Dovepress
Taylor & Francis Group

ORIGINAL RESEARCH

Crotonylation-Related Prognostic Model of Esophageal Squamous Cell Carcinoma Based on Transcriptome Analysis and Single-Cell Sequencing Analysis

Ruo-Yang Lin [ID][1], Ren-Pin Chen[1], Fu-Qiang Lin [ID][2]

[1]Department of Gastroenterology and Hepatology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, 325000, People's Republic of China; [2]Department of Thoracic Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, 325000, People's Republic of China

Correspondence: Fu-Qiang Lin, Email linfq66@qq.com

Background: Crotonylation is an emerging lysine acylation modification implicated in various diseases, yet its role in esophageal squamous cell cancer (ESCC) is unexplored. This study aimed to investigate the role of crotonylation-related genes (CRGs) in ESCC using bioinformatics approaches.

Methods: We included three ESCC datasets and 24 CRGs. Differentially expressed genes (DEGs) from TCGA-ESCA were intersected with key module genes related to CRGs to identify candidate genes. Univariate and LASSO regression analyses were conducted to select prognostic genes, which were then used to construct risk models. Independent prognostic analysis and nomogram construction followed. Functional enrichment and immune infiltration analyses were performed using the prognostic genes. Single-cell analysis was conducted to assess cell communication and pseudotemporal dynamics in key cells.

Results: Intersection of 1529 DEGs with 1,048 key module genes yielded 55 candidate genes. *OSM, FABP3, MICB*, and *FAM189A2* were identified as prognostic genes. These genes were used to classify ESCA patients into different risk groups and construct a nomogram. *FABP3* and FAM189A2 were enriched in neuroactive ligand-receptor interaction and ribosome terms. *MICB* and FABP3 showed strong positive correlations with natural killer T (NKT) cells, while *FAM189A2* negatively correlated with gamma delta T (γδT) cells. Single-cell analysis identified mast cells and neutrophils as key cells, differentiating into seven and three states, respectively.

Conclusion: Four genes (*OSM, FABP3, MICB*, and *FAM189A2*) were identified as prognostic crotonylation-related genes in ESCC, potentially involved in its pathogenesis. *OSM* was negatively correlated with ESCC, while *FABP3* and *MICB* were positively correlated.

Keywords: esophagus cancer, crotonylation, risk model, prognostic genes, single-cell sequencing analysis

Introduction

Esophageal carcinoma (ESCA) is a common malignant tumor of the digestive system; among all types of cancer, it ranked seventh and sixth in terms of mortality and morbidity, respectively, in 2020.[1] China has a high incidence of ESCA; new ESCA cases account for more than 50% of new cases worldwide, and new deaths constitute more than 50% of global deaths. ESCA ranks sixth and fourth among malignant cancers in terms of incidence and death.[2] According to histopathological classification, ESCA can be categorized into two main types, including ESCC and esophageal adenocarcinoma; between them, ESCC accounts for about 95% of all ESCA cases.[3] Most ESCC patients are already in the middle or advanced stages during admission. The treatment of ESCC remains a challenge despite significant advancements in this field and a considerable increase in the five-year survival rate. Elucidating the mechanisms

415

underlying the occurrence and development of ESCC, understanding the heterogeneity of ESCC, and finding new effective therapeutic targets are required to increase the overall survival rate of patients with ESCC.[4]

After the concept of epigenetics was first introduced by the Austrian developmental biologist Conrad Waddington in 1942, scientists have been interested in understanding epigenetic mechanisms.[5] Along with DNA methylation, histone modifications are also major components of the epigenetic code. Over the past decade, various histone acylation modifications have been discovered, including histone propionylation (Kpr), butylation (Kbu), 2-hydroxyisobutylation (Khib), succinylation (Ksucc), malondialdehyde (Kma), glutarylation (Kglu), crotonylation (Kcr), beta-hydroxybutylation (Kbhb), and benzene formylation (Kbz). Crotonylation is a novel and versatile lysine acylation modification involved in various biological processes, such as gene expression, the cell cycle, spermatogenesis, aging, and DNA damage. Compared to other histone modifications, including acetylation, crotonylation offers a distinct contribution. While both crotonylation and acetylation target lysine residues, the modification groups in crotonylation are larger. Crotonylation is distinguished from acetylation in histone modifications by its unique structural features, including a four-carbon planar chain and an alkenyl C=C double bond, which give it distinct functional properties.[6] Through in vitro and cellular models, Allis et al demonstrated that the transcriptional activation driven by histone crotonylation is even more potent than that induced by histone acetylation.[7] Besides, crotonylation has a greater impact on regulating the cell cycle and metabolism than acetylation, acting on specific genes or under certain environmental conditions, and is found at the promoters or enhancers of actively transcribed genes, where it modulates transcription.[8] Crotonylation plays a key role in multiple diseases, such as acute kidney injury, depression, hypertrophic cardiomyopathy, as well as cancer.[9] The loss of acyl-CoA oxidase 2 (Acox2) can decrease the level of crotonylation of peroxidase, thus compromising metabolic homeostasis and inducing liver cancer in mice.[10] Lysine crotonylation strongly influences cytoplasmic and mitochondrial proteins. In pancreatic cancer cells, enzymes involved in metabolic processes, including glycolysis, the tricarboxylic acid (TCA) cycle, fatty acid metabolism, and glutamine, undergo extensive crotonylation.[11] By affecting gene transcription, crotonylation plays an important role in the biological processes of cancer cell proliferation, differentiation and apoptosis. Research shows that glioblastoma stem cells (GSCs) reprogram lysine metabolism to accumulate crotonyl-CoA, promoting histone lysine crotonylation, which supports tumor growth and immune evasion.[12] Another study reveals that LncRNA LINC00922 promotes colorectal Cancer (CRC) metastasis by increasing crotonylation on Lys27 of histone H3 (H3K27cr), which activates ETS1 transcription and enhances cell migration and invasion.[13] Crotonylation is upregulated in ESCC.[9] However, no study has investigated the function and prognostic risk assessment of crotonylation-related genes in ESCC.

Hence, in this study, we examined crotonylation-related prognostic genes in ESCA and analyzed the functions and pathways of prognostic genes via bioinformatics based on transcriptome and single-cell data. Additionally, a prognostic model and a nomogram were constructed, followed by the analysis of the immune microenvironment and drug sensitivity in the high-risk and low-risk groups. Moreover, the key cells related to prognostic genes were identified, and the related functions of the key cells were analyzed. Finally, we verified the expression of these crotonylation-related DEGs in ESCC patients in clinical samples. These findings revealed the importance of crotonylation in the occurrence and development of ESCC and provided some novel ideas for the development of new targets and individualized treatments for ESCC.

Materials and Methods

Data Sources

In this study, data from patients with ESCA were downloaded from the Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/), including RNA-seq data from 11 normal esophageal tissues and 161 esophageal cancer tissues, as well as, relevant clinical information (age, sex, tumor grade, tumor stage, mutation data, and copy number variation (CNV)) and survival information. Complete clinical information is available for 160 cases. Additionally, both the single-cell dataset GSE196756 (GPL24676) and the validation dataset GSE53622 (GPL18109) were derived from GEO databases (http://www.ncbi.nlm.nih.gov/geo/), where GSE196756 included three esophageal squamous cell carcinoma tissue samples and three paracarcinoma tissue samples. The GSE53622 dataset had 238 samples, of

which 119 esophageal squamous cell carcinoma samples had information on overall survival (OS). Finally, we obtained 24 CRGs from reference.[9] All datasets were downloaded and analyzed starting from April 17th to April 30th, 2024.

Screening of CRG-Related Key Module Genes

First, to investigate the relationship between CRGs and ESCA prognostic survival, the single-sample gene set enrichment analysis (ssGSEA) scores for CRGs in the ESCA sample were calculated in TCGA-ESCA using the GSVA package (v 1.42.0).[14] The ESCA samples were classified into high/low score groups based on the best ssGSEA score cutoff values using the R package "survminer" (v 0.4.9). Next, survival differences between these two groups were compared using Kaplan-Meier (K-M) survival curves using the R package "survminer" (v 0.4.9). Subsequently, in order to identify co expressed gene modules and identify module genes highly correlated with CRGs, the sample expression matrix of TCGA-ESCA were subjected to Weighted correlation network analysis (WGCNA) via the WGCNA package (v 1.71).[15] The samples were first clustered through the expression of the Euclidean distance for hierarchical clustering to check for outliers in the sample. The appropriate soft threshold power was selected from 1–20 for intergenic correlations. Next, scale-free network construction was performed based on the filtered soft thresholds, and all genes were divided into different modules (the minimum number of genes was 200 for each module). Finally, the ssGSEA scores associated with the CRGs were used as traits to calculate the correlation coefficient matrices between the module eigenvectors and the phenotypic traits, and a correlation heat map was drawn. The modules with the highest correlation with the phenotypic traits were subsequently selected as key modules, and the genes associated with these modules were considered to be key module genes closely related to CRGs.

Differential Expression Analysis and Acquisition of Candidate Genes

In the TCGA-ESCA training set, to explore genes related to ESCA, the R package "DESeq2" (v 1.40.2) was used to screen for DEGs between 161 primary tumor samples (ESCA) and 11 normal tissue samples (normal) next to cancer samples ($p < 0.05$ and $|log2FoldChange| > 2.0$).[16] These DEGs were overlapped with key module genes to identify candidate genes. Next, we used the clusterProfiler package (v 4.7.1.3) to conduct gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analyses on the candidate genes to reveal their biological functions ($p < 0.05$).[17]

Construction and Validation of Risk Models

In this study, the TCGA-ESCA data were divided into training and at a ratio of 7:3 for internal validation of the risk model, with 112 and 48 samples, respectively. First, univariate Cox regression analysis was performed on the training set using the R packages "survival" (v 3.3.1) and "survminer" (v 0.4.9) to identify factors associated with patient survival time. The candidate genes were obtained and screened by the proportional hazards (PH) test ($p > 0.05$), and the genes that passed the test were incorporated into the univariate Cox analysis to correlate the survival information of each sample and screen for key genes ($p < 0.05$). Then, LASSO regression analysis with 10-fold cross-validation was performed to identify the prognostic genes used to construct the risk model, and the LASSO model with the smallest error was selected as the result of the analysis. Risk scores were calculated based on prognostic gene expression. In the training and test sets, the samples were categorized into high-risk and low-risk groups based on the median risk score, whereas, in the validation set, the grouping was based on the best cutoff value. The risk model was further validated in the test set and the GSE53622 dataset, where the survival differences between risk groups, the accuracy of the risk model, and the survival status between different risk groups were reflected by the K-M ($p < 0.05$), receiver operating characteristic (ROC) (larger AUC values indicated better predictive ability, and vice versa) and survival curves, respectively. In addition, to further verify the correlation between prognostic genes and survival outcomes, 161 esophageal cancer tissue samples in TCGA-ESCA were divided into high and low expression groups based on the optimal cutoff value of each prognostic gene expression level. The survival differences between the high and low expression groups were compared using K-M curves.

Construction of a Nomogram and Analysis of Clinical Features Between Risk Groups

In TCGA-ESCA, to construct a prognostic model for predicting patient survival, we combined the risk score with other clinical indicators to identify clinical indicators related to patient survival time. P-tests were initially conducted on eight variables (age, stage, sex, T stage, N stage, and histology). Based on the results of this analysis, a univariate Cox regression analysis ($p < 0.05$) was conducted to identify independent prognostic factors. A nomogram for predicting the one-, two-, and three-year survival of ESCA patients was subsequently constructed, visually demonstrating the impact weights of each variable on prognosis. The predictive performance was evaluated via calibration and receiver operating characteristic (ROC) curves. Finally, the differential performance of risk scores based on different clinical indicators was assessed, and the results were visualized by constructing heat maps.

