (cluster 5,17), epithelial cells (cluster 6,7,12,18), monocytes (cluster 8,14,10), macrophages (cluster 10), and endothelial cells (cluster 16) (Figure 1C).

Cellular Communication Associated with Epithelial Cells

Communication network diagrams were constructed to show the volume (Figure S1A) and intensity (Figure S1B) of contact between various cell types. Epithelial, fibroblast, and endothelial cells displayed strong intercellular communication. To investigate the connection between epithelial cells and various cell subtypes, their interactions were visualized with bubble diagrams that indicated mainly involvement of the Wnt signaling pathway (FZD6 + LRP6), the TGF-β signaling pathway (TGFBR1 + TGFBR2), and other joint and cancer-related pathways (Figure S1C). We mapped the cell-to-cell interactions of the Wnt, TGF, and EGF signaling pathways (Figure 2A–C) and constructed correlation heatmaps (Figure 2D–F). Epithelial cells occupied a central position, suggesting that they play a critical role in ESCC progression. Using bubble diagrams, we further documented the receptor ligands involved in the aforementioned three
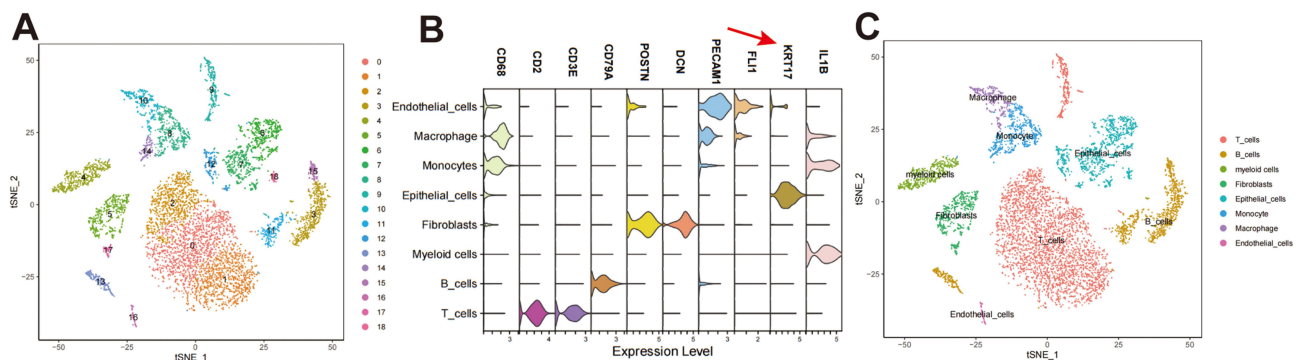

Figure 1 Analysis of single-cell sequencing data. (**A**) Reduced dimensional cluster analysis, t-distributed stochastic neighbor embedding (t-SNE) distribution of different cell clusters. (**B**) A violin diagram showing the marker genes for each cell type. (**C**) Distribution of t-SNE in different cell types. The red arrow indicates the reference gene of epithelial cells.
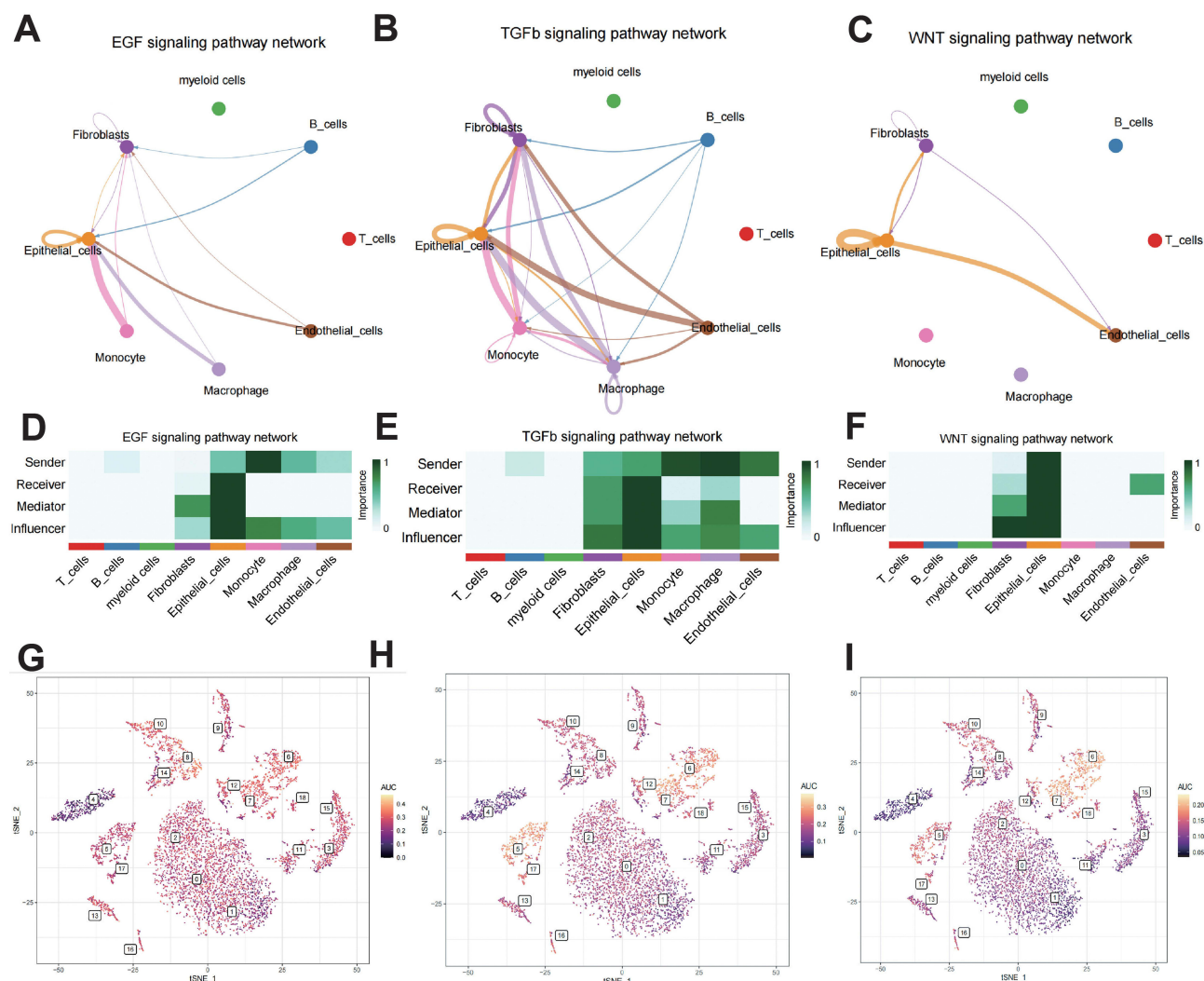
Figure 2 Analysis of epithelial cell signaling pathways. (**A–C**) Network diagram depicting the interaction of the Wnt, transformation growth factor (TGF), and epidermal growth factor (EGF) signaling pathways. (**D–F**) Heatmap of the interaction of the Wnt, TGF, and EGF signaling pathways. (**G–I**) AUCell analysis based on the Reactome pathway database.

pathways (Figure S2A–C). To verify the relevance of epithelial cells in these pathways, the pathway activity of specific gene sets of epithelial cells was analyzed using the R package AUCell (Figure 2G–I). Epithelial cells exhibited the highest activity in all three cancer-related signaling pathways, followed by fibroblasts, which further confirms that epithelial cells play an essential role in ESCC.

Access to Core Genes

After combining the GSE161533 and GSE20347 datasets (Figure 3A) and plotting a volcano map (Figure 3B), we identified 773 differentially expressed genes, of which 358 were upregulated and 415 were downregulated. Based on the single-cell dataset, we analyzed the differences in differentially expressed genes between all cell types, as visualized with volcano plots (Figure 3C). Wayne plots of the intersections between differentially expressed genes in ESCC and epithelial cells identified 231 core genes (Figure 3D).

