

Development and validation of a prognostic prediction model for endometrial cancer based on CD8⁺ T cell infiltration-related genes

Chao Chen, MM^a, Lipeng Pei, MM^a, Wei Ren, MM^{a,*} , Jingli Sun, MM^a

Abstract

Endometrial cancer (EC) is the most common gynecologic malignancy with increasing incidence and mortality. The tumor immune microenvironment significantly impacts cancer prognosis. Weighted Gene Co-Expression Network Analysis (WGCNA) is a systems biology approach that analyzes gene expression data to uncover gene co-expression networks and functional modules. This study aimed to use WGCNA to develop a prognostic prediction model for EC based on immune cell infiltration, and to identify new potential therapeutic targets. WGCNA was performed using the Cancer Genome Atlas Uterine Corpus Endometrial Carcinoma dataset to identify hub modules associated with T-lymphocyte cell infiltration. Prognostic models were developed using LASSO regression based on genes in these hub modules. The Search Tool for the Retrieval of Interacting Genes/Proteins was used for protein-protein interaction network analysis of the hub module. Gene Set Variation Analysis identified differential gene enrichment analysis between high- and low-risk groups. The relationship between the model and microsatellite instability, tumor mutational burden, and immune cell infiltration was analyzed using The Cancer Genome Atlas data. The model's correlation with chemotherapy and immunotherapy resistance was examined using the Genomics of Drug Sensitivity in Cancer and Cancer Immunome Atlas databases. Immunohistochemical staining of EC tissue microarrays was performed to analyze the relationship between the expression of key genes and immune infiltration. The green-yellow module was identified as a hub module, with 4 genes (ARPC1B, BATF, CCL2, and COTL1) linked to CD8⁺ T cell infiltration. The prognostic model constructed from these genes showed satisfactory predictive efficacy. Differentially expressed genes in high- and low-risk groups were enriched in tumor immunity-related pathways. The model correlated with EC-related phenotypes, indicating its potential to predict immunotherapeutic response. Basic leucine zipper activating transcription factor-like transcription factor (BATF) expression in EC tissues positively correlated with CD8⁺ T cell infiltration, suggesting BATF's crucial role in EC development and antitumor immunity. The prognostic model comprising ARPC1B, BATF, CCL2, and COTL1 can effectively identify high-risk EC patients and predict their response to immunotherapy, demonstrating significant clinical potential. These genes are implicated in EC development and immune infiltration, with BATF emerging as a potential therapeutic target for EC.

Abbreviations: BATF = basic leucine zipper activating transcription factor-like transcription factor, CCL2 = C-C motif chemokine ligand 2, CIBERSORT = Cell-type Identification by Estimating Relative Subsets Of RNA Transcripts, COTL1 = coactosin-like 1, EC = endometrial cancer, GDSC = Genomics of Drug Sensitivity in Cancer, GO = gene ontology, GSVA = Gene Set Variation Analysis, HH = Hedgehog-GLI, IHC = immunohistochemical, KEGG = Kyoto Encyclopedia of Genes and Genomes, MSI = microsatellite instability, MSI-H = high levels of microsatellite instability, OS = overall survival, PPI = protein-protein interaction, ROC = receiver operating characteristic, TCGA = The Cancer Genome Atlas, TCGA-UCEC = The Cancer Genome Atlas Uterine Corpus Endometrial Carcinoma, TCIA = The Cancer Immunome Atlas, TIME = tumor immune microenvironment, TMA = tissue microarrays, TMB = tumor mutational burden, UCEC = Uterine Corpus Endometrial Carcinoma, WGCNA = Weighted Gene Co-Expression Network Analysis.

Keywords: basic leucine zipper activating transcription factor-like transcription factor, CD8⁺ T cell infiltration, endometrial cancer, prognostic prediction model

CC, LP, WR, and JS contributed equally to this work.

The authors have no funding and conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

Ethical approval was not necessary for this study because public datasets were analyzed and all data were de-identified.

Supplemental Digital Content is available for this article.

^a Department of Obstetrics and Gynecology, General Hospital of Northern Theater Command, Shenyang, Liaoning, China.

* Correspondence: Wei Ren, Department of Obstetrics and Gynecology, General Hospital of Northern Theater, Shenyang, Liaoning 110000, China (e-mail: qrenweicc@163.com).

Copyright © 2024 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Chen C, Pei L, Ren W, Sun J. Development and validation of a prognostic prediction model for endometrial cancer based on CD8⁺ T cell infiltration-related genes. *Medicine* 2024;103:49(e40820).