Functional Analysis

We investigated the biological functions of genes associated with different risk groups. Using the DESeq2 package, differential expression analysis was conducted between the risk groups, and the calculated log2FC values were sorted from largest to smallest. Next, the KEGG-enriched gene set (cc2.cp.kegg.v7.4.symbols.gmt) in the MSigDB was used as a reference gene set to perform GSEA using the clusterProfiler package (v 4.7.1.3).[17] GSEA ($p < 0.05$) was also performed for prognostic genes via c2.cp.kegg.v2023.1.Hs.symbols.gmt as the reference gene set to determine the biological pathways involved.

Immune-Related Analyses

In order to gain a deeper understanding of the function and role of the immune system in ESCC, the ssGSEA algorithm in the GSVA package was used to calculate 28 immune infiltration cell enrichment scores, and their differences among different risk groups were compared. Additionally, the correlation of prognostic genes with differential immune cells was assessed. Moreover, 48 immune checkpoints extracted from previous studies[18] were overlapped with genes in TCGA-ESCA to obtain the levels of expression of 48 immune checkpoints, and their differences between risk groups were further investigated ($p < 0.05$). The immune score, stromal score, and ESTIMATE score of the ESCA samples were calculated in TCGA-ESCA via the ESTIMATE algorithm, and the Wilcoxon test ($p < 0.05$) was performed to analyze the differences between risk groups. Their correlation with risk scores was calculated using Spearman's method (correlation > $|0.3|$, $p < 0.05$).

Tumor Mutational Burden (TMB) and Drug Sensitivity Analyses

To determine the mutation status of the genes in the risk groups, the mutation data were downloaded from the TCGA database based on TCGA-ESCA, and the top 20 mutated genes in the risk groups were illustrated using the MAfTools package (v 3.12).[19] The differences ($p < 0.05$) in TMB based on the risk group were further examined using K-M survival curves. Next, the chemotherapeutic drugs with significantly different responses in the risk groups were screened based on the GDSC database. The IC50 values of the common chemotherapeutic drugs were calculated for each sample using the pRRophetic (v 0.5) algorithm, and their differences between risk groups were determined by conducting the Wilcoxon test ($p < 0.05$).

Regulatory Network Analysis

To elucidate the molecular regulatory mechanisms of the prognostic genes, the microRNAs (miRNAs) were predicted using the mirTarbase database on the NetworkAnalyst platform. The long non-coding RNAs (lncRNAs) of the miRNAs were also obtained from the miRNet database. Finally, the complete lncRNA–miRNA-mRNA network was constructed using Cytoscape. The ENCODE database was used to search for transcription factors (TFs) that regulate prognostic genes, and the TF-mRNA network was visualized using Cytoscape (v 3.0).[20]

Single-Cell Analysis

The single-cell dataset (GSE196756) first underwent quality control through the Seurat package (v 4.1.0),[21] where the screening criterion for the number of nFeature_RNAs was greater than 100 and less than 5000), the number of

nCount_RNAs was less than 20000), and the percentage.mt was less than 5%. Next, the data were normalized by the normalizeData function, and the highly variable genes were identified by the FindVariableFeatures function. After the data were normalized by the ScaleData function, the dataset was analyzed by principal component analysis (PCA) for dimensionality reduction. The cells were subsequently classified into different cell clusters via UMAP cell clustering analysis, where the resolution was set to 0.2. Then, based on the marker genes, the cell clusters were annotated as different cell types using the singleR package (v 3.11) and the CellMarker database.[21] The Wilcoxon test was conducted to analyze the expression of prognostic genes in different cell types in ESCC samples and control samples to further screen key cells. Additionally, cell communication analysis was performed for cell types using the CellChat package (v 1.6.1).[22] Based on the results of the above analysis, the proposed timing analysis for key cells was performed using the monocle package (v 0.4.3).[23]

Quantitative Real-Time Polymerase Chain Reaction (qRT–PCR)

The level of expression of the prognostic genes was confirmed by qRT-PCR. Five pairs of samples (including five ESCC samples and five normal samples) from patients at the First Affiliated Hospital of Wenzhou Medical University were used in the qRT-PCR analysis from April 19, 2024 to May 5, 2024. The study was approved by the Ethics Committee of the First Affiliated Hospital of Wenzhou Medical University. All studies were carried out in accordance with the principles stipulated in the Declaration of Helsinki, and all patients signed an informed consent form.

First, the total RNA of the frozen sample was extracted with TRIzol. Then, an equal amount of mRNA was reverse-transcribed to synthesize cDNA. Next, qPCR was performed using a 2x Universal Blue SYBR Green qPCR Master Mix Kit according to the following reaction system: predenaturation at 95 °C for 1 min; denaturation at 95 °C for 20s, annealing at 55 °C for 20s, and extension at 72 °C for 30s; the reaction was conducted for 40 cycles. Finally, the gene expression level was calculated using the $2^{-\Delta\Delta CT}$ method. The primer sequences used for the genes are provided in Supplementary Table 1.

Statistical Analysis

Bioinformatic analysis was conducted in the R software (v 4.1.0). All results were considered to be statistically significant at $p < 0.05$. To further validate the reliability of the statistical method, G * power (v 3.1) software was used for statistical efficacy analysis, with an effect size set to 0.5 and a significance level set to 0.05.[24]

Results

The MEred Module Contained 1048 Key Module Genes

Based on the calculation results, the statistical test power reaches 0.8 with 128 participants (Supplementary Figure 1). The actual total number of participants in this study is 160, ensuring robust statistical power. Based on the ssGSEA scores of the 24 CRGs, the ESCA samples were divided into a high-score group (171 samples) and a low-score group (53 samples). Moreover, the K-M curves revealed significant differences in survival between these two score groups (Figure 1a). No outlier samples were observed in TCGA-ESCA (Figure 1b). When the optimal soft threshold was set to 4, the grid approximated the scale-free distribution (Figure 1c). The samples in TCGA-ESCA was subsequently classified into 10 coexpression modules (Figure 1d), among which the MEred module had the strongest correlation with the CRG ssGSEA score (cor = 0.3, p = 0.0011) and was considered to be the key module gene (Figure 1e). Additionally, 1,048 genes in the MEred module were considered to be key module genes.

Candidate Genes Identified for Functional Enrichment Analysis

In TCGA-ESCA, we identified 1,529 DEGs, including 891 upregulated genes and 638 downregulated genes (Figure 2a and B). Moreover, 55 candidate genes were identified based on their intersection with key module genes (Figure 2c). GO enrichment analysis revealed that the candidate genes were enriched in 250 functional terms, including tertiary granule and collagen-containing extracellular matrix (Figure 2d). Additionally, 89 signaling pathways, such as cytokine–cytokine receptor interactions and the chemokine signaling pathway, were enriched in the KEGG analysis (Figure 2e).
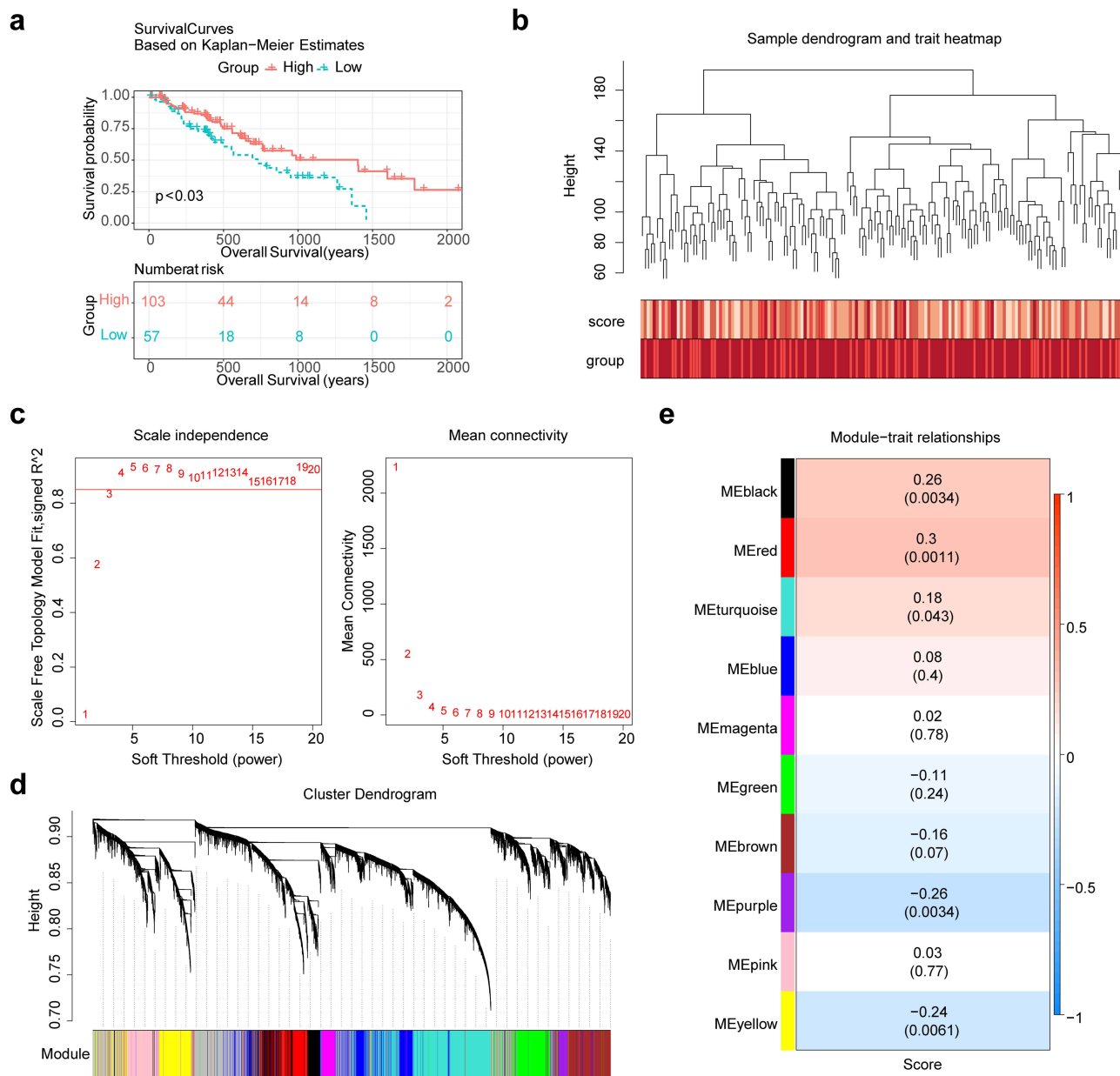
Figure 1 Screening of key module genes in the MEred module. (**a**) OS of patients in the high-score and low-score groups was determined from the K-M curves. (**b**) Weighted correlation network analysis of samples from TCGA-ESCC. (**c**) A scale-free network was constructed based on the filtered soft thresholds. (**d**) Classification of genes into different modules. (**e**) Correlations between CRGs and the modules assessed by the ssGSEA scores.

Abbreviations: OS, overall survival; CRGs, crotonylation-related genes; ssGSEA, single-sample gene set enrichment analysis.