Enrichment Analysis

The results of the GO enrichment analysis were depicted with chordal plots (Figure S3A). Co-differentially expressed genes were mainly enriched in epidermal development and microtubule cytoskeleton organization involved in mitosis
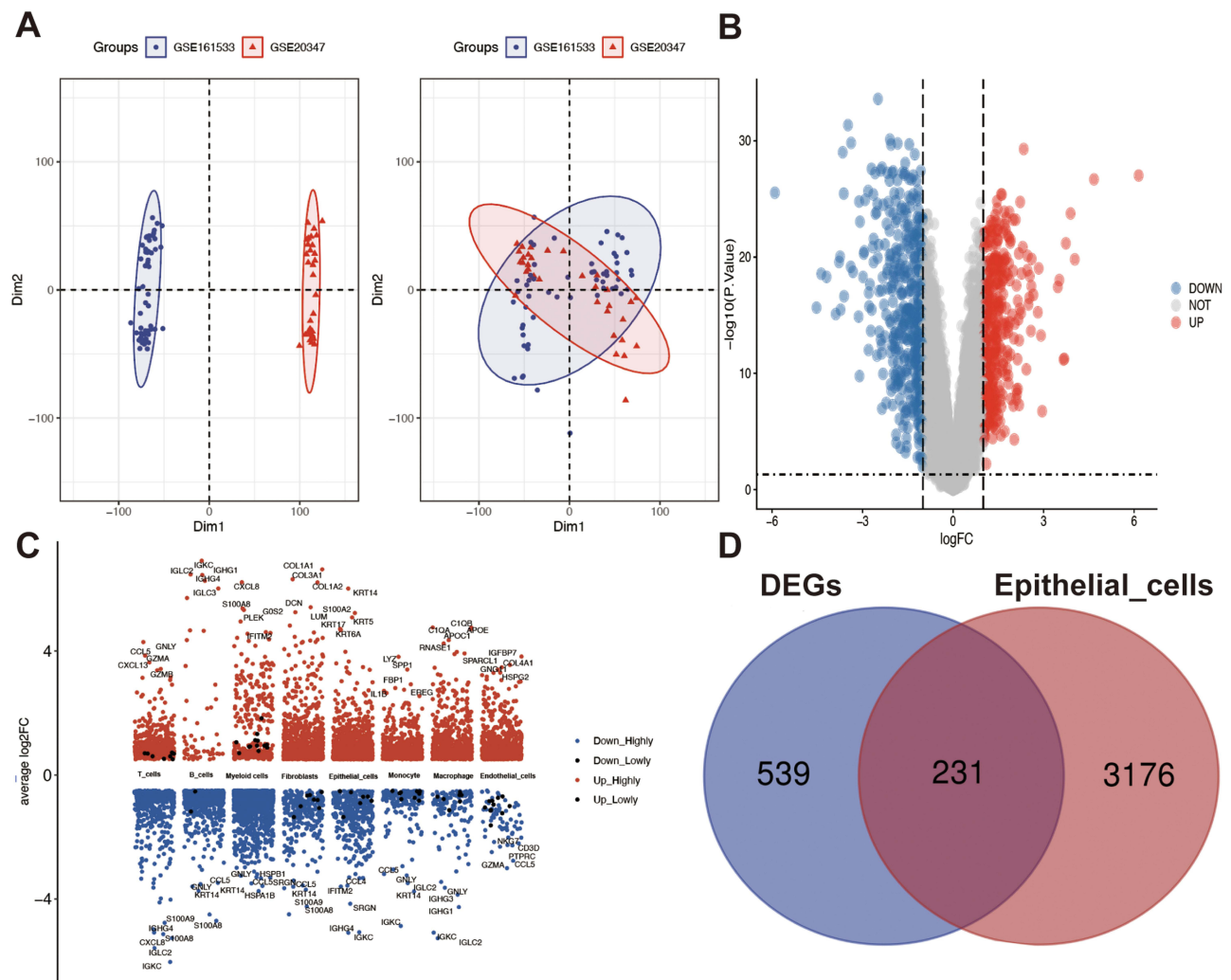
Figure 3 Acquisition of differentially expressed genes. (**A**) Principal component analysis plot of the GSE161533 and GSE20347 dataset concatenation. (**B**) Volcano map of merged differentially expressed genes. (**C**) Uniform Manifold Approximation and Projection plot showing differentially expressed genes of different cell types. (**D**) Wayne diagram of differentially expressed genes and epithelial cells in ESCC.

(biological processes); and extracellular matrix structural constituents and CXCR chemokine receptor binding (molecular functions. These effects are primarily related to cellular functions and immunity. KEGG results indicated gene enrichment related mainly to the IL-17 signaling pathway, transcriptional misregulation in cancer, and Toll-like receptor signaling pathway, which were focused on inflammation, immunity, and other related ways; these results are depicted as bar graphs in Figure S3B.

Results of Pseudo-Time Trajectories Analysis

Figure 4 presents timing diagrams that are colored according to the cell population (Figure S4A), cell state (Figure S4B), and pseudo-time course (Figure S4C). As time progressed, the number of epithelial cells and fibroblasts gradually decreased, and B cells differentiated mainly into T cells and monocytes. The decrease in monocytes may be due to their gradual transformation into macrophages. The core genes were placed alongside 231 intersecting targets in the STRING database to acquire the protein–protein interaction network, and the cytoHubba plugin was used to predict the critical protein nodes and subnetworks. The top 50 core genes were then selected using the Degree parameter (Figure S4D), and a heatmap was constructed that displayed the expression of the top 50 genes with pseudotemporal variations (Figure S4E) as based on pseudotemporal trends. These genes exhibited different trends, for example, the expression of genes such as CKS2, CENPM, ATAD2, and MCM2 varied from high to low, and those of genes such as CXCL8 and CCL3 varied from low to high.
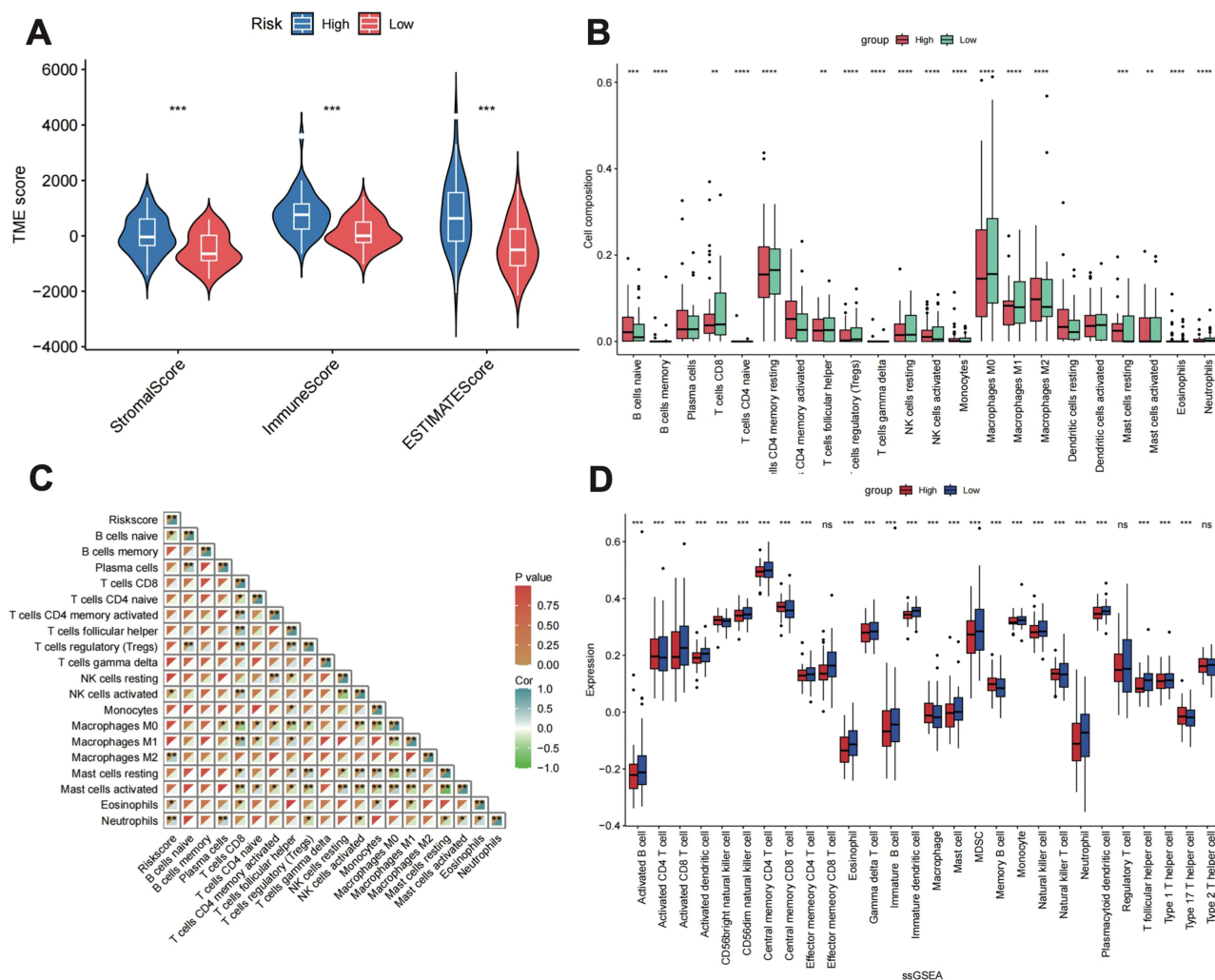
Figure 4 Immuno-infiltration analysis. (**A**) ESTIMAT analysis of the high- and low-risk group. (**B**) CIBERSORTx analysis of the abundance of immune cells. (**C**) Heatmap of correlation between risk score and the presence of immune cell types. (**D**) The single-sample Gene Set Enrichment Analysis scores of 28 immune cells in the high- and low-risk groups. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
Abbreviation: ns, Not Statistically Significant.