Received: 9 October 2023 / Received in final form: 4 August 2024 / Accepted: 15 November 2024

<http://dx.doi.org/10.1097/MD.0000000000040820>

1. Introduction

Endometrial cancer (EC) is a prevalent gynecological malignancy that poses a significant threat to women's health.^[1] Unlike most cancers, the incidence and mortality rates of EC have been rising.^[2] In 2013, there were 49,560 new cases and 11,350 deaths,^[3] which increased to 65,620 new cases and 12,590 deaths by 2018.^[4] Data from 2007 to 2016 indicate a 1.3% annual increase in EC incidence.^[5] Initially rising obesity rates^[6] were considered a major factor, but recent studies attribute the increase primarily to nonendometrioid adenocarcinoma, which is not linked to obesity.^[7]

The rise in EC mortality rates largely reflects the stagnation in treating recurrent and metastatic tumors.^[8] Early identification of high-risk ECs is crucial for improving treatment outcomes. Although factors like pathological type, stage, grade, and molecular type can indicate high risk, accurately predicting recurrence and metastasis risk in patients remains challenging.^[9] Consequently, gynecologists are seeking better methods for early identification of high risk EC. Advances from the Human Genome Project and the availability of The Cancer Genome Atlas Uterine Corpus Endometrial Carcinoma (TCGA-UCEC) data have facilitated the development of genetic models for early identification of high risk EC. For example, Wang et al^[10] developed a long noncoding RNA signature related to autophagy for predicting EC prognosis, while Jiang et al^[11] and Liu et al^[12] created prognostic models based on glycolysis-related genes.

Tumor immune microenvironment (TIME) is critical for tumor progression and prognosis.^[13–15] Immune cells and factors within the TIME play vital roles in tumor development.^[16] Research shows that TIME is associated with tumor prognosis, chemoresistance, and immunotherapy efficacy.^[17–21] EC tissues, rich in immune cells and factors,^[22] suggest that TIME could be a key area for treatment and diagnosis breakthroughs. CD8⁺ T cells are essential mediators of tumor immunity, responsible for tumor cell killing. Tumor immune tolerance often involves mechanisms that evade CD8⁺ T cell killing, and immunotherapy aims to counteract this tolerance.^[23] Studies by Kondratiev et al,^[24] Workel et al,^[25] and Jong et al^[26] demonstrate that the distribution and activation status of CD8⁺ T cells significantly affect EC prognosis. These findings indicate that CD8⁺ T cells are crucial for EC prognosis and could inform treatment strategies.

Systems biology, which systematically analyzes high-throughput data, helps identify potential biomarkers for cancer diagnosis, prognosis, and treatment response.^[27,28] Weighted Gene Co-Expression Network Analysis (WGCNA) is a novel systems biology method that identifies clusters of highly associated genes, providing systems-level insights. This method can identify important biomarkers and therapeutic targets.^[29,30]

This study aimed to establish a prognostic model of genes related to CD8⁺ T cell infiltration using WGCNA and lasso regression, identify new genes associated with EC prognosis and immune infiltration, and introduce potential therapeutic targets.

2. Methods

2.1. Data source

The mRNA expression data of UCEC were downloaded from the TCGA database (<https://portal.gdc.cancer.gov>), including 35 cases of normal or paracancerous tissue samples and 552 cases of tumor samples. Overall survival (OS) data from tumor samples were downloaded and merged with the mRNA expression data for model construction and subsequent analysis. The immunogenomic information of the samples was downloaded from the Cancer Immunome Atlas (TCIA) database (<https://tcia.at/home>) to estimate immunotherapy efficacy. Gene signatures associated with chemotherapeutic drug sensitivity were obtained

from the Genomics of Drug Sensitivity in Cancer (GDSC) database (<https://www.cancerrxgene.org/>). Ethical approval was not necessary for this study because public datasets were analyzed and all data were de-identified.

2.2. Immuno-infiltration analysis

Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) is an algorithm designed to quantify various cell types from complex tissue samples based on transcriptomic data.^[31] It exhibits high-resolution cell type quantification capabilities, particularly useful for tumor samples where other cell types are present in low abundance.^[32] In this study, we employed the CIBERSORT algorithm to infer the proportions of various immune cells within TCGA-UCEC tumor samples.