OSM, FABP3, MICB, and FAM189A2 Were Used to Construct Risk Models

In the training set, four genes were identified via univariate Cox regression analysis (Figure 3a). Additionally, *OSM, FABP3, MICB*, and *FAM189A2* were considered as prognostic genes in the LASSO regression analysis when the λmin value was 0.01054443 (Figure 3b). The risk model for prognostic genes was calculated using the following formula: riskScore = *OSM* × (0.543) + *FABP3* × (0.228) + *MICB* × (0.169) + *FAM189A2* × (–0.377). In the training set, the survival rate of the high-risk group was considerably lower than that of the low-risk group (Figure 3c). The AUC values of the risk model according to the ROC curve at one, two, and three years were greater than 0.7 (Figure 3d). The survival time of patients decreased as the risk score increased, as determined by the risk curve (Figure 3e). These results indicated that the risk model was reliable and accurate. The areas under the ROC curves (AUCs) at one, two, and three years were
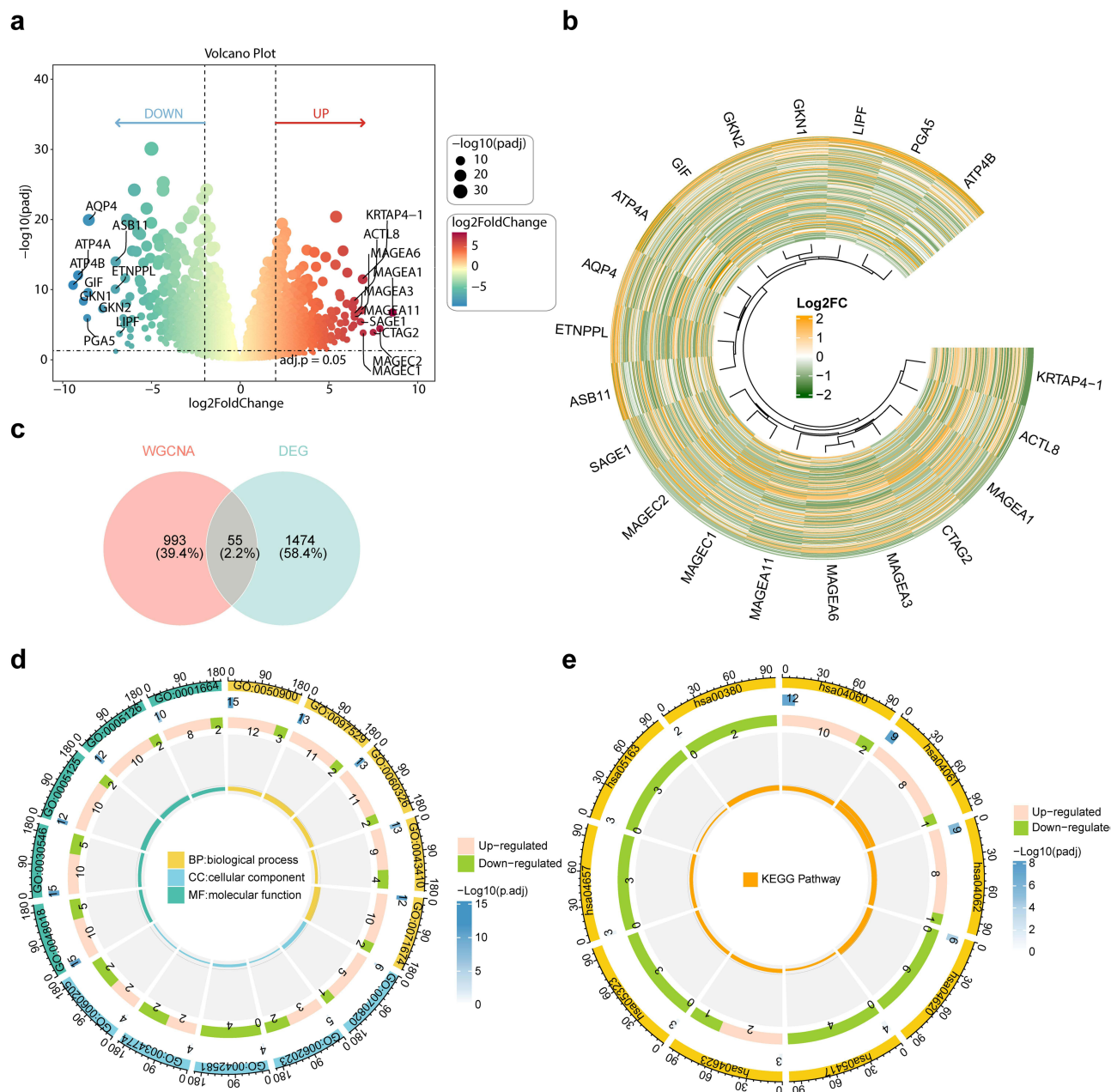
Figure 2 Screening and pathway enrichment analysis of candidate genes. (**a** and **b**) DEGs screened from TCGA-ESCC. (**c**) Intersection of DEGs with key module genes to further screen candidate genes. (**d**) The results of the GO functional enrichment analysis of the candidate genes. (**e**) The results of the KEGG functional enrichment analysis of the candidate genes.

Abbreviations: DEGs, differentially expressed genes; GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes.

greater than 0.7 and 0.6 in the test set (Figure 4a–c) and GSE53622 validation set (Figure 4d–f), respectively. The survival rate was significantly lower in the high-risk group than in the low-risk group ($p < 0.05$), and patient survival time decreased as the risk score increased. These results demonstrated the generalizability and validity of the risk model. The K-M survival curve of prognostic genes showed significant differences in the survival status of samples between high and low expression groups of *OSM, FABP3*, and *FAM189A2* genes (Figure 4g–i).

Construction of a Nomogram

After P-tests for the risk scores and clinical indicators (sex, stage, age, T stage, N stage, and risk score) were performed, the indicators risk score, stage IV, and N1 stage were found to be significantly associated with the survival prognosis of ESCA
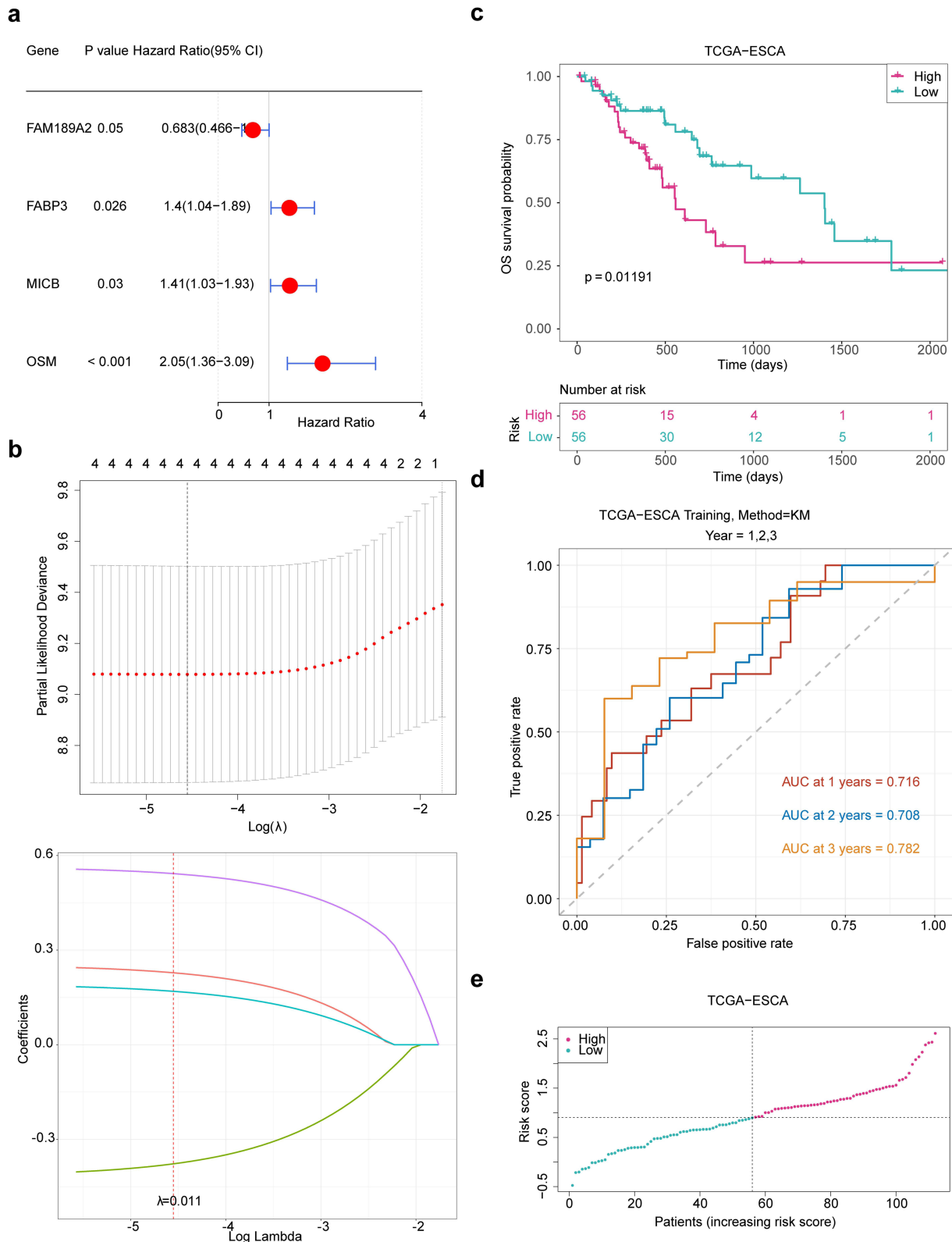
Figure 3 Construction of the risk model and its validation in the training set. (**a**) Univariate Cox regression analysis of candidate genes was performed. (**b**) The prognostic genes (*OSM, FABP3, MICB*, and *FAM189A2*) used to construct the risk model were screened by LASSO regression analysis. (**c**) Survival in the low-risk and high-risk groups of the training set was determined by the K-M curves. (**d**) The accuracy of the risk model assessed in the training set was evaluated using ROC curves. (**e**) The survival status of different risk groups was analyzed in the training set based on survival curves.