Prognostic Model Construction and Validation

When using the data from 93 patients with TCGA-ESCC to perform univariate Cox regression analysis of 321 core-related genes, we discovered that 11 of the core genes (LAPTM4B, TPX2, MYO10, TRIP13, PBK, CTSL, DLGAP5, MRGBP, MFAP2, NCF2, and PDLIM2) had a significant impact on the survival of patients with ESCC (p < 0.05) (Figure 5A). We used LASSO regression analysis to construct the fitted model (Figure 5B) and cross-validation analysis to determine an optimal lambda value of 5 (Figure 5C and Table 1). Based on the external dataset GSE45670, we further validated the five screened modeling genes (Figure 5D–H) and found that the expression of LAPTM4B, MYO10, NCF2, and PDLIM2 differed significantly different between the standard and control groups. Using the GEPIA dataset, we further validated five core genes (Figure 5I–M) that similarly showed significantly different expressions.

Risk scores were calculated using Equation. Additionally, the patients were divided into high- and low-risk groups based on the median value. The assumption was made that the age, sex, and tumor stage of patients with ESCC influence their lifespan. Heatmaps were constructed that related the calculated risk scores to clinical patient information (Figure S5A), with respective maps of sex, tumor stage, T stage, and age for the high- and low-risk groups (Figure S5B–E). Based on these findings, a multifactorial prognostic column chart was created (Figure 6A). Then, in the training group (TCGA), we plotted ROC curves and calibration curves to evaluate the stability of the nomogram (Figure 6B) and further validated it in an
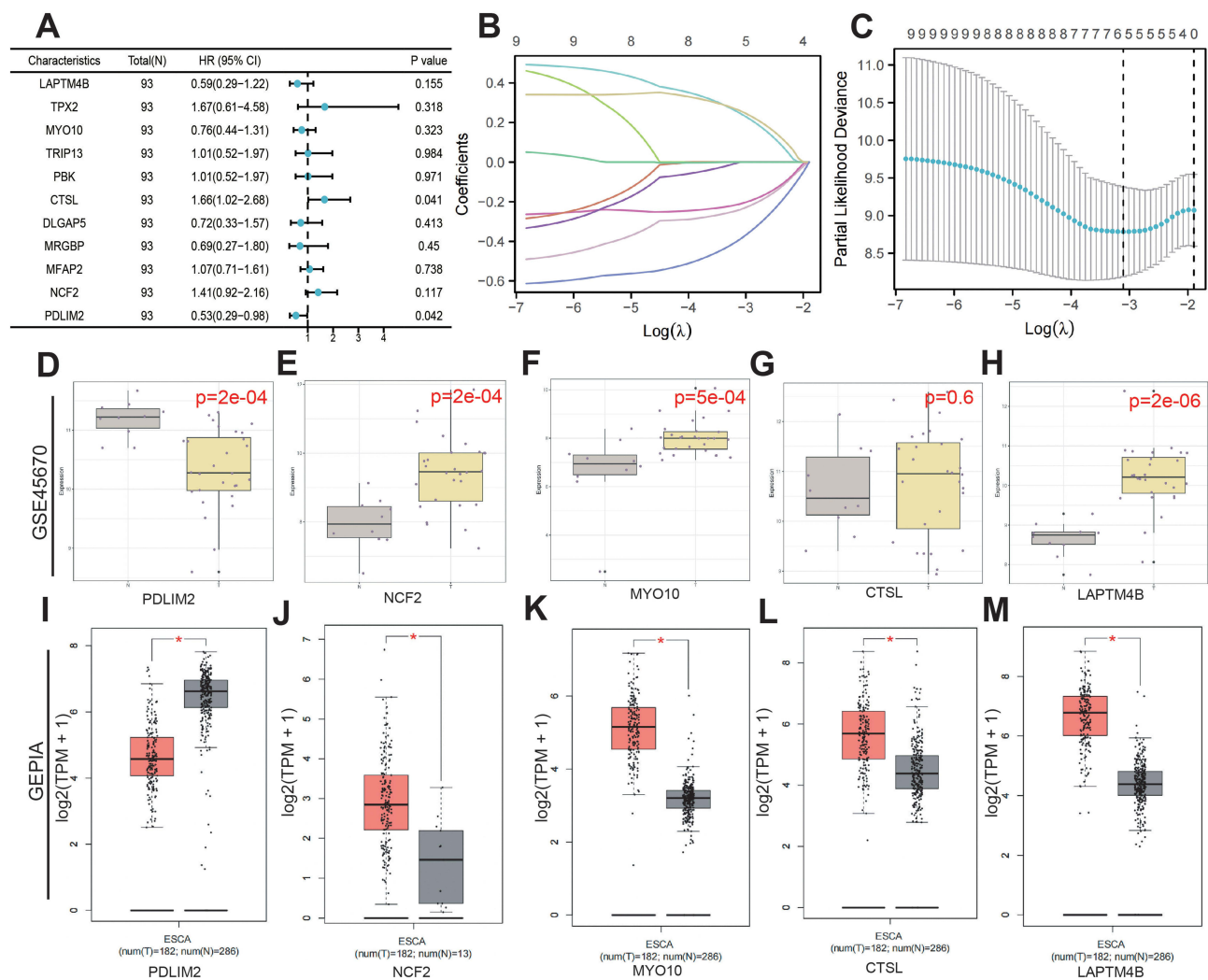
Figure 5 Screening and validation of core prognostic genes. (**A**) Prognostic forest map. (**B**) LASSO analysis to construct a fitted model. (**C**) Determining the best lambda value for the fitted model: the x-axis represents the logarithmized lambda value, the y-axis represents the error of the model, and the dotted line on the left represents the lambda value that minimizes the error and the number of features screened. (**D–H**) Dataset GSE45670 validation. (**I–M**) Validation of the GEPIA dataset. *p < 0.05.
Abbreviations: N, Total sample size; HR, Hazard Ratio.

external dataset (GSE53624) (Figure 6C), and the results showed that the predicted nomograms had good calibration performance. Meanwhile, DCA showed that the prognostic column line plot had strong clinical utility (Figure S6A and B).

Analysis of Transcriptome Heterogeneity in ESCC Epithelial Cells

Epithelial cells were re-annotated into 10 clusters (Figure 7A), and marker genes of epithelial tumors (EPCAM, MDK) and paratumors (IGLC2, FABP5) were identified in the literature.[28] Bubble plots were constructed to demonstrate the expression of these marker genes among the 10 clusters (Figure 7B), and epithelial cells were accordingly categorized into tumor (clusters 0, 1, 2, 3, 4, 5, 7, and 9) and paratumor (clusters 6, 8) groups for visualization (Figure 7C). We validated this finding using inferCNV analysis. The tumor group had more chromosomal variants than the paratumor group, with more amplifications on chromosomes 1 and 9 and more deletions on chromosomes 3 and 5 (Figure 7D). Next, we analyzed the expression of four marker genes (EPCAM, MDK, IGLC2, and FABP5) and five modeling genes (CTSL, NCF2, PDLIM2, LAPTM4B, and MYO10) using violin plots (Figure 7E).

Table 1 Detailed Information on LAPTM4B, CTSL, MYO10, NCF2, and PDLIM2

Gene Symbol	Gene ID	Full Name	Function
LAPTM4B	7805	Lysosomal protein transmembrane 4 beta	Involved in vesicle trafficking, lysosome maturation and fusion, as well as apoptosis and tumor development. LAPTM4B is highly expressed in a wide range of human tumors and correlates with tumor invasion and metastasis, treatment resistance, and other characteristics.
MYO10	4646	Myosin X	MYO10 plays an important role in the process of cell movement and cell architecture. Its main functions involve directional cell movement, cell adhesion and migration.
CTSL	1514	Cathepsin L	CTSL is an important lysosomal enzyme that performs multiple functions within the cell. It is involved in intracellular protein degradation, lysosome formation and functional maintenance. CTSL plays an important regulatory role in physiological processes such as immune response, apoptosis, embryonic development, tissue remodeling and inflammation.
NCF2	4759	Neutrophil cytosolic factor 2	It is a multisubunit enzyme complex involved in electron transfer during intracellular reactions to produce reactive oxygen species. The main function of NCF2 is to provide electrons for NOX and to promote intracellular redox reactions.
PDLIM2	10,350	PDZ and LIM domain 2	PDLIM2 protein is a binding protein that interacts with other proteins and forms complexes involved in the regulation of multiple cellular processes and signaling pathways. PDLIM2 plays an important role in physiological processes such as cell development, differentiation and proliferation.