The CIBERSORT computations were conducted using the “CIBERSORT R script v1.03.” We used RNA-seq data matrix of TCGA-UCEC tumor samples as the mixture file. The Support Vector Machine training was executed with nu values set to 0.25, 0.5, and 0.75, iterating 3 times. The following parameters were set: Permutations to 1000, Quantile Normalization to False, and a *P*-value cutoff of <.05. The proportions of various types of immune cells in the TCGA-UCEC tumor samples were obtained by calculation, and the proportions of T-lymphocyte infiltration were further extracted for the screening of hub-module.

2.3. Co-expression module construction

In this study, we used processed mRNA expression data from the TCGA UCEC cohort, measured in fragments per kilobase million units. The co-expression network was constructed using the “WGCNA” R package (version: 1.70-3). Initially, we created a matrix of correlation coefficients between genes, which called weighted adjacency matrix. To ensure a scale-free network within the adjacency matrix, we used the “pickSoftThreshold” function from the “WGCNA” R package to choose an appropriate soft thresholding power β (range of 1–20). Then, converted the weighted adjacency matrix to a topological overlay matrix. Subsequently, constructing a hierarchical clustering tree structure from the topological overlay matrix, and using the dynamic tree cut method identifies gene co-expression modules from the hierarchical clustering tree. The minimum number of genes in each gene co-expression module was set as 50, and the cut height threshold for merging similar gene modules was set as 0.25. Finally, Pearson correlation analysis was used to determine the relationship between gene co-expression modules and the percentage of T-lymphocyte infiltration to identify hub-module.

2.4. Model construction

Tumor samples from TCGA-UCEC were randomly divided into training and test sets in a 1:1 ratio. Genes in the hub-module associated with prognosis were identified by lasso regression (10-fold cross-validation), and the association of these genes with survival was assessed by multifactorial COX regression. A prognostic model was constructed with the formula Risk-Score = $\sum(\text{Coefi} * \text{Exp}_i)$, where Coefi represents the risk coefficient, and Exp_i represents the expression of each gene. COX analysis was conducted using the “survival” package, and lasso regression with the “glmnet” package.

2.5. Prognostic effectiveness evaluation

Risk scores for the samples were calculated based on the model, and patients were categorized into high- and low-risk groups based on the median risk score. Kaplan–Meier survival analysis based on OS was performed on the training and test sets, and receiver operating characteristic (ROC) curves based

on OS were plotted to assess the prognostic efficacy of the model using the “survival,” “survminer,” and “pROC” packages, respectively. The area under the ROC curve (1, 3, 5 years) was calculated to evaluate the predictive performance of the model.

2.6. Protein–protein interaction network

To further examine the intrinsic connections of the genes in the hub-module, we performed protein–protein interaction (PPI) network analysis. PPI were retrieved from the Search Tool for the Retrieval of Interacting Genes/Proteins database (version