Abbreviations: LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic.
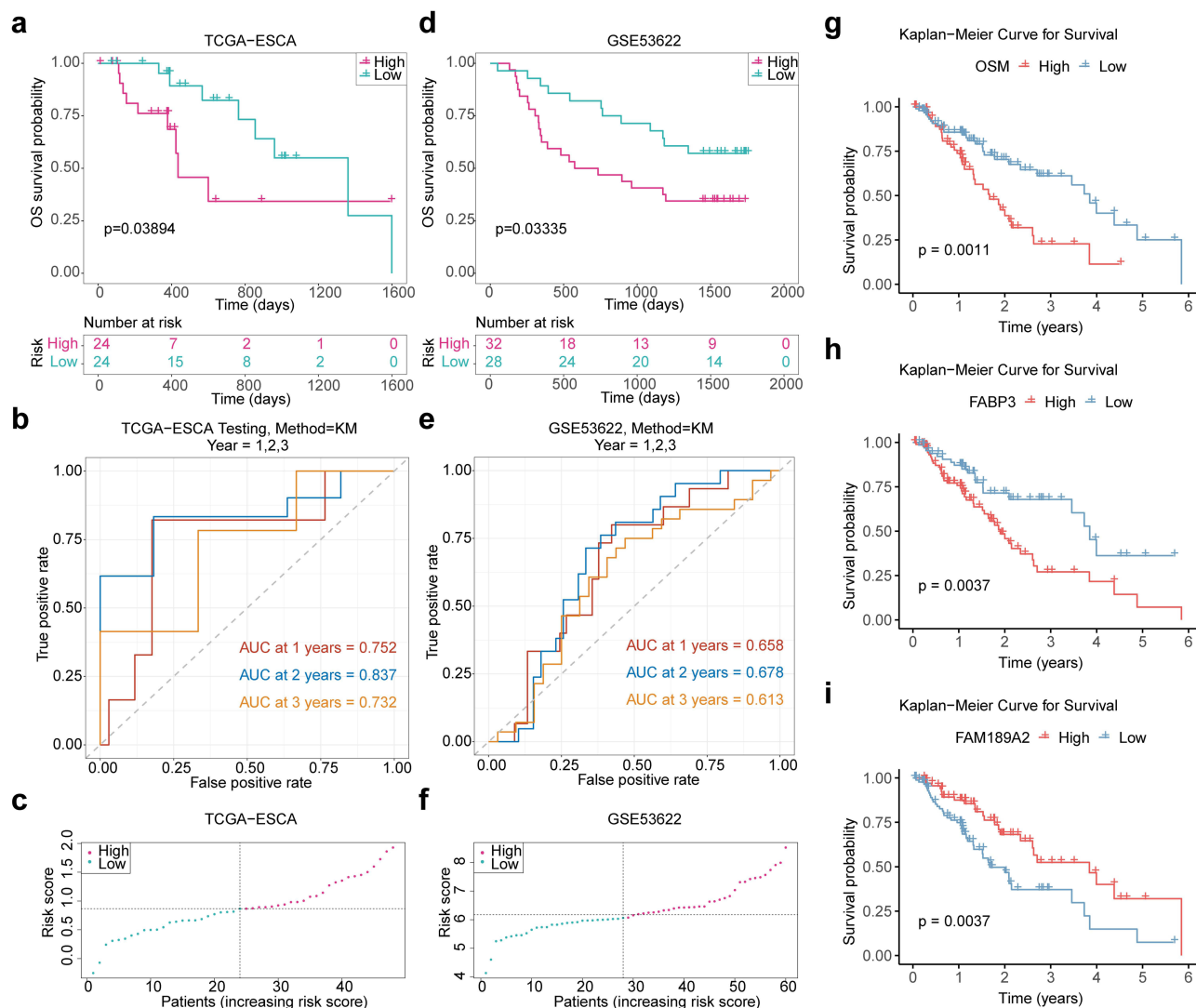
Figure 4 Validation of the risk model in the test set and the GSE53622 dataset. (**a**) Survival in the low-risk and high-risk groups of the test set was evaluated based on K-M curves. (**b**) The accuracy of the risk model was analyzed in the test set with ROC curves. (**c**) The survival status of different risk groups was assessed in the test set with survival curves. (**d**) Survival in the low-risk and high-risk groups in the GSE53622 dataset was analyzed using K-M curves. (**e**) Accuracy of the risk model was assessed in the GSE53622 dataset with ROC curves. (**f**) The survival status of different risk groups was analyzed in the GSE53622 dataset with survival curves. (**g–i**) Survival curve of high and low expression groups of prognostic genes. (**g**) *OSM* gene. (**h**) *FABP3* gene. (**i**) *FAM189A2* gene. The horizontal axis represents survival time, and the vertical axis represents survival rate.

Abbreviation: ROC, receiver operating characteristic.

patients, as determined by univariate Cox analysis (Figure 5a). Next, we constructed a nomogram to predict patient survival at one, two, and three years (Figure 5b). The slope of the nomogram was close to 1 in the calibration curve (Figure 5c), and the AUC value was greater than 0.6 in the ROC curve (Figure 5d), which indicated that the nomogram had good predictive performance. Finally, the analysis of risk scores and clinical indicators showed that high risk scores were concentrated in the high-risk group, and the expression of *OSM, FABP3*, and *MICB* showed the opposite trend as that of *FAM189A2* (Figure 5e).

Signaling Pathways Significantly Enriched in Different Risk Groups Included Steroid Biosynthesis and Others

In the functional enrichment analysis of different risk groups, the significantly enriched signaling pathways included steroid biosynthesis and terpenoid backbone biosynthesis (Figure 6a). In the GSEA of the prognostic genes, *FABP3* and *FAM189A2* were significantly enriched in neuroactive ligand-receptor interaction and the ribosome signaling pathway,
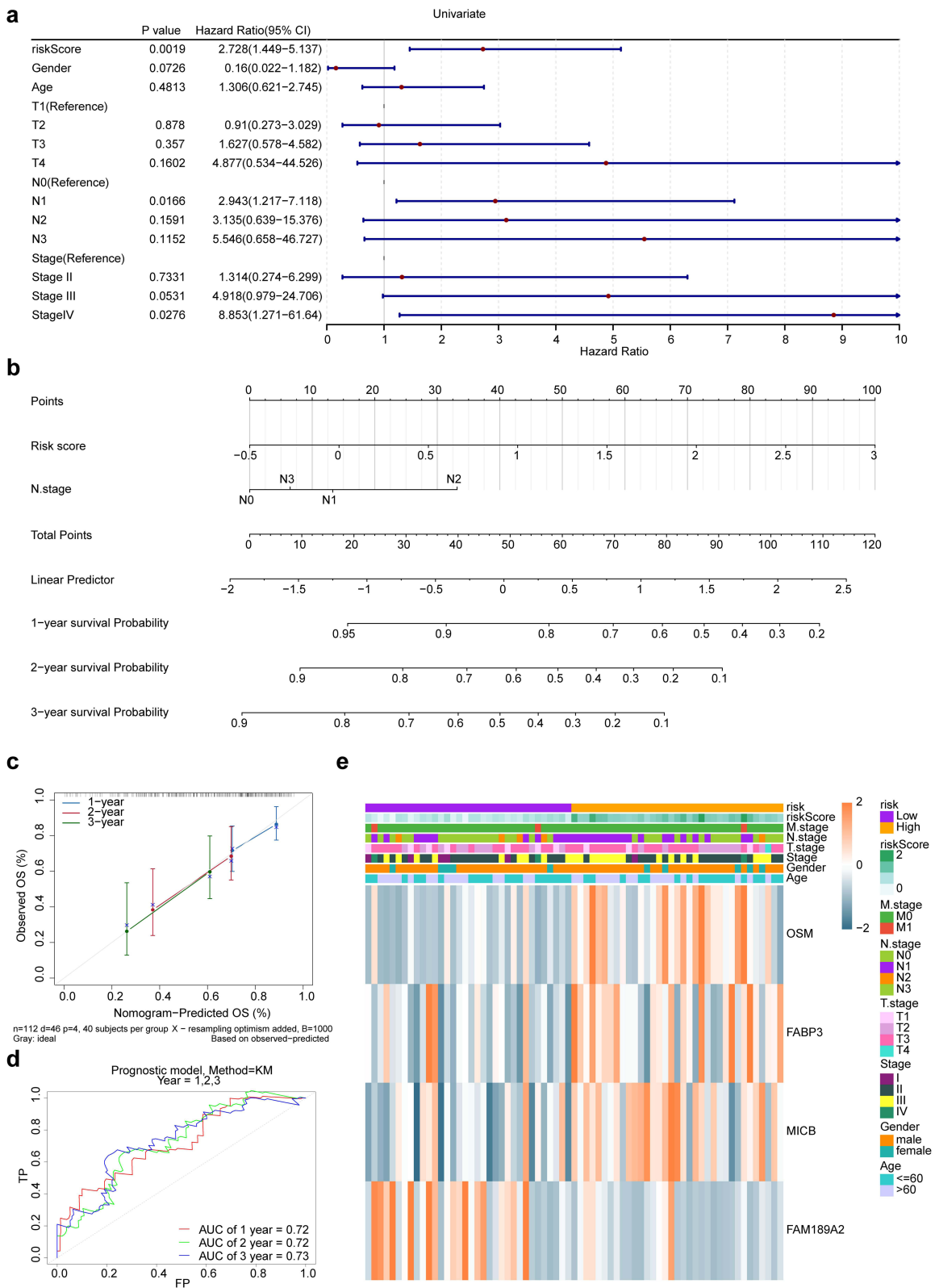
Figure 5 Construction of a nomogram and analysis of clinical features between risk groups. (**a**) Univariate Cox regression analysis was performed to identify independent prognostic factors (risk score, stage IV, and N1) among the risk score and variables obtained via P-tests. (**b**) A nomogram was constructed to predict the one-, two-, and three-year survival of ESCC patients. (**c** and **d**) The predictive performance of the nomogram was evaluated by calibration (**c**) and ROC (**d**) curves. (**e**) Differential risk scores based on different clinical indicators are shown.

Abbreviation: ESCC, esophageal squamous cell cancer.

whereas *FABP3* and *MICB* were involved in the olfactory transduction signaling pathway. *OSM* was also significantly enriched in the chemokine signaling pathway and cytokine-cytokine receptor interaction (Figure 6b–e).

Screening of Differential Immune Cells

The ssGSEA scores of 28 immune cells were calculated (Figure 7a); among these cells, 11 immune cells, including activated dendritic cells and central memory CD8$^+$ T cells, exhibited significantly greater infiltration in the high-risk
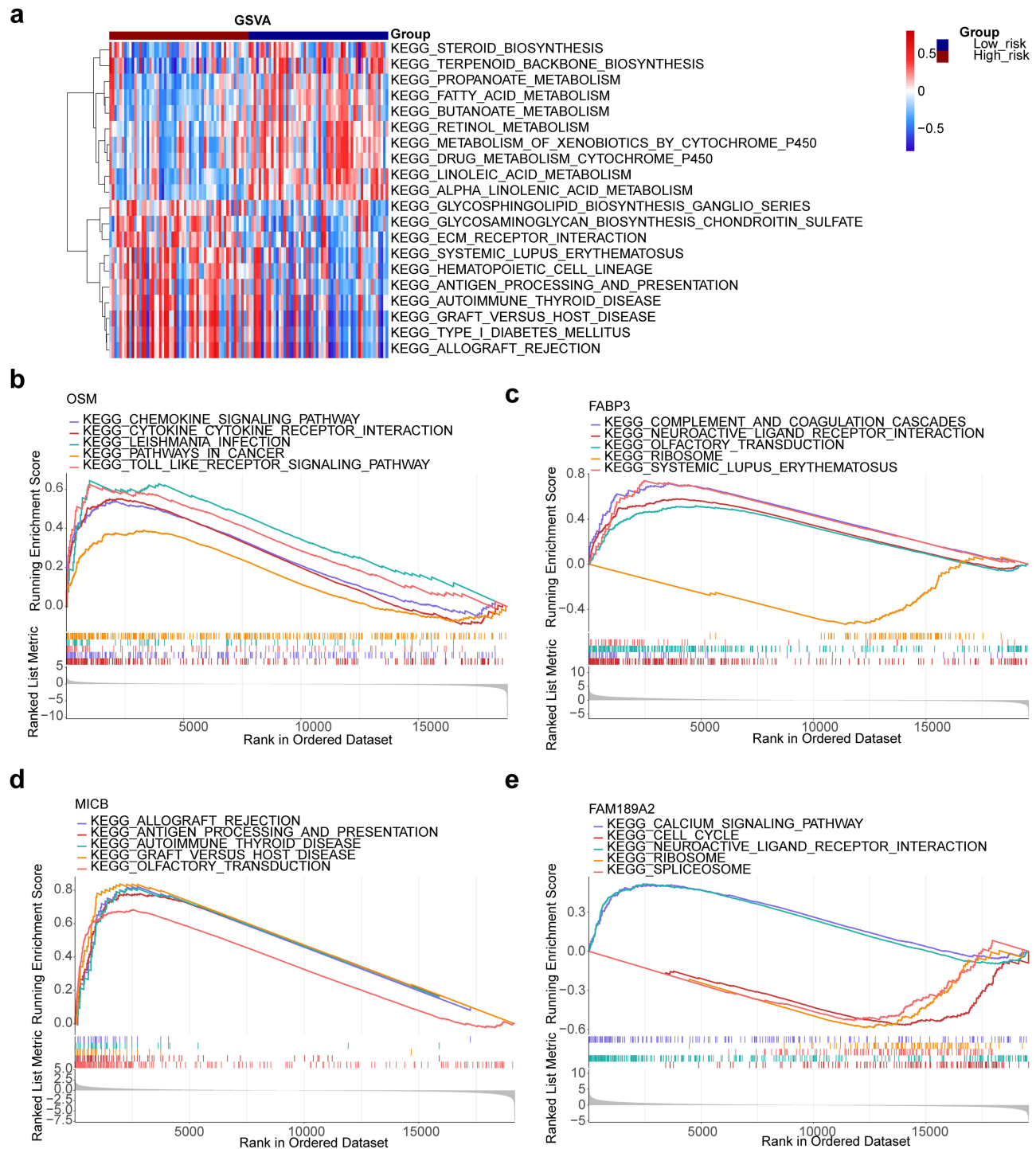