Immune Infiltration and Immune Checkpoint Analysis

We investigated the connection between the risk score and TME characteristics using ESTIMAT software. The findings revealed a positive correlation between immune and risk scores, with the immune score being higher in the high-risk than the low-risk group. Additionally, the risk and ESTIMAT scores were positively correlated, with risk score being more predominant in the high-risk than the low-risk group (Figure 4A). Cumulatively, these findings confirm a strong correlation between the risk score and TME.
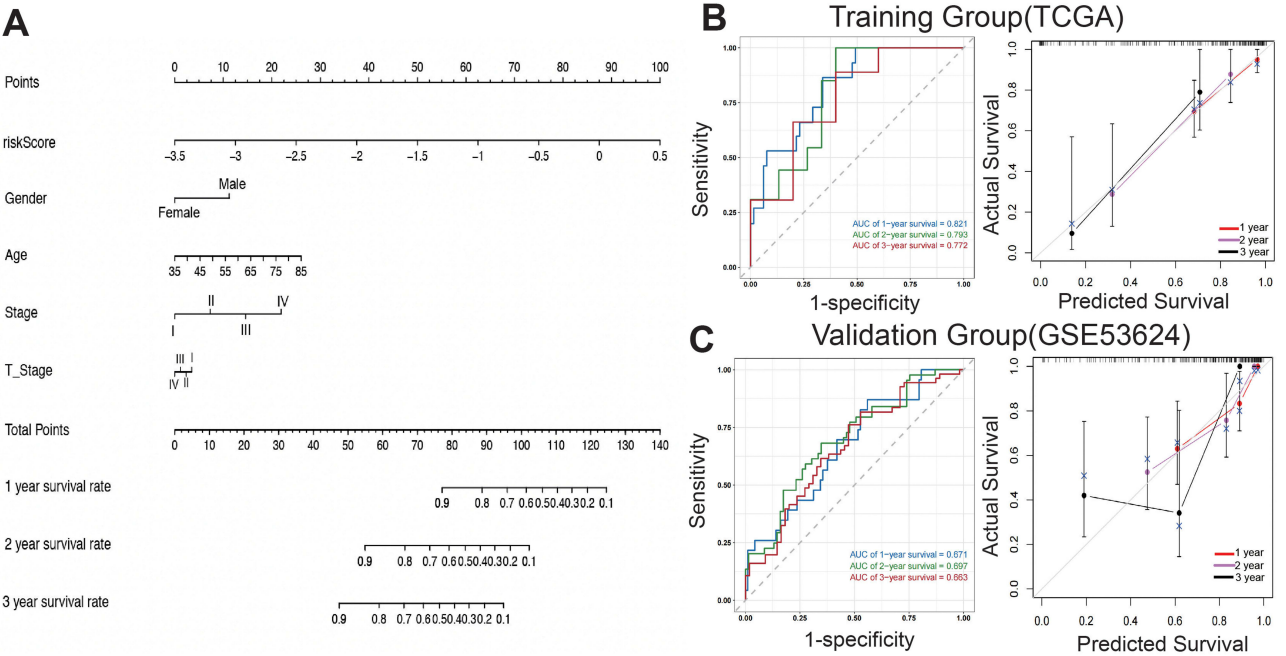

Figure 6 Nomogram construction of related genes. (**A**) Nomogram construction. (**B**) ROC curves and calibration curves for the nomograms in the training group. (**C**) ROC curves and calibration curves for the nomograms in the validation group.
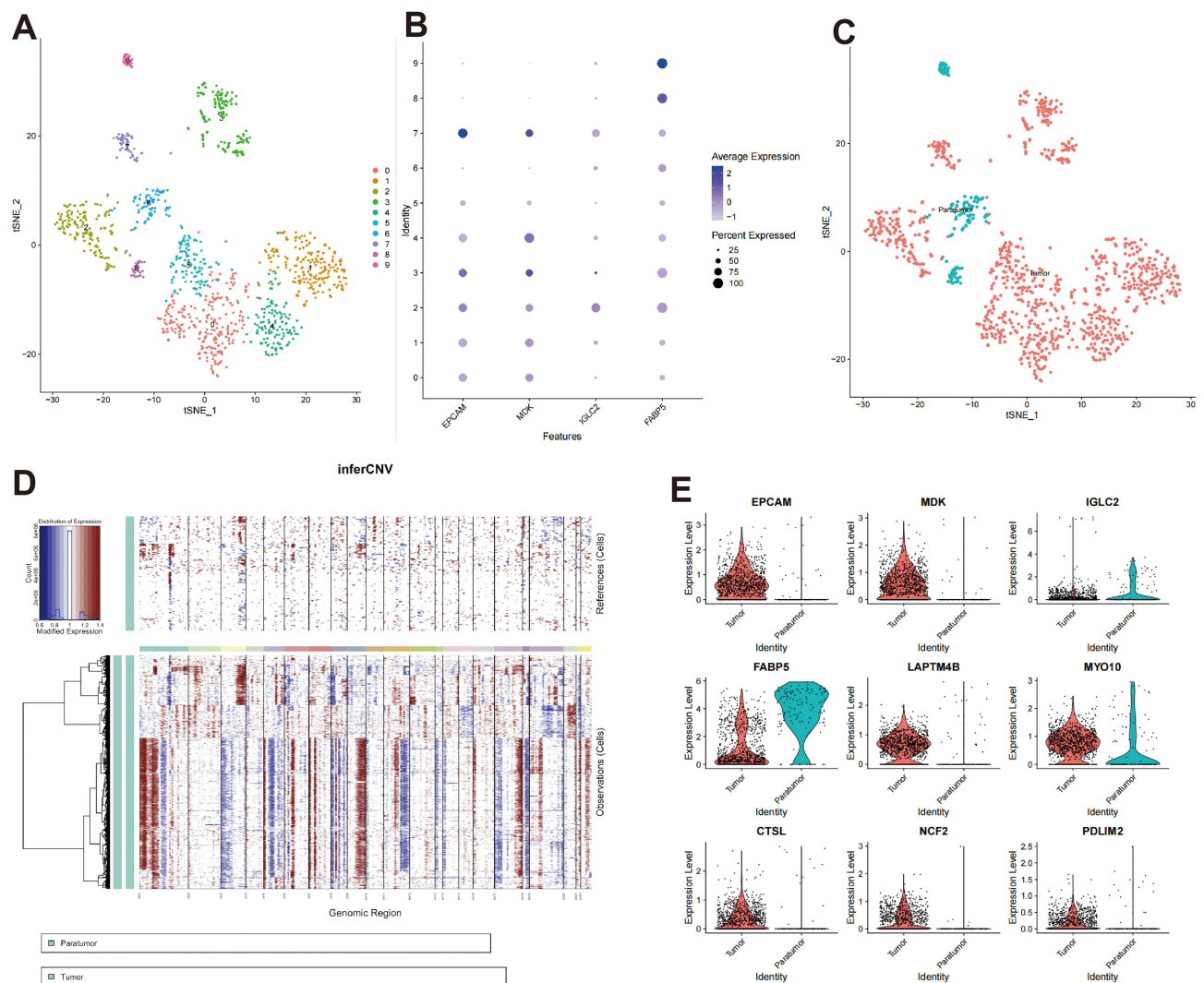
Figure 7 Epithelial cell analysis. (**A**) The t-distributed stochastic neighbor embedding (t-SNE) of 10 cell clusters. (**B**) Dot plot of different marker genes. (**C**) Identification of tumor and paratumor marker genes; (**D**) Copy number variation profile showing relative differences between the tumor and paratumor groups. (**E**) Violin plot of essential genes.

To assess the immune infiltration of high- and low-risk patients in the TCGA-ESCC dataset (Figure 4B), we used the CIBERSORTx approach. The abundance of most immune cell types differed significantly among the subgroups. Naive B cells, CD4+ T cells, CD8+ T cells, monocytes, macrophages, eosinophils. A heatmap of the correlation between risk score and immune cell type (Figure 4C) revealed significant positive correlations between activated CD4 T cells and resting natural killer (NK) cells and between monocytes and follicular helper T cells ($r > 2$, $p < 0.05$). Activated mast cells showed a significant negative correlation with CD8 T cells, activated NK cells, and M0 macrophages ($r > 2$, $p < 0.05$). Risk score was significantly associated with naive B cells ($r = -0.26$, $p < 0.05$), M2 macrophages ($r = 0.33$, $p < 0.05$), eosinophils ($r = 0.21$, $p < 0.05$), and neutrophils ($r = 0.27$, $p < 0.05$). These results were further validated using ssGSEA (Figure 4D).

Immune checkpoint inhibitors are often used in tumor immunotherapy. In this study, we examined the relationship between the risk scores and 21 standard immunological checkpoints (Figure S7A). Risk score showed a strong positive association with each immune checkpoint ($p < 0.05$). We also analyzed the correlation between risk score and programmed death-ligand 1 (PD-L1) (Figure S7B), which was strong and positive ($r = 0.362$, $p < 0.001$), suggesting that immunotherapy targeting PD-L1 may have a higher sensitivity in high-risk patients. We then downloaded the MSI data of patients with ESCC, which showed a significant negative correlation with risk score ($r = -0.255$, $p = 0.014$) (Figure S7C). Based on these findings, we can hypothesize that risk score may be related to the degree of immune cell infiltration into the immunological microenvironment, which may affect ESCC treatment.

Drug Sensitivity Analysis

Patients with ESCC can display different therapeutic responses to the same medication because of the diverse characteristics of this cancer. Using the GDSC v2 database and R tool OncoPredict, we determined the IC50 values of several chemotherapeutic agents and small-molecule anticancer medications. These values differed significantly between patients with high and low wind (Spearman correlation analysis, $p < 0.01$). The top eight drugs selected for demonstration were AZD8186 (Figure 8A), entospletinib (Figure 8B), gemcitabine (Figure 8C), GSK2606414 (Figure 8D), mitoxantrone (Figure 8E), PRT062607 (Figure 8F), SB505124 (Figure 8G), and UMI-77 (Figure 8H). Detailed information on these drugs is provided in Table 2.