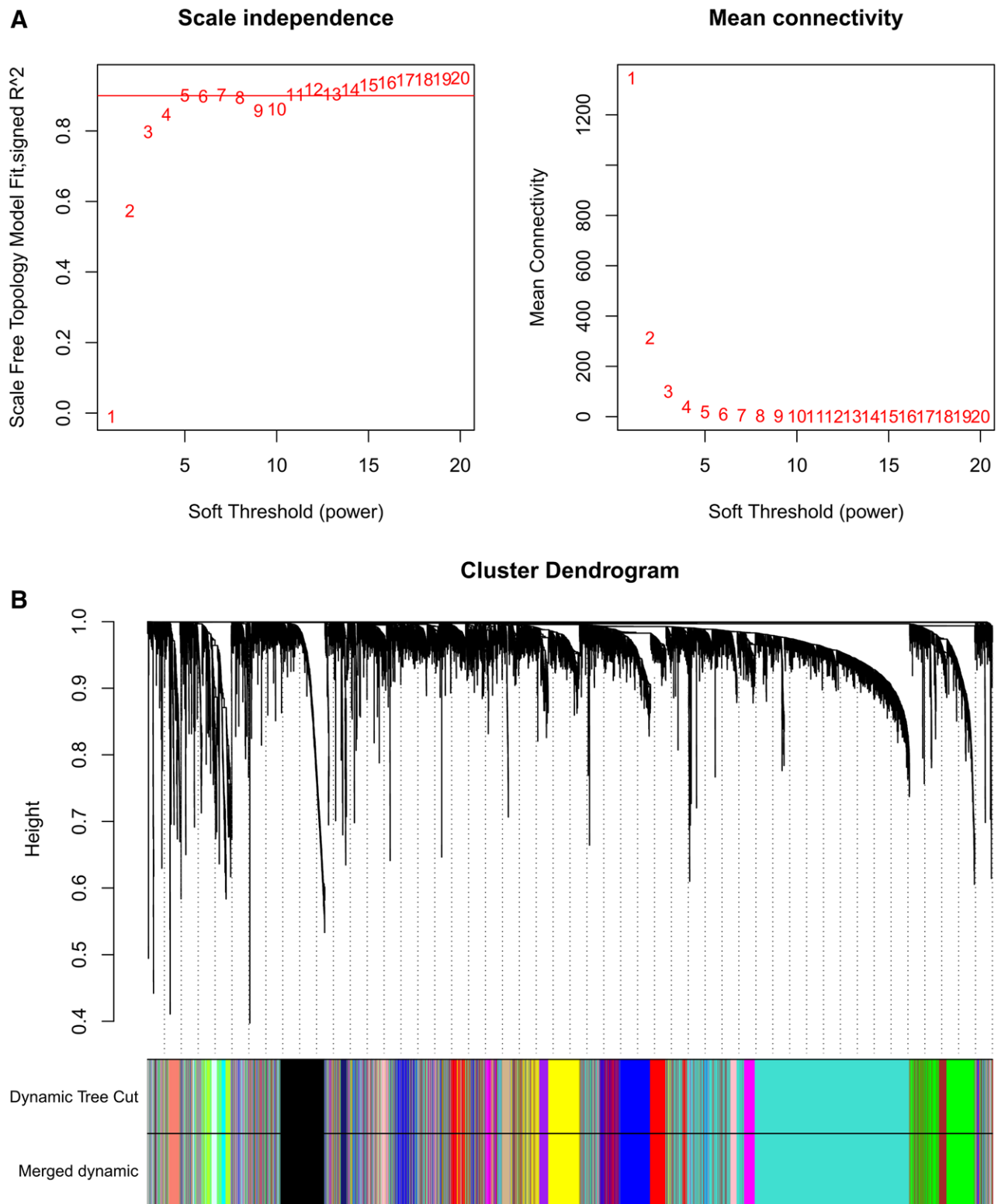


Figure 1. Weighted gene co-expression network analysis (WGCNA). (A) Determination of suitable soft thresholds. (B) Stratified clustering of genes associated with T lymphocyte infiltration. (C) Module-trait relationships. The relevant *P* value and correlation coefficient are listed in each cell. (D) Scatterplot of the correlation between genes in the module and the infiltration ratio of CD8⁺ T cells.