Figure 6 Functional enrichment analysis and pathway analysis of *OSM*, *FABP3*, *MICB*, and *FAM189A2*. (**a**) Functional enrichment analysis of different risk groups. (**b–e**) GSEA of *OSM* (**b**), *FABP3* (**c**), *MICB* (**d**), and *FAM189A2* (**e**).
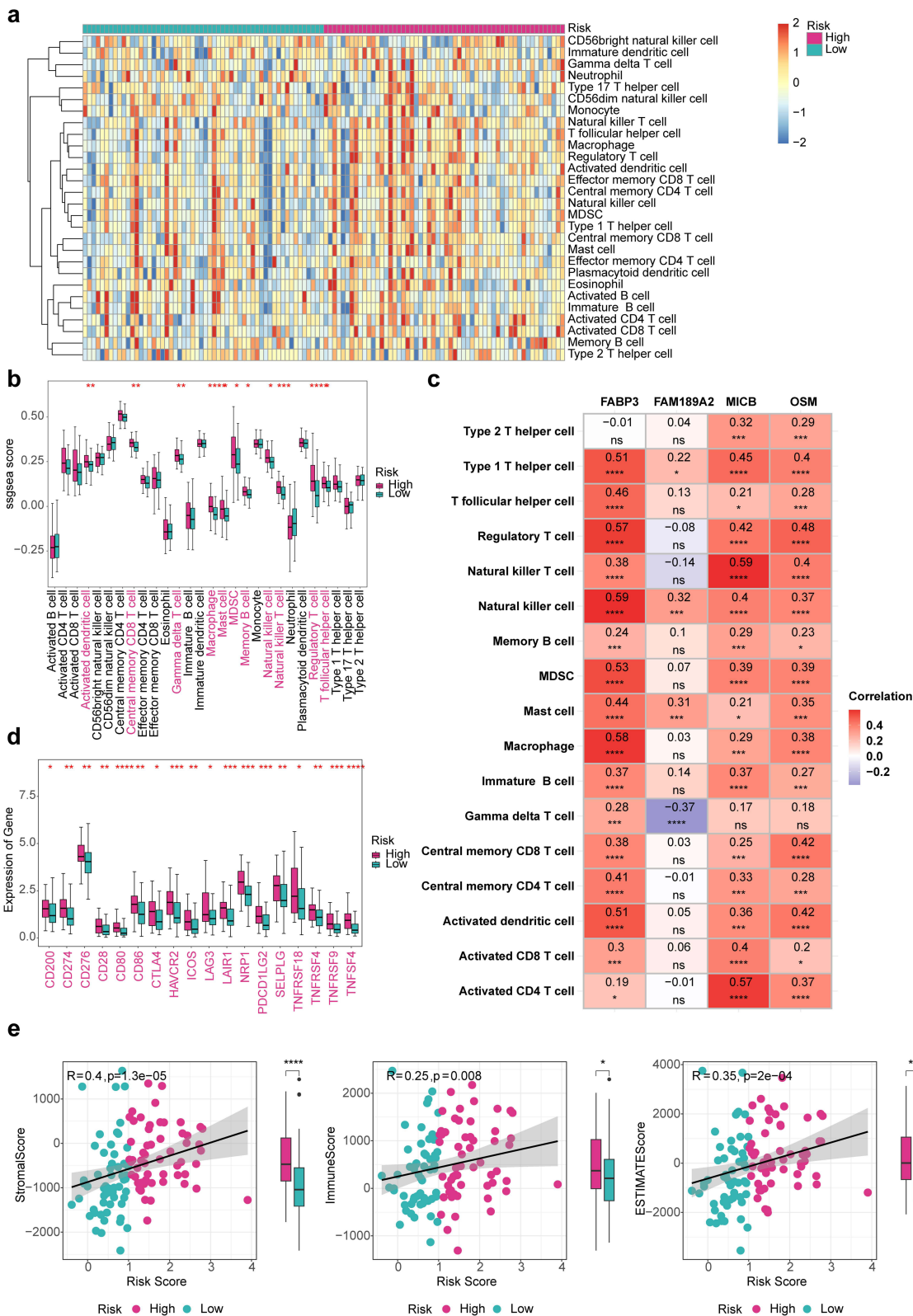
Figure 7 Screening of differential immune cells. (**a**) Enrichment scores of 28 infiltrating immune cells were calculated using the ssGSEA algorithm. (**b**) The differences in enrichment scores between the low-risk and high-risk groups. "*" represented $p < 0.05$, "**" represented $p < 0.01$, "***" represented $p < 0.001$, "****" represented $p < 0.0001$. (**c**) Correlations of *OSM, FABP3, MICB*, and *FAM189A2* with different immune cells. "ns" represented no significance, "*" represented $p < 0.05$, "***" represented $p < 0.001$, "****" represented $p < 0.0001$. (**d**) The expression of 48 immune checkpoints extracted from the literature was analyzed with TCGA-ESCA, and differences in expression between the low-risk and high-risk groups were determined. "*" represented $p < 0.05$, "**" represented $p < 0.01$, "***" represented $p < 0.001$, "****" represented $p < 0.0001$. (**e**) Correlations between the immune score, stromal score, and ESTIMATE score of ESCA samples in TCGA-ESCA and risk scores were assessed using Spearman's method. **Abbreviation**: ssGSEA, single-sample gene set enrichment analysis.

group (Figure 7b). Additionally, *MICB* and *FABP3* had the greatest positive correlations with natural killer T(NKT) cells (cor = 0.59), whereas *FAM189A2* had the greatest negative correlation with gamma delta T ($\gamma\delta$T) cells (cor = –0.37) (Figure 7c). Among the 48 immune checkpoints, 16 immune checkpoints were significantly different between the different risk groups (Figure 7d). The stromal score, immune score, and ESTIMATE score were significantly correlated with the risk score (cor > |0.3|, $p < 0.05$) (Figure 7e).

Regulatory Network of Prognostic Genes

In the TMB analysis, TP53 and TTN had the highest mutation frequencies among the risk groups, and among mutation types, missense mutation had the highest frequency (Figure 8a and b). The K-M curves revealed significant differences in TMB between risk groups (Figure 8c). ABT.888, an AKT inhibitor, was identified based on the GDSC database. VIII and AMG.706 were among the top 10 drugs most effective for treating ESCA patients (Figure 8d). We also investigated the regulatory mechanisms of the prognostic genes. First, 16 miRNAs were predicted, and based on these results, 75 lncRNAs corresponding to miRNAs were obtained. The regulatory relationships contained in the lncRNA–miRNA-mRNA network formed by them with prognostic genes were CYTOR-hsa-miR-2467-3P-*FABP3*, A1BG-AS1-hsa-miR-450b-5p-*FAM189A2*, etc. (Figure 8e). In total, 45 TFs were obtained from the ENCODE database, and they formed relationship pairs, including *FAM189A2*-EZH2 and *OSM*-E2F5 (Figure 8f).

Selection, Cell Communication and Pseudotemporal Analysis of Key Cells

After quality control of the single-cell dataset, the first 2000 highly variable genes were screened (Figure 9a). In the principal component analysis (PCA), the first two principal components (PCs) in the tumor and normal samples could be fused, indicating that batch effects had little impact (Figure 9b). Moreover, the top 40 PCs were screened for subsequent analyses (Figure 9c). Then, the cells were divided into 14 cell clusters (Figure 9d). In total, 11 cell types, including neutrophils and fibroblasts, were annotated based on the marker genes (Figure 9e). Among them, significant differences were found in the expression of *OSM* and *MICB* in mast cells and neutrophils between tumor and normal samples (Figure 9f). Therefore, mast cells and neutrophils were used as key cells for subsequent analysis.

Based on the results of the cell communication analysis, mast cells and neutrophils presented greater interactions, as well as, greater numbers of cells in normal samples than in disease samples (Figure 10a). These cells showed the strongest communication interaction relationship among different groups, with the ligand-receptor pairs being ANXA1-FPR1 (Figure 10b). In the pseudotemporal analysis, neutrophils and mast cells differentiated into three and seven different states, respectively (Figure 10c).

Verification of the Expression of Prognostic Genes

We analyzed the expression of prognostic genes by qRT-PCR analysis, and the results revealed that *MICB* and *FAM189A2* were significantly overexpressed in disease samples, whereas, *OSM* showed the opposite pattern of expression. Additionally, no significant difference was found in the level of expression of *FABP3* between the control and disease samples (Figure 11).

Discussion

ESCC is the predominant histological subtype of esophageal cancer. ESCC cases account for about 90% of all esophageal cancers in developing countries.[25] ESCC is associated with a poor prognosis and high mortality.[26] Crotonylation is a newly identified epigenetic modification and plays a role in various diseases and biological processes, such as carcinogenesis, tissue injury, and inflammation.[27–29] However, the involvement of crotonylation in ESCC is not clear. Therefore, in this study, we identified CRGs in ESCC, established a CRG-related prognostic signature for ESCC, and assessed the clinical predictive value of this signature. By conducting bioinformatics analysis, we identified a novel signature comprising four CRGs (*OSM, FABP3, MICB,* and *FAM189A2*) to predict the prognosis of patients with ESCC.