Molecular Docking of β-Sitosterol to Core Genes

We molecularly docked β-sitosterol with each of the five core modeling genes, namely, CTSL (Figure 9A), LAPTM4B (Figure 9B), MYO10 (Figure 9C), NCF2 (Figure 9D), and PDLIM2 (Figure 9E), with all genes linking with β-sitosterol via hydrogen bonds. The docking results were presented in a heatmap (Figure 9F), which showed that the docking energies of β-sitosterol with CTSL, LAPTM4B, MYO10, NCF2, and PDLIM2 were –7.13 kcal/mol, –4.81 kcal/mol, –4.71 kcal/mol, –6.36 kcal/mol, and –6.35 kcal/mol. This indicates that β-sitosterol could bind stably to the genes, implying that β-sitosterol may exert a potential therapeutic effect on ESCC. To further validate the effect of core genes on survival of ESCC patients, we divided ESCC patients in the TCGA database into high- and low-expression groups. Kaplan-Meier survival analysis showed that CTSL, MYO10, NCF2, and PDLIM2 were all significant (Figure 9G), but in the external dataset GSE53624, only PDLIM2 has stability (Figure 9H). Therefore, we judged that the PDLIM2 gene plays an important role in the development of ESCC patients.

PDLIM2 is Underexpressed in ESCC and Inhibits Cancer Cell Proliferation and Migration

We assessed the expression levels of PDLIM2 in the HEEC, KYSE510, KYSE150, KYSE30, and Eca9706 cell lines using the qRT-PCR method. Compared to HEEC, the expression of PDLIM2 was significantly reduced in the KYSE30,
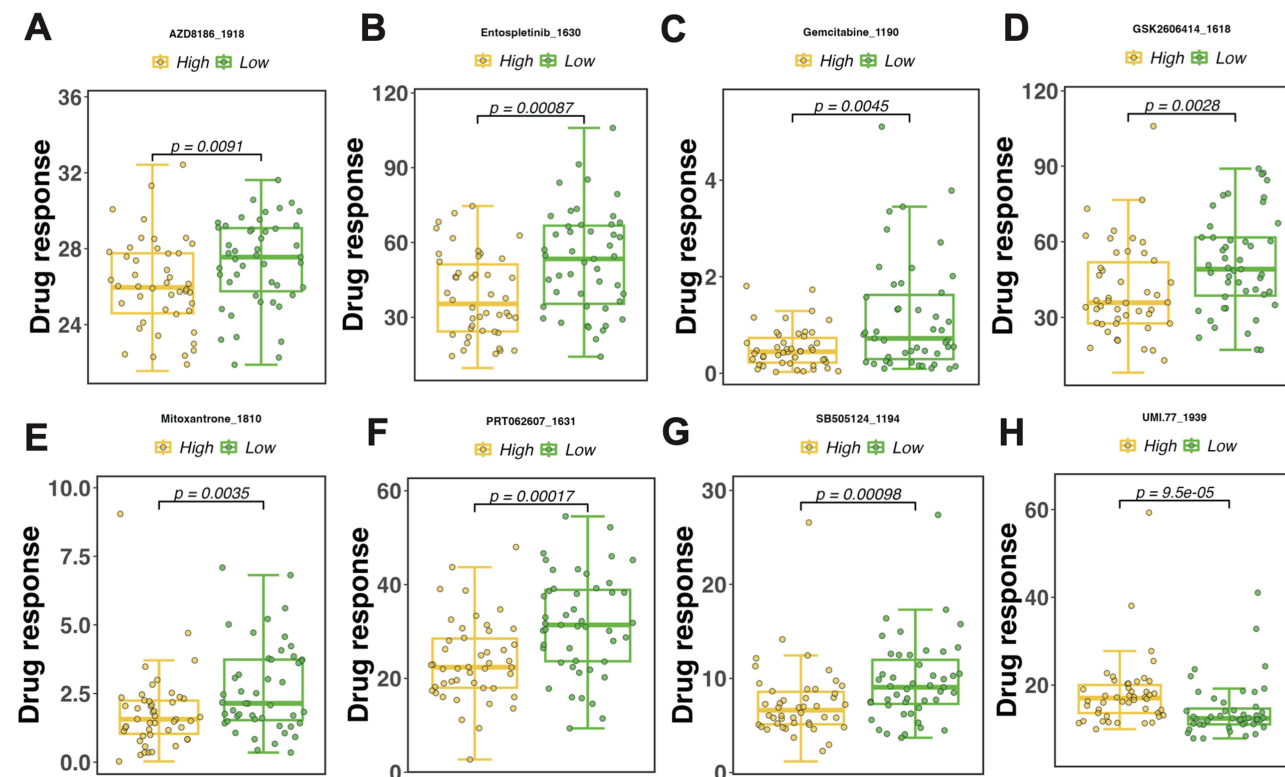

Figure 8 Drug sensitivity analysis. (**A–H**) Top 8 drugs showing the most sensitivity with Risk score. High/Low indicates the grouping of samples into high and low categories. p represents the result of statistical analysis, with p<0.05 indicating significance.

Table 2 Detailed Information on the Top 8 Sensitive Drugs

Drug Name	Drug Target	Target Pathway	Introduction
AZD8186	PI3Kalpha, PI3Kbeta	PI3K/MTOR signaling	AZD8186 interferes with intracellular signaling pathways involved in processes such as cell growth, differentiation, and apoptosis by inhibiting PI3K kinase.
Entospletinib	SYK	Other, kinases	Entospletinib slows the growth and proliferation of cancer cells by inhibiting signaling pathways by targeting BTK. BTK is a tyrosine kinase involved in B-cell receptor signaling and other inflammation-related pathways.
Gemcitabine	Pyrimidine antimetabolite	DNA replication	It is a nucleoside analog that inhibits the proliferation and division of cancer cells by interfering with their DNA and RNA synthesis.
GSK2606414	PERK	Metabolism	GSK2606414 interferes with the proliferation and growth of cancer cells by inhibiting the mTOR signaling pathway. mTOR is a key intracellular signaling pathway involved in cell growth, division, metabolism and survival.
Mitoxantrone	TOP2	DNA replication	The mechanism of action of Mitoxantrone mainly involves DNA cross-linking and damage. It blocks the division and proliferation of cancer cells by interacting with DNA and interfering with the stability and replication process of DNA strands.
PRT062607	SYK	Other, kinases	PRT062607 blocks signaling in these diseases by inhibiting BTK activity, reducing the proliferation and survival of pathological B cells.
SB505124	TGFBR1, ACVR1B, ACVR1C	RTK signaling	It is a selective transforming growth factor beta (TGF-β) receptor kinase (ALK4/5/7) inhibitor, which means that it interferes with the normal function of the TGF-β signaling pathway
UMI-77	MCL1	Apoptosis regulation	UMI-77 acts as a dopamine D2S receptor agonist that mimics the effects of dopamine and selectively activates D2S receptors.

KYSE150, Eca9706, and KYSE510 cell lines, with the most notable reductions observed in the KYSE150 and Eca9706 cell lines (Figure 10A). Therefore, we selected these two cell lines for subsequent experiments. The overexpression efficiency of PDLIM2 in KYSE150 and Eca9706 cells was detected by qRT-PCR and validated by WB (Figure 10B and C), yielding results with significant differences. Subsequently, we further analyzed the function of the gene PDLIM2 in ESCC through in vitro assays. Wound healing assays demonstrated that overexpression of PDLIM2 significantly reduced the healing rate of the KYSE150 and Eca9706 cell lines (Figure 10D). In subsequent migration and invasion assays, we observed that, compared to the control group, overexpressed PDLIM2 significantly inhibited the migration and invasion of ESCC (Figure 10E). Additionally, colony formation assays indicated that overexpression of PDLIM2 reduced the proliferative capacity of the KYSE150 and Eca9706 cells (Figure 10F). These results suggest that PDLIM2 plays a significant role in the proliferation and invasion of ESCC, thereby greatly enhancing the persuasiveness of our model.

Discussion

ESCC is one of the deadliest cancers and the most common histological subtype of esophageal cancer in China.[29] Although treatment modalities have come a long way, ESCC continues to metastasize later in life because of recurrence, metastasis, and nonresponsiveness to adjuvant therapy, resulting in poor 5-year survival rates.[30] Numerous studies have shown that ESCC develops when normal epithelial cells transform into intraepithelial neoplasia and finally into an invasive carcinoma.[6] Malignant epithelial cells are also the primary cell type involved in tumor metastasis, and the aberrant activation of epithelial–mesenchymal transition (EMT) is critical for cancer progression and metastasis.[31] Therefore, deciphering the molecular mechanisms underlying ESCC metastasis and establishing how normal epithelial cells are transformed into precancerous or cancerous cells are essential for improving ESCC treatment and preventing tumor metastasis.
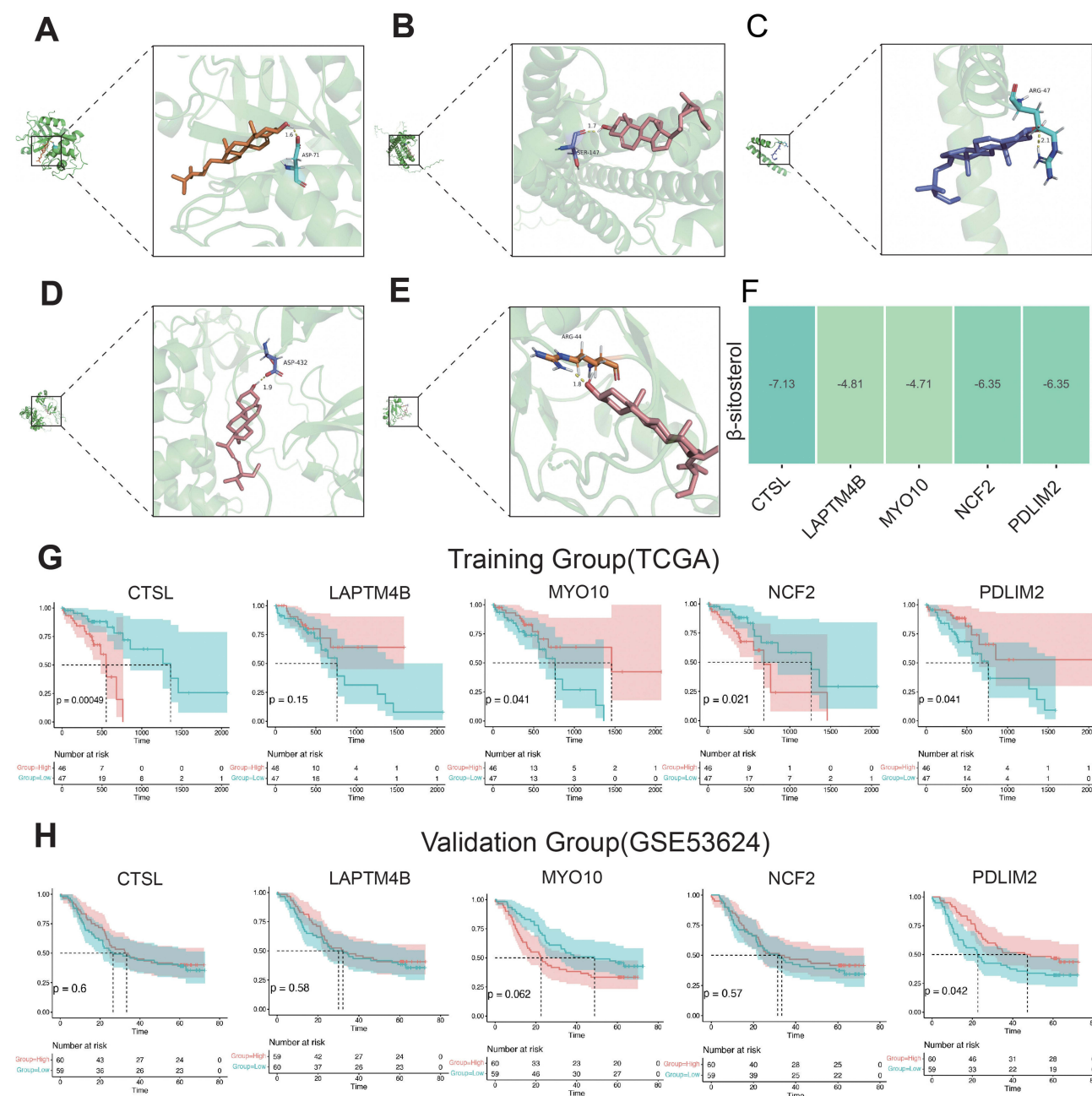
Figure 9 Molecular docking of β-sitosterol with five core genes and the mRNA levels were measured by real-time fluorescence quantitative PCR. (**A-E**) Molecular docking of β-sitosterol to five core genes. (**F**) The docking results were presented in a heatmap. (**G**) K-M survival analysis of the five core genes in the TCGA database. (**H**) Further validation of the core genes in the external dataset GSE53624.

The heterogeneity of ESCC was examined using the single-cell dataset GSE196756. After annotation, eight distinct cell types were identified: T cells, B cells, myeloid cells, fibroblasts, epithelial cells, monocytes, macrophages, and endothelial cells. Based on the labeling of the epithelial cell genes, these cell types were categorized into tumor and paratumor subgroups, and inferCNV analysis revealed that the tumor group possessed more variants. An analysis of intercellular communication revealed a substantial link between fibroblasts and epithelial cells. Moreover, epithelial cells were strongly correlated with cancer-related pathways such as the Wnt, TGF, and EGF signaling pathways— followed by fibroblast cells. Precancerous and cancerous epithelial cells have been demonstrated to express less ANXA 1, which interacts with formyl peptide receptor 2 (FPR2) on fibroblasts to control and maintain normal fibroblast homeostasis in normal esophageal epithelial cells; diminished ANXA 1-FPR 2 signaling favors the conversion of normal fibroblasts to
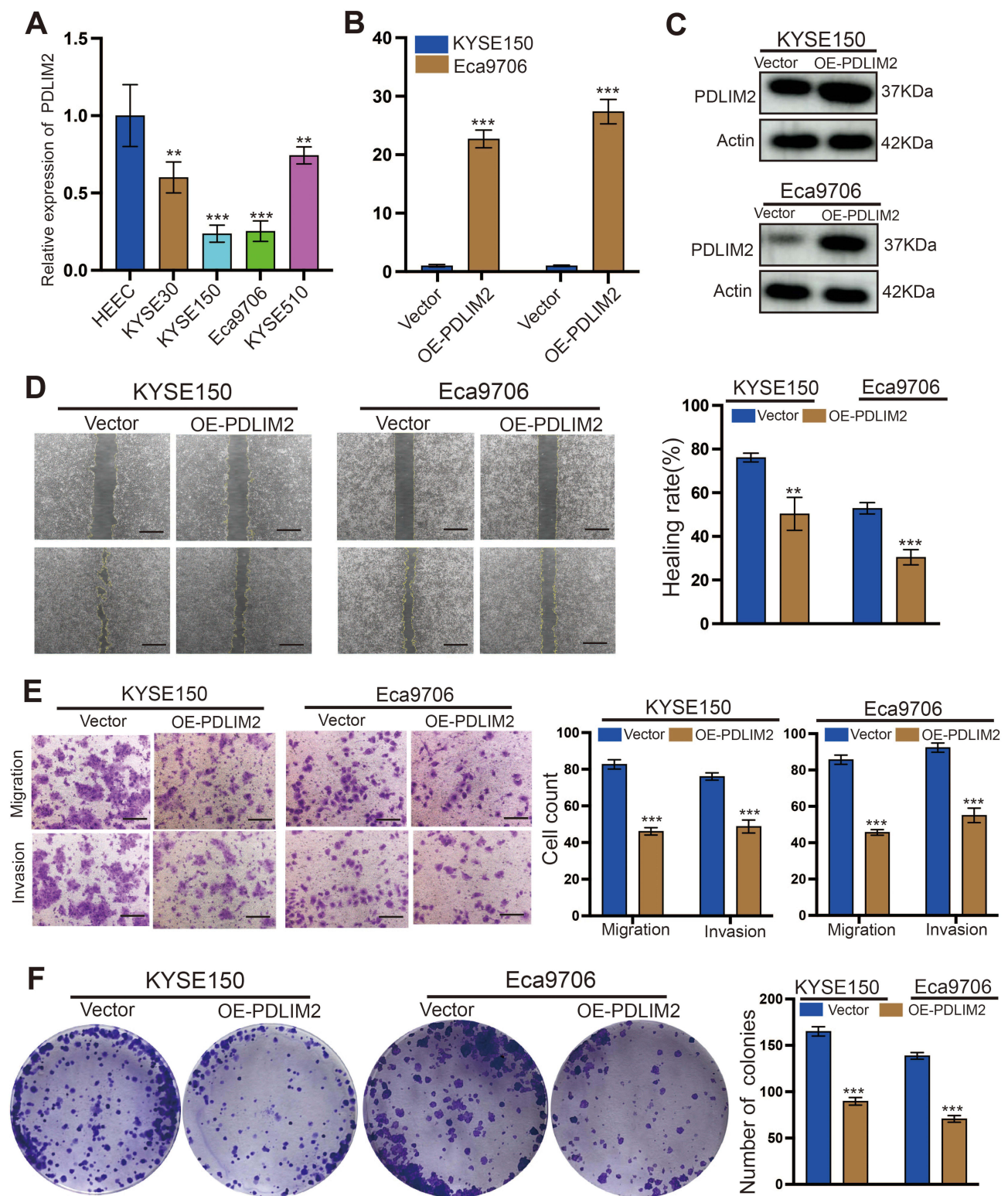
Figure 10 Overexpression of PDLIM2 significantly inhibits the proliferation and migration of KYSE150 and Eca9706 cells. (**A**) q-PCR detection of PDLIM2 expression in esophageal squamous cell carcinoma cell lines and normal esophageal epithelial cell lines. (**B** and **C**) q-PCR and WB were used to detect the overexpression efficiency of PDLIM2 in KYSE150 and Eca9706 cells. (**D**) The effect of PDLIM2 on the migration of KYSE150 and Eca9706 cells was explored in wound healing assays. (**E**) Compared with the control group, overexpressed PDLIM2 significantly inhibits the migration and invasion of esophageal squamous carcinoma cells. (**F**) The effect of PDLIM2 on the proliferation of KYSE150 and Eca9706 cells was investigated in colony formation assays. **p < 0.01, ***p < 0.001. Scale bars: 100 μm.

cancer-associated fibroblasts.[6] Our findings therefore suggest that the degeneration of epithelial cells is a critical event in ESCC tumorigenesis and progression.