OSM, a member of the interleukin-6 family of cytokines, plays important roles in several physiological and pathological processes, including mesenchymal stem cell differentiation, fibrosis, hematopoiesis, metabolism, and cancer.[30] Several studies have shown that *OSM* is an ESCC-related gene and its level is negatively correlated with
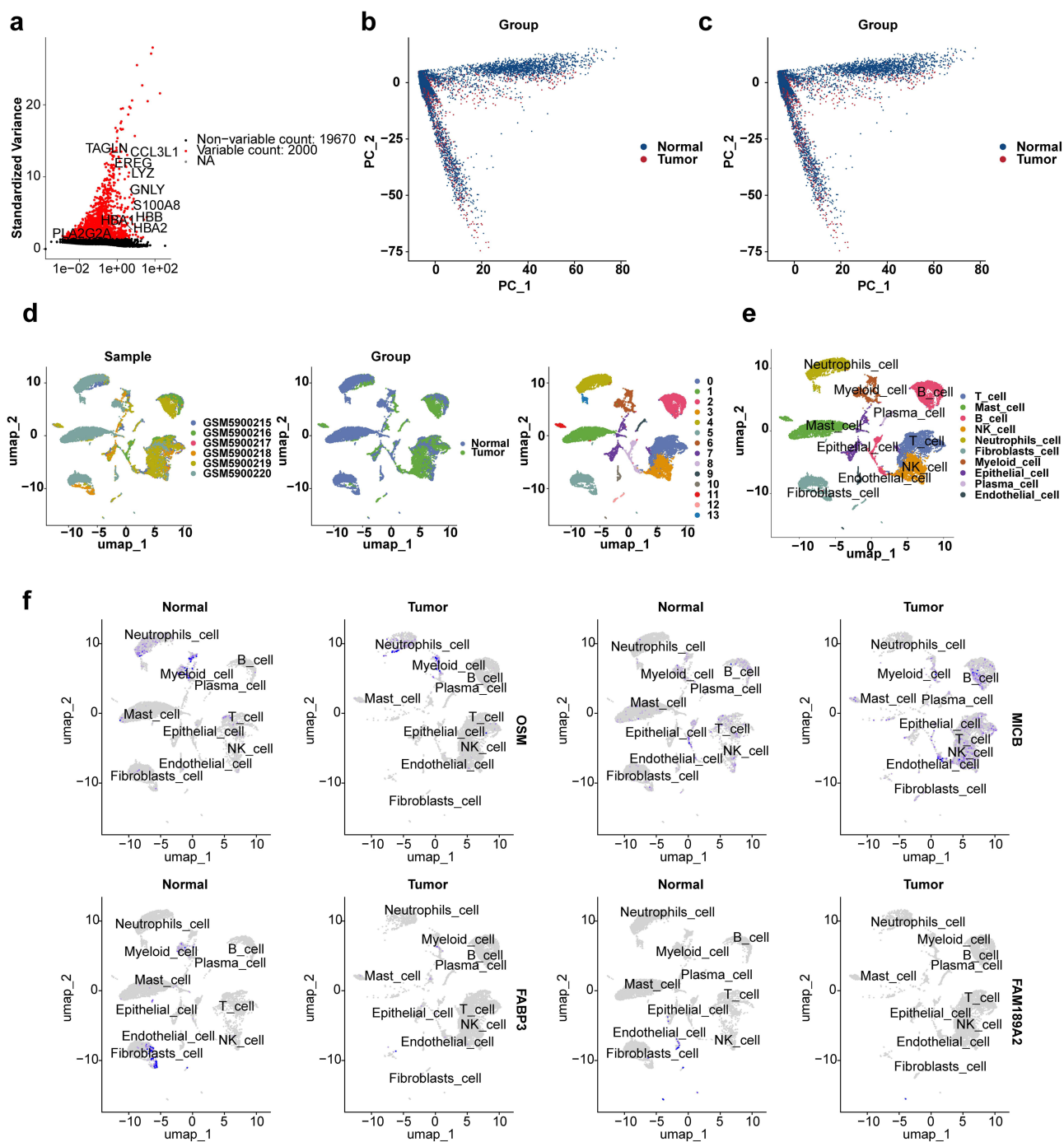
Figure 8 Construction of the regulatory network of *OSM, FABP3, MICB*, and *FAM189A2*. (**a** and **b**) TMB analysis of the mutation status of the genes in the low-risk and high-risk groups, with the top 20 mutated genes in each group are illustrated. (**c**) The differences in TMB between the low-risk and high-risk groups were assessed using K-M curves. (**d**) The IC50 values of the common chemotherapeutic drugs for each sample were considered to screen drugs that are effective for treating ESCA patients. "*" represented *p* < 0.05, "**" represented *p* < 0.01. (**e**) Construction of the complete lncRNA–miRNA–mRNA network of *OSM, FABP3, MICB*, and *FAM189A2*. (**f**) Construction of the transcription factor–mRNA network of *OSM, FABP3, MICB*, and *FAM189A2*.

Figure 9 Mast cells and neutrophils were identified as key cells. (**a**) Highly variable genes of *OSM, FABP3, MICB*, and *FAM189A2* were screened in the single-cell dataset (GSE196756). (**b** and **c**) PCA of GSE196756 was performed for dimensionality reduction. (**d**) UMAP clustering analysis of cells was conducted (the cells were divided into 14 clusters). (**e**) Annotation of cell clusters as different cell types based on the marker genes. (**f**) The key cells (mast cells and neutrophils) were screened by analyzing the expression of *OSM, FABP3, MICB*, and *FAM189A2* in different cell types in ESCA and normal samples.
Abbreviation: PCA, principal component analysis.

risk of ESCC.[31,32] Like previous studies, this study characterized *OSM* as a CRG for the prognosis of ESCC and found that it was downregulated in ESCC samples. *FABP3* is a member of a large family of fatty acid-binding proteins that are expressed in a tissue-specific manner.[33] *FABP3* is highly expressed in aged skeletal muscles that exhibit loss of muscle mass and function, disrupting homeostasis through lipid remodeling; therefore, it is a valuable target for sarcopenia.[34] Another study suggested that high expression of *FABP3* may be responsible for health problems related to cardiovascular
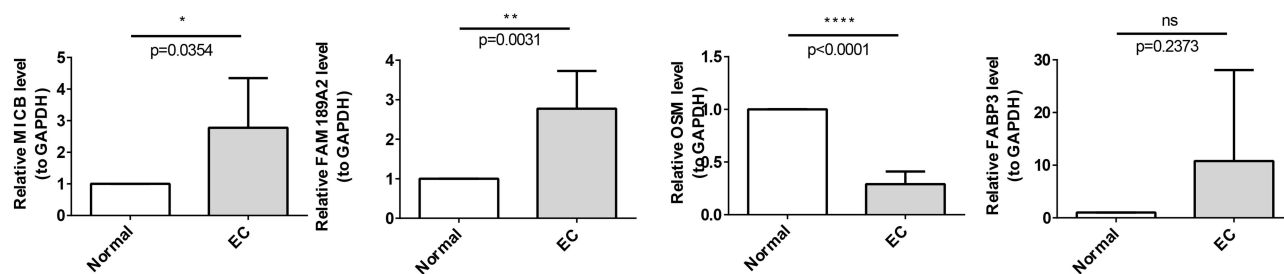
Figure 10 Analyses of mast cells and neutrophils. (**a** and **b**) Cell communication analysis of cell types, showing strong interactions between mast cells and neutrophils and the ligand-receptor pair of ANXA1–FPR1. (**c**) Pseudotemporal analysis of the differentiation status of mast cells and neutrophils.

Figure 11 Verification of the expression of *OSM, FABP3, MICB*, and *FAM189A2* in normal and EC samples. The expression of *OSM, FABP3, MICB*, and *FAM189A2* was measured in normal and EC samples by qRT-PCR analysis. "*" represented $p < 0.05$, "**" represented $p < 0.01$, "***" represented $p < 0.001$, "****" represented $p < 0.0001$.
Abbreviation: EC, Esophageal carcinoma.

disease in patients who succumb to chronic schizophrenia.[35] The expression of *FABP3* is upregulated in ESCC cases; additionally, cell function and immunohistochemical staining assays suggest that *FABP3* functions as an oncogene in ESCC by promoting the malignant progression of ESCC cells.[36] In contrast, this study found no significant difference in *FABP3* expression between control and ESCC samples. This difference may be due to the detection methods used. Different primers and different qRT-PCR methods (TaqMan qRT-PCR) should be used to detect the expression level of *FABP3*. In addition to qRT-PCR, RNA-Fluorescence in situ hybridization (FISH) and other techniques should also be applied to observe the differential expression of *FABP3* at the tissue level. Besides, the detection of protein level can better reflect the relationship between *FABP3* level and ESCC disease. Methods similar like Western blot and immuno-fluorescence (IF) should be applied to detect *FABP3* content to determine its relationship with ESCC. *MICB* belongs to the major histocompatibility complex class I chain-related gene B family, and its expression is induced in response to cellular stress.[37] As a result, *MICB* is expressed by different types of human cancer.[38] In line with our results, ESCC tissues are characterized by the significant overexpression of *MICB* compared to its expression in adjacent normal tissues.[39] *FAM189A2* is a unique activator of HECT-type ubiquitin E3 ligases that can regulate multiple aspects of cellular function by ubiquitinating various substrates.[40] The prognosis of prostate cancer patients may be related to the abnormal expression of *FAM189A2* since its expression was lower in the high-risk score group than in the low-risk score group.[41] Additionally, the expression of *FAM189A2* was found to be lower in endometrial carcinoma samples than in normal samples, whereas high expression of *FAM189A2* predicted long overall survival in patients with endometrial carcinoma.[42] Despite these findings, the expression and prognostic value of *FAM189A2* in ESCC remain undetermined. This study provided the first evidence suggesting abundant expression of *FAM189A2* in ESCC samples and identifying *FAM189A2* as a potential CRG related to the prognosis of ESCC. Thus, these results may aid in the development of biomarkers for the early detection of ESCC.

By conducting functional enrichment analysis, we found that *FABP3* and *FAM189A2* were significantly enriched in neuroactive ligand-receptor interactions and the ribosome signaling pathway. *FABP3* and *MICB* dominated the olfactory transduction signaling pathway. Additionally, *OSM* was predominant in the chemokine signaling pathway and cytokine-cytokine receptor interaction. Neuroactive ligand-receptor interactions are involved in the initiation and progression of several diseases, such as pulmonary arterial hypertension and colon cancer.[43,44] Sevoflurane imparts cardioprotection in patients following coronary artery bypass graft, partly by regulating neuroactive ligand-receptor interactions.[45] Genes involved in the regulation of neuroactive ligand–receptor interactions are highly expressed in esophageal cancer.[46] Additionally, the pathway is associated with significant ratios of target genes of *Hedyotis diffusa* Willd. for treating esophageal adenocarcinoma.[47] The ribosome is responsible for protein synthesis in all cells.[48,49] An increase in ribosome biogenesis contributes to aberrant increases in nucleolar size and number, which are hallmarks of cancer and are linked to a poor prognosis of diseases, including ESCC.[50] According to a recent study, olfactory transduction was identified as the top signaling pathway enriched with genes exclusively expressed in TCGA-ESCC patient tumor tissues.[51] Moreover, the hypomethylated and highly expressed genes in the tissues of ESCC patients are associated primarily with the cytokine-cytokine receptor interaction pathway.[52,53] Several studies have provided evidence that the activation of chemokines greatly contributes to aggressive tumor behavior in ESCC.[54–57] Based on these findings, targeting these pathways may

have a prognostic value and therapeutic potential in ESCC. In summary, our model achieves higher accuracy through a unique feature selection approach that combines single-factor Cox regression and LASSO, effectively identifying biomarkers with strong prognostic relevance. Secondly, we conducted both internal and external validation to ensure the model's reliability. Internal validation evaluates the model's stability and accuracy within the constructed dataset, mitigating overfitting risks, while external validation assesses its generalizability using independent datasets. This dual-validation approach enhances the model's robustness compared to those relying on limited or single validation methods. Additionally, we utilized tools such as nomograms, calibration curves, and ROC curves to comprehensively assess and verify the model's predictive accuracy. The integration of these validation tools provides a thorough evaluation of its performance. In summary, our prognostic model distinguishes itself through its robust construction, rigorous validation processes, and superior predictive capabilities compared to existing models.

Crotonylated X-linked gene 2(BEX2) plays a key role in lung cancer cells by promoting mitochondrial autophagy to inhibit drug-induced apoptosis during chemotherapy. Mu et al[58] found that brain expressed BEX2 is overexpressed in lung adenocarcinoma and is associated with a poor prognosis in patients with cancer without lymph node metastasis. Therefore, drug therapy targeting BEX2-induced mitosis in combination with anticancer drugs may be an effective strategy for treating non-small cell lung cancer. Jiang et al[59] constructed nine genetic risk profiles associated with the prognosis of head and neck squamous cell carcinoma (HNSCC) based on a regulator of Kcr. The risk profile model predicted five-year OS more accurately than traditional clinical parameters (age, sex, T stage, N stage, and histological grade). Thus, it may provide a more scientific basis for the stratification of prognosis in patients with HNSCC and promote the optimization and upgrading of treatment regimens. A review of the relevant studies revealed that crotonylation is an important post-translational protein modification mode that probably plays a pivotal role in the prognostic assessment of diseases and the formulation of treatment strategies. This discovery not only broadens our understanding of the mechanism of protein function regulation but also provides a strong scientific basis to explore new methods for disease treatment.

NKT cells are large granular lymphocytes of the innate immune system.[60,61] They can produce cytokines and chemokines that modulate the immune response and play pivotal roles in cancer, autoimmunity, and inflammation.[62] γδT cells are a specialized subset of T lymphocytes that can directly recognize and kill target cells independently of human leukocyte antigen presentation; thus, they are a promising effector immune cell type for cancer immunotherapy.[63,64] Among the different types of innate immune cells involved in esophageal cancer progression, NKT and γδT cells have antitumor functions because of their strong cytotoxicity.[65] The immune analysis results of this study revealed that both *MICB* and *FABP3* had the greatest positive associations with NKT cells, whereas, *FAM189A2* had the greatest negative association with γδT cells. These findings revealed the predictive value of CRGs in ESCC. Studies[66] have shown that the level of expression of *OSM* and the mutation status of the TP53 gene play key roles in regulating the infiltration of immune cells into tumor tissues and are directly related to the survival prognosis of patients with cholangiocarcinoma. Another study revealed that *MICB*, as a key molecule in the activation of NK cells, promotes the recognition and killing of tumor cells by NK cells by binding with NKG2D and is an important mechanism underlying the function of the immune system against cancer metastasis.[67] Additionally, low-dose radiotherapy (LDR) influences the destruction of the immunosuppressive tumor microenvironment (TME). *FABP3* is a key molecule whose expression level influences tumor cell sensitivity to LDR and the subsequent immunotherapy response.[68] Jiakun Liu et al reported that the *FAM189A2* gene plays a role as a tumor suppressor in lung adenocarcinoma, which is achieved by regulating tight junction protein (TJP) and CXCR4. Moreover, *FAM189A2* is strongly associated with the immune microenvironment of lung adenocarcinoma, which may be related to patient prognosis and immunotherapeutic effects.[69]

Subsequent single-cell RNA-seq analysis based on the CRGs suggested that mast cells and neutrophils are key cells in ESCC. Mast cells are tissue-resident effector cells of immune responses; they accumulate in the tumor stroma of numerous types of human cancer, and an increase in mast cell density is associated with a good or poor prognosis, which depends on the tumor type and stage.[70] High mast cell density at the invasive edge of ESCC is responsible for aggressive tumor behavior and a decrease in the survival rate of patients.[71] Neutrophils, characterized as short-lived effector cells, elicit cancer initiation, progression, and metastasis through direct effects on cancer cells or indirect effects

on the TME.[72,73] An increase in the number of neutrophils after neoadjuvant therapy is associated with a poor prognosis in patients with esophageal cancer.[74] Therefore, depletion of these cells or inhibition of their function can be assessed as therapeutic strategies for ESCC.

However, this study had several limitations. Model development based on the TCGA database is limited to covering the genetic and environmental complexity and diversity of ESCC, which affects the wide applicability and representativeness of the model to a certain extent. Due to the lack of direct clinical validation, the reliability and validity of the model need to be confirmed. Although we evaluated the efficacy of the model via external datasets and verified differences in the expression of prognostic genes between the disease and control groups by qRT-PCR, these results did not fully reveal the actual expression of prognostic genes in the clinical setting. To overcome these limitations, we aim to strengthen the following aspects in future studies: first, expand the scale and diversity of clinical validation and evaluate the applicability of the model across patient populations; second, investigate the role of genes in ESCC in detail to improve the scientific model. Moreover, in vitro cell experiments and in vivo animal model experiments need to be conducted to better understand the effect of these prognostic genes on ESCC cell behavior and tumor growth.

Conclusion

Four genes (*OSM, FABP3, MICB*, and *FAM189A2*) were identified as prognostically crotonylation related genes for ESCC and might be involved in its pathogenesis. Specifically, *OSM* may negatively regulate crotonylation, thereby inhibiting cancer progression. *FABP3* and *MICB* may be positively related to crotonylation and promote tumor progression as ESCC marker genes. In this study, *FAM189A2* is also considered to be a potential CRG related to ESCC prognosis, but its content in ESCC samples did not prove statistical difference, and its role in regulating ESCC remains to be studied. In addition to providing biomarker diagnostic genes for clinical diagnosis of ESCC, our study also suggests basic research to explore the role and mechanism of crotonylation-related genes in ESCC, including these four genes in our results.

In conclusion, these findings can help optimize the risk prediction model and provide stronger support for precision medicine for ESCC.

Abbreviations

ESCC, esophageal squamous cell cancer; CRGs, crotonylation-related genes; DEGs, differentially expressed genes; EC, Esophageal carcinoma; Kpr, histone propionylation; Kbu, butylation, Khib, 2-hydroxyisobutylation; Ksucc, succinylation; Kma, malondialdehyde; Kglu, glutarylation; Kcr, crotonylation; Kbhb, beta-hydroxybutylation; Kbz, benzene formylation; Acox2, acyl-CoA oxidase 2;TCA, tricarboxylic acid; TCGA, the Cancer Genome Atlas; CNV, copy number variation; OS, overall survival; ROC, receiver operating characteristic; TMB, Tumor mutational burden; TFs, transcription factors; AUCs, The areas under the ROC curves; BEX2, Crotonylated X-linked gene 2; HNSCC, head and neck squamous cell carcinoma; LDR, low-dose radiotherapy; TME, tumor microenvironment; TJP, tight junction protein; LASSO, least absolute shrinkage and selection operator; NKT, natural killer T; γδT, gamma delta T; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PCA, principal component analysis.

Ethics Approval and Informed Consent

The study was approved by the Ethics Committee of the first affiliated hospital of Wenzhou Medical University. All studies were carried out in accordance with the principles stipulated in the Declaration of Helsinki, and all patients had signed an informed consent form.

Disclosure

The author(s) report no conflicts of interest in this work.

References

1. Sung H, Ferlay J, Siegel RL. et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca a Cancer J Clinicians*. 2021;71:209–249. doi:10.3322/caac.21660

2. Qi J, Li M, Wang L, et al. National and subnational trends in cancer burden in China, 2005–20: an analysis of national mortality surveillance data. *Lancet Public Health*. 2023;8:e943–e955. doi:10.1016/s2468-2667(23)00211-6

3. Arnold M, Ferlay J, van Berge Henegouwen MI, Soerjomataram I. Global burden of oesophageal and gastric cancer by histology and subsite in 2018. *Gut*. 2020;69:1564–1571. doi:10.1136/gutjnl-2020-321600

4. Nie Y, Yao G, Xu X, et al. Single-cell mapping of N6-methyladenosine in esophageal squamous cell carcinoma and exploration of the risk model for immune infiltration. *Front Endocrinol*. 2023;14:1155009. doi:10.3389/fendo.2023.1155009

5. Chen Y, Sprung R, Tang Y, et al. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *mol Cell Proteomics*. 2007;6:812–819. doi:10.1074/mcp.M700021-MCP200

6. Zhao D, Li Y, Xiong X, Chen Z, Li H. YEATS Domain-A histone acylation reader in health and disease. *J mol Biol*. 2017;429:1994–2002. doi:10.1016/j.jmb.2017.03.010

7. Sabari BR, Tang Z, Huang H, et al. Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol Cell*. 2015;58:203–215. doi:10.1016/j.molcel.2015.02.029

8. Tan M, Luo H, Lee S, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*. 2011;146:1016–1028. doi:10.1016/j.cell.2011.08.008

9. Wang S, Mu G, Qiu B, et al. The function and related diseases of protein crotonylation. *Int J Bio Sci*. 2021;17:3441–3455. doi:10.7150/ijbs.58872

10. Zhang Y, Chen Y, Zhang Z, et al. Acox2 is a regulator of lysine crotonylation that mediates hepatic metabolic homeostasis in mice. *Cell Death Dis*. 2022;13. doi:10.1038/s41419-022-04725-9

11. Zheng Y, Zhu L, Qin Z-Y, et al. Modulation of cellular metabolism by protein crotonylation regulates pancreatic cancer progression. *Cell Rep*. 2023;42:112666. doi:10.1016/j.celrep.2023.112666

12. Yuan H, Wu X, Wu Q, et al. Lysine catabolism reprograms tumour immunity through histone crotonylation. *Nature*. 2023;617:818–826. doi:10.1038/s41586-023-06061-0

13. Liao M, Sun X, Zheng W, et al. LINC00922 decoys SIRT3 to facilitate the metastasis of colorectal cancer through up-regulation the H3K27 crotonylation of ETS1 promoter. *mol Cancer*. 2023;22:163. doi:10.1186/s12943-023-01859-y

14. Hänzelmann S, Castelo R, Guinney J, et al. GSVA gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. 2013;16:7. doi:10.1186/1471-2105-14-7

15. 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

16. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. doi:10.1186/s13059-014-0550-8

17. Wu T, Hu E, Xu S, et al. ClusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2:100141. doi:10.1016/j.xinn.2021.100141

18. Wu J, Li L, Zhang H, et al. A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma. *Oncogene*. 2021;40:4413–4424. doi:10.1038/s41388-021-01853-y

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

20. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–2504. doi:10.1101/gr.1239303

21. Tan Z, Chen X, Zuo J, et al. Comprehensive analysis of scRNA-Seq and bulk RNA-Seq reveals dynamic changes in the tumor immune microenvironment of bladder cancer and establishes a prognostic model. *J Transl Med*. 2023;21:223. doi:10.1186/s12967-023-04056-z

22. 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

23. Ni L, Yang H, Wu X, et al. The expression and prognostic value of disulfidptosis progress in lung adenocarcinoma. *Aging*. 2023;7:7741–7759. doi:10.18632/aging.204938

24. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39:175–191. doi:10.3758/bf03193146

25. An L, Li M, Jia Q. Mechanisms of radiotherapy resistance and radiosensitization strategies for esophageal squamous cell carcinoma. *mol Cancer*. 2023;22:104. doi:10.1186/s12943-023-01839-2

26. Yang T, Hui R, Nouws J, et al. Untargeted metabolomics analysis of esophageal squamous cell cancer progression. *J Transl Med*. 2022;20:127. doi:10.1186/s12967-022-03311-z

27. Jiang G, Li C, Lu M, Lu K, Li H. Protein lysine crotonylation: past, present, perspective. *Cell Death Dis*. 2021;12:703. doi:10.1038/s41419-021-03987-z

28. Yang P, Qin Y, Zeng L, et al. Crotonylation and disease: current progress and future perspectives. *Biomed Pharmacother*. 2023;165:115108. doi:10.1016/j.biopha.2023.115108