To fully understand the relationship between epithelial cells and clinical prognosis in patients with ESCC, we extracted the marker genes of epithelial cells, obtained the intersection with genes from the GSE161533 and GSE20347 datasets for ESCC, and identified five core genes by LASSO regression analysis: CTSL, LAPTM4B, MYO10, NCF2, and PDLIM2. Validation was conducted using GSE45670 and the GEPIA database as external datasets.

CTSL (encoding cathepsin L) is located on human chromosome 9 and plays an essential function in lysosomes by degrading intracellular proteins; it is involved in various biological processes, including immunomodulation, apoptosis, and tissue remodeling.[32] CTSL furthermore degrades the extracellular matrix in breast cancer; increases cell proliferation, migration, and invasion; and disrupts intercellular adhesion.[33] In head and neck squamous cell carcinoma, CTSL is closely associated with the infiltration of immune cells and is a potential oncogenic factor.[34] The overexpression or aberrant activation of CTSL has therefore been associated with the development and progression of various cancers, and its regulation can potentially provide new therapeutic options.

LAPTM4B (encoding lysosome-associated transmembrane protein 4 beta) is found on human chromosome 8. Its protein interacts with yes-associated protein to create a positive feedback loop that promotes tumor cell invasion and metastasis.[35] In renal cell carcinoma, downregulation of LAPTM4B expression significantly reduces proliferation and promotes apoptosis.[36] Moreover, LAPTM4B regulates the RPS9/STAT3 axis to promote acute myeloid leukemia progression.[37] LAPTM4B is therefore essential for tumor cell proliferation, survival, invasion, metastasis, and immune escape.

MYO10 (myosin X) is located on human chromosome 7 and encodes actin, which is involved in intracellular protein transport and cellular motility. MYO10 plays a crucial role in mitotic progression, the regulation of genomic stability, and the development of inflammation and tumors.[38] It is significantly expressed in colorectal cancer tumor tissues and promotes colorectal cancer progression and metastasis by mediating the integrin/Src/FAK signaling pathway. High MYO10 expression is generally associated with tumor aggressiveness and a poorer prognosis.[39] Similar to the aforementioned core genes, MYO10 plays a vital role in cancer, especially in tumor invasion and metastasis.

NCF2 (encoding neutrophil cytosolic factor 2) is located on human chromosome 1 and is an essential subunit of the NADPH oxidase complex, which plays a crucial role in neutrophils of the immune system. The positive expression of NCF2 may promote ESCC aggressiveness. PDLIM2 is located on human chromosome 5 and interacts with other proteins to participate in intracellular signaling, cell adhesion, and other functions.[40] The gene plays a crucial role as a tumor suppressor or promoter in several cancers. For example, PDLIM2 attenuates ovarian cancer cell proliferation, migration, invasion, and EMT by inhibiting the TGF-β/Smad pathway.[41] In breast cancer, PDLIM2 expression affects tumor-associated stroma (particularly M2 macrophage infiltration), contributing to the progression of triple-negative and other breast cancer subtypes.[42] Therefore, the role of PDLIM2 in cancer is complex, and its specific function depends on the cell type, tumor type, and interactions involved, among other factors. In summary, CTSL, LAPTM4B, MYO10, NCF2, and PDLIM2 all represent noteworthy prognostic markers for ESCC.

The TME plays a vital role in cancer progression, and there is growing evidence that cancer immunotherapy is important for patients with tumors.[43] This study examined the differences in immune cell infiltration between high- and low-risk patients. Monocyte and macrophage counts of the high-risk group were significantly enriched. Monocytes release large amounts of inflammatory mediators, creating an inflammatory environment that promotes tumor growth and creates a fertile ecological environment for cancer development.[44] Meanwhile, the expression of regulatory T cells was higher in the high-risk group, and these cells are capable of generating an immunosuppressive microenvironment that hinders the CD8+ T cell-mediated eradication of tumor cells, which in turn promotes tumorigenesis progression.[45]

In recent years, considerable breakthroughs have been made in cancer treatment using monoclonal antibodies against immune checkpoint molecules. For example, PD-1/PD-L1 inhibitors have shown favorable therapeutic effects in solid tumors.[46] Our findings demonstrated that risk scores can be helpful biomarkers for immunotherapy. Risk scores were strongly positively correlated with PD-1/PD-L1, and there were significant variations in immunity and stromal scores between the high- and low-risk groups. Risk scores can also aid drug screening. This study revealed the potential therapeutic effects of AZD8186, tospletinib, gemcitabine, GSK2606414, mitoxantrone, RPT062607, SB505124, and

UMI.77 on ESCC, and a better understanding of the mechanism of action of these drugs can stimulate the development of novel cancer treatments.

β-Sitosterol is known to possess anticancer activity against various cancers.[47] A higher dietary intake of β-sitosterol is associated with a lower risk of ESCC, and phytosterol-rich foods or supplements are essential for preventing squamous esophageal cancer.[48] The ability of β-sitosterol to act on the five core genes identified was evaluated in this study. Our results indicated strong binding affinities between β-sitosterol and all of these genes, suggesting that β-sitosterol may be another potential therapeutic target for ESCC. It is the most abundant phytosterol and is found in a wide variety of plant foods. With the option to administer it via food intake, β-sitosterol can reduce the enormous public health and economic burden tremendously. Future research is warranted to investigate the mechanism of β-sitosterol action in ESCC treatment.

Our study has some limitations. The epithelial prognostic model was developed using a publicly available dataset and needs to be validated by extensive prospective clinical studies, and patients should not use nomograms to assess their risk of death without further validation in a representative patient population. In addition, due to the limitations of this study, prognostic genes could not be validated by independent studies.

Conclusions

Using single cell and bulk RNA-seq, this study explored the link between epithelial cells and the TME in ESCC. It revealed and validated essential prognostic genes, further elucidating epithelial cell heterogeneity through functional enrichment, cellular differentiation trajectories, and intercellular communication. In addition, we used a multigene combinatorial model to build predictive models, combined epithelial cell-related genes and prognostic risk models, and performed drug sensitivity prediction, cell mutation, and immune cell infiltration analyses. We furthermore analyzed the potential therapeutic role of β-sitosterol in ESCC to provide additional treatment options for patients with ESCC. And we verified the stability of the core genes by further experiments. In conclusion, our findings offer compelling support for the development of novel ESCC biomarkers and guide the early intervention or reversal of ESCC. We hope that our work will help advance the development of precise immunotherapies for ESCC and provide a reliable basis for individualized ESCC patient management and treatment.

Abbreviations:

C-index, concordance index; ESCC, esophageal squamous cell carcinoma; GDSC, Genomics of Drug Sensitivity in Cancer; GO, gene ontology; GSEA, gene set enrichment analyses; IC50, 50% inhibiting concentration; KEGG, Kyoto Encyclopedia of Genes and Genomes; K-M, Kaplan–Meier; LASSO, least absolute shrinkage and selection operator; OS, overall survival; scRNA-seq, single-cell RNA sequencing; TMB, tumor mutation burden; TME, tumor microenvironment; tSNE, t-distributed stochastic neighborhood embedding.

Data Sharing Statement

Data will be made available on request. For further information, please contact the corresponding author.

Ethics Approval

This study was performed in line with the principles of the Declaration of Helsinki. The study was approved by the Ethical Review Committee of Shandong Provincial hospital affiliated to Shandong First Medical University (approval number: NSFC: NO. 2024-619). Informed consent was waived by the approving ethics committee and data was de-identified.

Acknowledgments

Zhenhu Zhang and Bin Shang are co-first authors for this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically

reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors have no relevant financial or non-financial interests to disclose for this work.

References

1. Wang Y, Cheng J, Xie D, et al. NS1-binding protein radiosensitizes esophageal squamous cell carcinoma by transcriptionally suppressing c-Myc. *Cancer Commun*. 2018;38:33. doi:10.1186/s40880-018-0307-y

2. Huang C, Zhu Y, Li Q, et al. Feasibility and efficiency of concurrent chemoradiotherapy with a single agent or double agents vs radiotherapy alone for elderly patients with esophageal squamous cell carcinoma: experience of two centers. *Cancer Med*. 2019;8(1):28–39. doi:10.1002/cam4.1788

3. Oppedijk V, van der Gaast A, van Lanschot JJB, et al. Patterns of recurrence after surgery alone versus preoperative chemoradiotherapy and surgery in the CROSS trials. *J Clin Oncol*. 2014;32(5):385–391. doi:10.1200/JCO.2013.51.2186

4. He S, Xu J, Liu X, Zhen Y. Advances and challenges in the treatment of esophageal cancer. *Acta Pharm Sin B*. 2021;11(11):3379–3392. doi:10.1016/j.apsb.2021.03.008

5. Martínez-Reyes I, Chandel NS. Cancer metabolism: looking forward. *Nat Rev Cancer*. 2021;21(10):669–680. doi:10.1038/s41568-021-00378-6

6. Chen Y, Zhu S, Liu T, et al. Epithelial cells activate fibroblasts to promote esophageal cancer development. *Cancer Cell*. 2023;41(5):903–918.e8 e908. doi:10.1016/j.ccell.2023.03.001

7. Ren Q, Zhang P, Zhang X, et al. A fibroblast-associated signature predicts prognosis and immunotherapy in esophageal squamous cell cancer. *Front Immunol*. 2023;14:1199040. doi:10.3389/fimmu.2023.1199040

8. Kuksin M, Morel D, Aglave M, et al. Applications of single-cell and bulk RNA sequencing in onco-immunology. *Eur J Cancer*. 2021;149:193–210. doi:10.1016/j.ejca.2021.03.005

9. Ren X, Zhang L, Zhang Y, Li Z, Siemers N, Zhang Z. Insights gained from single-cell analysis of immune cells in the tumor microenvironment. *Annu Rev Immunol*. 2021;39(1):583–609. doi:10.1146/annurev-immunol-110519-071134

10. Chen Z, Zhao M, Liang J, et al. Dissecting the single-cell transcriptome network underlying esophagus non-malignant tissues and esophageal squamous cell carcinoma. *EBiomedicine*. 2021;69:103459. doi:10.1016/j.ebiom.2021.103459

11. Tabah A, Buetti N, Staiquly Q, et al. Eurobact-2 Study Group EEE and the ON epidemiology and outcomes of hospital-acquired bloodstream infections in intensive care unit patients: the EUROBACT-2 international cohort study. *Intensive Care Med*. 2023;49:178–190. doi:10.1007/s00134-022-06944-2

12. Rosenfeld JA, Wang Z, Schones DE, Zhao K, DeSalle R, Zhang MQ. Determination of enriched histone modifications in non-genic portions of the human genome. *BMC Genomics*. 2009;10(1):143. doi:10.1186/1471-2164-10-143

13. Wen J, Yang H, Liu MZ, et al. Gene expression analysis of pretreatment biopsies predicts the pathological response of esophageal squamous cell carcinomas to neo-chemoradiotherapy. *Ann Oncol*. 2014;25:1769–1774. doi:10.1093/annonc/mdu201

14. Clough E, Barrett T. The gene expression omnibus database. *Methods Mol Biol*. 2016;1418:93–110. doi:10.1007/978-1-4939-3578-9_5

15. Shi K, Li Y, Yang L, et al. Profiling transcriptional heterogeneity of epithelium, fibroblasts, and immune cells in esophageal squamous cell carcinoma by single-cell RNA sequencing. *FASEB J*. 2022;36:e22620. doi:10.1096/fj.202200898R

16. Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016;44(8):e71. doi:10.1093/nar/gkv1507

17. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573–3587.e29. doi:10.1016/j.cell.2021.04.048

18. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47(D1):D419–D426. doi:10.1093/nar/gky1038

19. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 1999;27(1):29–34. doi:10.1093/nar/27.1.29

20. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021;12:1088. doi:10.1038/s41467-021-21246-9

21. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14:1083–1086. doi:10.1038/nmeth.4463

22. Qiu X, Mao Q, Tang Y, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979–982. doi:10.1038/nmeth.4402

23. Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4(1):2612. doi:10.1038/ncomms3612

24. Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol*. 2019;37(7):773–782. doi:10.1038/s41587-019-0114-2

25. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747–1756. doi:10.1101/gr.239244.118

26. Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2013;41:D955–D961. doi:10.1093/nar/gks1111

27. Geeleher P, Cox N, Huang RS, Barbour JD. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One*. 2014;9(9):e107468. doi:10.1371/journal.pone.0107468

28. Chen Z, Huang Y, Hu Z, et al. Dissecting the single-cell transcriptome network in patients with esophageal squamous cell carcinoma receiving operative paclitaxel plus platinum chemotherapy. *Oncogenesis*. 2021;10(10):71. doi:10.1038/s41389-021-00359-2

29. Xia C, Dong X, Li H, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. *Chin Med J*. 2022;135 (5):584–590. doi:10.1097/CM9.0000000000002108

30. Gu J, Sun Y, Song J, et al. Irradiation induces DJ-1 secretion from esophageal squamous cell carcinoma cells to accelerate metastasis of bystander cells via a TGF-beta1 positive feedback loop. *J Exp Clin Cancer Res*. 2022;41(1):259. doi:10.1186/s13046-022-02471-6

31. Huang Y, Hong W, Wei X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *J Hematol Oncol*. 2022;15:129. doi:10.1186/s13045-022-01347-8

32. Ezz MA, Takahashi M, Rivera RM, Balboula AZ. Cathepsin L regulates oocyte meiosis and preimplantation embryo development. *Cell Prolif*. 2023;e13526. doi:10.1111/cpr.13526

33. Zhang L, Zhao Y, Yang J, et al. CTSL, a prognostic marker of breast cancer, that promotes proliferation, migration, and invasion in cells in triple-negative breast cancer. *Front Oncol*. 2023;13:1158087. doi:10.3389/fonc.2023.1158087

34. Gao F, Wang X, Qin N, et al. The analysis of cathepsin L that mediates cellular SARS-CoV-2 infection leading to COVID-19 in head and neck squamous cell carcinoma. *Front Immunol*. 2023;14:1156038. doi:10.3389/fimmu.2023.1156038

35. Liao J, Wang J, Xu Y, et al. LAPTM4B-YAP loop feedback amplification enhances the stemness of hepatocellular carcinoma. *iScience*. 2023;26:106754. doi:10.1016/j.isci.2023.106754

36. Wang H, Wang Q, Wu Y, Lou J, Zhu S, Xu Y. Autophagy-related gene LAPTM4B promotes the progression of renal clear cell carcinoma and is associated with immunity. *Front Pharmacol*. 2023;14:1118217. doi:10.3389/fphar.2023.1118217

37. Huang Y, Peng M, Qin H, et al. LAPTM4B promotes AML progression through regulating RPS9/STAT3 axis. *Cell Signal*. 2023;106:110623. doi:10.1016/j.cellsig.2023.110623

38. Mayca Pozo F, Geng X, Miyagi M, Amin AL, Huang AY, Zhang Y. MYO10 regulates genome stability and cancer inflammation through mediating mitosis. *Cell Rep*. 2023;42(5):112531. doi:10.1016/j.celrep.2023.112531

39. Ou H, Wang L, Xi Z, et al. MYO10 contributes to the malignant phenotypes of colorectal cancer via RACK1 by activating integrin/Src/ FAK signaling. *Cancer Sci*. 2022;113(11):3838–3851. doi:10.1111/cas.15519

40. Song G, Xu J, He L, et al. Systematic profiling identifies PDLIM2 as a novel prognostic predictor for oesophageal squamous cell carcinoma (ESCC). *J Cell Mol Med*. 2019;23(8):5751–5761. doi:10.1111/jcmm.14491

41. Lv W, Guo H, Wang J, Ma R, Niu L, Shang Y. PDLIM2 can inactivate the TGF-beta/Smad pathway to inhibit the malignant behavior of ovarian cancer cells. *Cell Biochem Funct*. 2023;41(5):542–552. doi:10.1002/cbf.3801

42. Cox OT, O'Sullivan N, Tresse E, Ward S, Buckley N, O'Connor R. PDLIM2 is highly expressed in breast cancer tumour-associated macrophages and is required for M2 macrophage polarization. *Front Oncol*. 2022;12:1028959. doi:10.3389/fonc.2022.1028959

43. Lavareze L, Scarini JF, de Lima-Souza RA, et al. Salivary gland cancer in the setting of tumor microenvironment: translational routes for therapy. *Crit Rev Oncol Hematol*. 2022;171:103605. doi:10.1016/j.critrevonc.2022.103605

44. Chen W, Li Y, Liu C, et al. In situ engineering of tumor-associated macrophages via a nanodrug-delivering-drug (beta-Elemene@Stanene) strategy for enhanced cancer chemo-immunotherapy. *Angew Chem Int Ed Engl*. 2023;62(41):e202308413. doi:10.1002/anie.202308413

45. Plitas G, Konopacki C, Wu K, et al. Regulatory T cells exhibit distinct features in human breast cancer. *Immunity*. 2016;45(5):1122–1134. doi:10.1016/j.immuni.2016.10.032

46. Darvin P, Toor SM, Sasidharan Nair V, Elkord E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Exp Mol Med*. 2018;50 (12):1–11. doi:10.1038/s12276-018-0191-1

47. Wang H, Wang Z, Zhang Z, Liu J, Hong L. Beta-sitosterol as a promising anticancer agent for chemoprevention and chemotherapy: mechanisms of action and future prospects. *Adv Nutr*. 2023;14:1085–1110. doi:10.1016/j.advnut.2023.05.013

48. Wang S, Zhao W, Sun L, et al. Independent and opposing associations of dietary phytosterols intake and PLCE1 rs2274223 polymorphisms on esophageal squamous cell carcinoma risk. *Eur J Nutr*. 2021;60:4357–4366. doi:10.1007/s00394-021-02561-9