29. Ntorla A, Burgoyne JR. The regulation and function of histone crotonylation. *Front Cell Develop Biol*. 2021;9:624914. doi:10.3389/fcell.2021.624914

30. Lantieri F, Bachetti T. OSM/OSMR and interleukin 6 family cytokines in physiological and pathological condition. *Int J mol Sci*. 2022;23:11096. doi:10.3390/ijms231911096

31. Tian LI, Huang Y, Zhang B, et al. Targeting LncRNA LLNLR-299G3.1 with antisense oligonucleotide inhibits malignancy of esophageal squamous cell carcinoma cells in vitro and in vivo. *Oncol Res*. 2023;31:463–479. doi:10.32604/or.2023.028791

32. Aversa J, Song M, Shimazu T, et al. Prediagnostic circulating inflammation biomarkers and esophageal squamous cell carcinoma: a case-cohort study in Japan. *Int J Cancer*. 2020;147:686–691. doi:10.1002/ijc.32763

33. Baru A, Devi C, Mukhopadhyay T, et al. Anti bacterial function of secreted human FABP3. *Biochim Biophys Acta Gen Subj*. 2023;1867:130472. doi:10.1016/j.bbagen.2023.130472

34. Lee S-M, Lee SH, Jung Y, et al. FABP3-mediated membrane lipid saturation alters fluidity and induces ER stress in skeletal muscle with aging. *Nat Commun*. 2020;11:5661. doi:10.1038/s41467-020-19501-6

35. Hung W-C, Yu T-H, Wu C-C, et al. FABP3, FABP4, and heart rate variability among patients with chronic schizophrenia. *Front Endocrinol*. 2023;14:1165621. doi:10.3389/fendo.2023.1165621

36. Liu T, Fang P, Han C, et al. Four transcription profile–based models identify novel prognostic signatures in oesophageal cancer. *J Cell Mol Med*. 2019;24:711–721. doi:10.1111/jcmm.14779

37. Jumnainsong A, Jearanaikoon P, Khahmahpahte S, et al. Associations of MICB with cervical cancer in north-eastern Thais: identification of major histocompatibility complex class I chain-related gene B motifs influencing natural killer cell activation. *Clin Exp Immunol*. 2008;153:205–213. doi:10.1111/j.1365-2249.2008.03682.x

38. Ferrari de Andrade L, Tay RE, Pan D, et al. Antibody-mediated inhibition of MICA and MICB shedding promotes NK cell–driven tumor immunity. *Science*. 2018;359:1537–1542. doi:10.1126/science.aao0505

39. Lim KS, Mimura K, Kua LF, et al. Implication of highly cytotoxic natural killer cells for esophageal squamous cell carcinoma treatment. *J Immunother*. 2018;41:261–273. doi:10.1097/CJI.0000000000000227

40. Tsunoda T, Riku M, Yamada N, et al. ENTREP/FAM189A2 encodes a new ITCH ubiquitin ligase activator that is downregulated in breast cancer. *EMBO Rep*. 2021;23:e51182. doi:10.15252/embr.202051182

41. Luo R, Huang M, Wang Y, Yang WH. A transcription factor-based risk model for predicting the prognosis of prostate cancer and potential therapeutic drugs. *Evid Based Complement Alternat Med*. 2021;2021:1–14. doi:10.1155/2021/6894278

42. Wang A, Guo H, Long Z, Huang T. Integrative analysis of differently expressed genes reveals a 17-gene prognosis signature for endometrial carcinoma. *Biomed Res Int*. 2021;2021:1–18. doi:10.1155/2021/4804694

43. Xiao G, Wang T, Zhuang W, et al. RNA sequencing analysis of monocrotaline-induced PAH reveals dysregulated chemokine and neuroactive ligand receptor pathways. *Aging*. 2020;12:4953–4969. doi:10.18632/aging.102922

44. Yang Y, et al. Inhibition of neuroactive ligand-receptor interaction pathway can enhance immunotherapy response in colon cancer: an in silico study. *Expert Rev Anticancer Ther*. 2023;23:1205–1215. doi:10.1080/14737140.2023.2245567

45. Wang J, Cheng J, Zhang C, Li X. Cardioprotection effects of sevoflurane by regulating the pathway of neuroactive ligand-receptor interaction in patients undergoing coronary artery bypass graft surgery. *Comput Math Methods Med*. 2017;2017:1–7. doi:10.1155/2017/3618213

46. Chattopadhyay I, Phukan R, Singh A, et al. Molecular profiling to identify molecular mechanism in esophageal cancer with familial clustering. *Oncol Rep*. 2009;21:1135–1146. doi:10.3892/or_00000333

47. Zhuang Y, Sun Y-G, Wang C-G, et al. Molecular targets and mechanisms of *Hedyotis diffusa* Willd. for esophageal adenocarcinoma treatment based on network pharmacology and weighted gene co-expression network analysis. *Curr Drug Targets*. 2024;11:431–443. doi:10.2174/0113894501265851240102101122

48. Pelletier J, Thomas G, Volarević S, et al. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nat Rev Cancer*. 2018;18:51–63. doi:10.1038/nrc.2017.104

49. Yang Y-M, Jung Y, Abegg D, et al. Chaperone-directed ribosome repair after oxidative damage. *Molecular Cell*. 2023;83:1527–1537.e1525. doi:10.1016/j.molcel.2023.03.030

50. Zhao Y, Li Y, Zhu R, et al. RPS15 interacted with IGF2BP1 to promote esophageal squamous cell carcinoma development via recognizing m6A modification. *Signal Transduct Target Ther*. 2023;8:224. doi:10.1038/s41392-023-01428-1

51. Ma F, Laster K, Nie W, et al. Heterogeneity analysis of esophageal squamous cell carcinoma in cell lines, tumor tissues and patient-derived xenografts. *J Cancer*. 2021;12:3930–3944. doi:10.7150/jca.52286

52. Sang L, Yu Z, Wang A, et al. Identification of methylated-differentially expressed genes and pathways in esophageal squamous cell carcinoma. *Pathol Res Pract*. 2020;216:153050. doi:10.1016/j.prp.2020.153050

53. Chen X, Cai S, Wang L, et al. Analysis of the function of MAGE-A in esophageal carcinoma by bioinformatics. *Medicine*. 2019;98:e15774. doi:10.1097/md.0000000000015774

54. Baba Y, Iwatsuki M, Yoshida N, Watanabe M, Baba H. Review of the gut microbiome and esophageal cancer: pathogenesis and potential clinical implications. *Ann Gastroenterol Surg*. 2017;1:99–104. doi:10.1002/ags3.12014

55. Chen J, Zhu Y, Zhao D, et al. Co-targeting FAK and Gli1 inhibits the tumor-associated macrophages-released CCL22-mediated esophageal squamous cell carcinoma malignancy. *MedComm*. 2023;4:e381. doi:10.1002/mco2.381

56. Wu X, Zhang H, Sui Z, et al. CXCR4 promotes the growth and metastasis of esophageal squamous cell carcinoma as a critical downstream mediator of HIF-1α. *Cancer Sci*. 2022;113:926–939. doi:10.1111/cas.15265

57. Fang L, Che Y, Zhang C, et al. PLAU directs conversion of fibroblasts to inflammatory cancer-associated fibroblasts, promoting esophageal squamous cell carcinoma progression via uPAR/Akt/NF-κB/IL8 pathway. *Cell Death Discovery*. 2021;7:32. doi:10.1038/s41420-021-00410-6

58. Mu N, Wang Y, Li X, et al. Crotonylated BEX2 interacts with NDP52 and enhances mitophagy to modulate chemotherapeutic agent-induced apoptosis in non-small-cell lung cancer cells. *Cell Death Dis*. 2023;14:645. doi:10.1038/s41419-023-06164-6

59. Jiang L, Yin X, Zhang H, et al. Development and validation of a prognostic signature based on the lysine crotonylation regulators in head and neck squamous cell carcinoma. *Biomed Res Int*. 2023;2023:4444869. doi:10.1155/2023/4444869

60. Gapin L, Godfrey DI, Rossjohn J. Natural Killer T cell obsession with self-antigens. *Curr Opin Immunol*. 2013;25:168–173. doi:10.1016/j.coi.2013.01.002

61. Mundy-Bosse BL, Scoville SD, Chen L, et al. MicroRNA-29b mediates altered innate immune development in acute leukemia. *J Clin Investig*. 2016;126:4404–4416. doi:10.1172/jci85413

62. Singh AK, Gaur P, Das SN, et al. Natural killer T cell anergy, co-stimulatory molecules and immunotherapeutic interventions. *Hum Immunol*. 2014;75:250–260. doi:10.1016/j.humimm.2013.12.004

63. Yan W, Dunmall LSC, Lemoine NR, et al. The capability of heterogeneous γδ T cells in cancer treatment. *Front Immunol*. 2023;14:1285801. doi:10.3389/fimmu.2023.1285801

64. Saura-Esteller J, de Jong M, King LA, et al. Gamma delta T-cell based cancer immunotherapy: past-present-future. *Front Immunol*. 2022;13:915837. doi:10.3389/fimmu.2022.915837

65. Cui K, Hu S, Mei X, Cheng M. Innate immune cells in the esophageal tumor microenvironment. *Front Immunol*. 2021;12:654731. doi:10.3389/fimmu.2021.654731

66. Liu Q, Lan T, Song Y, et al. Oncostatin M expression and TP53 mutation status regulate tumor-infiltration of immune cells and survival outcomes in cholangiocarcinoma. *Aging*. 2020;12:21518–21543. doi:10.18632/aging.103936

67. Maurer S, Ferrari de Andrade L. NK cell interaction with platelets and myeloid cells in the tumor milieu. *Front Immunol*. 2020;11:608849. doi:10.3389/fimmu.2020.608849

68. Zhang Q, Wang X, Zhao Y, et al. Nanointegrative in situ reprogramming of tumor-intrinsic lipid droplet biogenesis for low-dose radiation-activated ferroptosis immunotherapy. *ACS Nano*. 2023;17:25419–25438. doi:10.1021/acsnano.3c08907

69. Liu J, Zhao W, Luan Y, et al. AM189A2 plays a tumour suppressor role in lung adenocarcinoma by influencing cell apoptosis, CXCR4 expression and tight junction proteins. *Tissue Cell*. 2024;89:102441. doi:10.1016/j.tice.2024.102441

70. Aponte-López A, Muñoz-Cruz S. Mast cells in the tumor microenvironment. *Adv Exp Med Biol*. 2020;1273:159–173. doi:10.1007/978-3-030-49270-0_9

71. Fakhrjou A, Niroumand-Oscoei SM, Somi MH, et al. Prognostic value of tumor-infiltrating mast cells in outcome of patients with esophagus squamous cell carcinoma. *J Gastrointest Cancer*. 2014;45:48–53. doi:10.1007/s12029-013-9550-2

72. Treffers LW, Hiemstra IH, Kuijpers TW, et al. Neutrophils in cancer. *Immunol Rev*. 2016;273:312–328. doi:10.1111/imr.12444

73. Ng MSF, Tan L, Wang Q, et al. Neutrophils in cancer-unresolved questions. *Sci China Life Sci*. 2021;64:1829–1841. doi:10.1007/s11427-020-1853-4

74. Cabalag CS, Prall OWJ, Ciciulla J, et al. Tumor-infiltrating neutrophils after neoadjuvant therapy are associated with poor prognosis in esophageal cancer. *Ann Surg Oncol*. 2022;30:1614–1625. doi:10.1245/s10434-022-12562-5