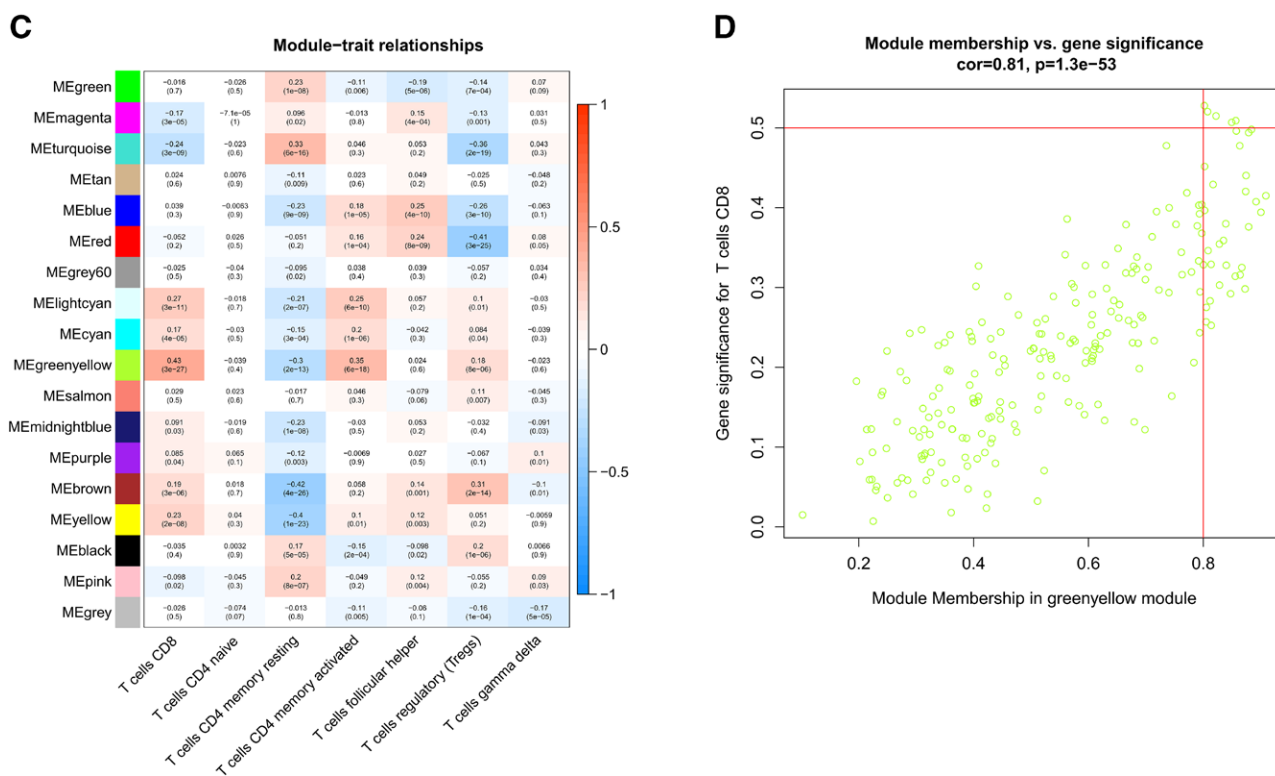


Figure 1. Continued

12.0) (<https://www.string-db.org>) for *Homo sapiens*. In order to balance reliability and coverage, the network was filtered to include only interactions with a median confidence score greater than .4, the resultant network data was exported from Search Tool for the Retrieval of Interacting Genes/Proteins and imported into Cytoscape (version 3.8.0) for visualization. In Cytoscape, each node represents a protein, and the lines between nodes indicate interactions between proteins. During computation, a grid layout algorithm is used to clearly display each node.

2.7. GO and KEGG enrichment analysis

Gene ontology (GO) (<https://geneontology.org>) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (<https://www.kegg.jp>) analyses were performed on the hub-module using the “ClusterProfiler,” “org.Hs.e.g.db,” “enrich plot,” and “ggplot2” package. Pathways with *p* and *q* values <.05 were considered significantly enriched. Display the enrichment results using bar plots. In the GO enrichment analysis, the color of the bars represents the *P*-values, the length of the bars represents the number of enriched genes, and the biological processes are sorted by *P*-values in descending order. The top 6 biological processes are displayed. In the KEGG enrichment analysis, the color of the bars represents the *q*-values, the length of the bars represents the number of enriched genes, and the pathways are sorted by *q*-values in descending order. The top 8 pathways are displayed.

2.8. Correlation analysis

Spearman correlation was performed between the risk score and the proportion of each type of immune cells on the TCGA tumor sample using the “Cor” function. Microsatellite instability (MSI) and tumor mutational burden (TMB) are important biomarkers often used to evaluate the response of certain

types of cancer to immunotherapy. High levels of microsatellite instability (MSI-H) and high levels of TMB generally indicate that tumors are more likely to respond positively to immunotherapy.^[33–35]

In this study, we acquired the corresponding TMB scores using TCGA data and performed a correlation analysis with the risk scores, after which we further divided the sample into high- and low-risk groups based on the median risk scores, and analyzed the differences in TMB between the 2 groups using the “limma” software package. In addition, we obtained the corresponding MSI scores and analyzed the relationship between MSI and high- and low-risk groups in conjunction with the model.

2.9. Chemotherapy resistance analysis

Using the GDSC,^[36] we employed the “pRRophetic” R package to predict the chemotherapy sensitivity of each EC tumor sample. The pRRopheticPredict function was used to predict the IC50 values of all drug (selection = 1, scale = TRUE). A Wilcoxon test was then performed to compare drug sensitivities between the 2 groups (high-risk and low-risk) (pfilter = .05). The gene expression threshold was set to 0.5, once the gene was duplicated, we averaged the gene expressions. In order to reduce the data noise introduced by the batch differences, and to improve the reliability and consistency of the prediction results, we first handle the batch effect. The box plot was used to show the drugs with significant differences in sensitivity between the 2 groups.

2.10. Differential analysis of enrichment pathways in high- and low-risk groups

Gene Set Variation Analysis (GSVA) assessed the enrichment of transcriptomic gene sets using data from the molecular signature database (<https://www.gsea-msigdb.org/gsea/index.jsp>).

This kind of algorithm is capable of estimating gene set enrichment variation across different samples in a nonparametric and unsupervised manner.^[37,38] This study used GSVA algorithm to identify potential biological functional changes in different

groups. The calculations and plotting for GSVA were conducted using the “GSVA,” “limma,” and “ggplot2” packages, with all parameters set to their default values. Differential results ($P < .05$) were displayed using a barplot. The pathways were

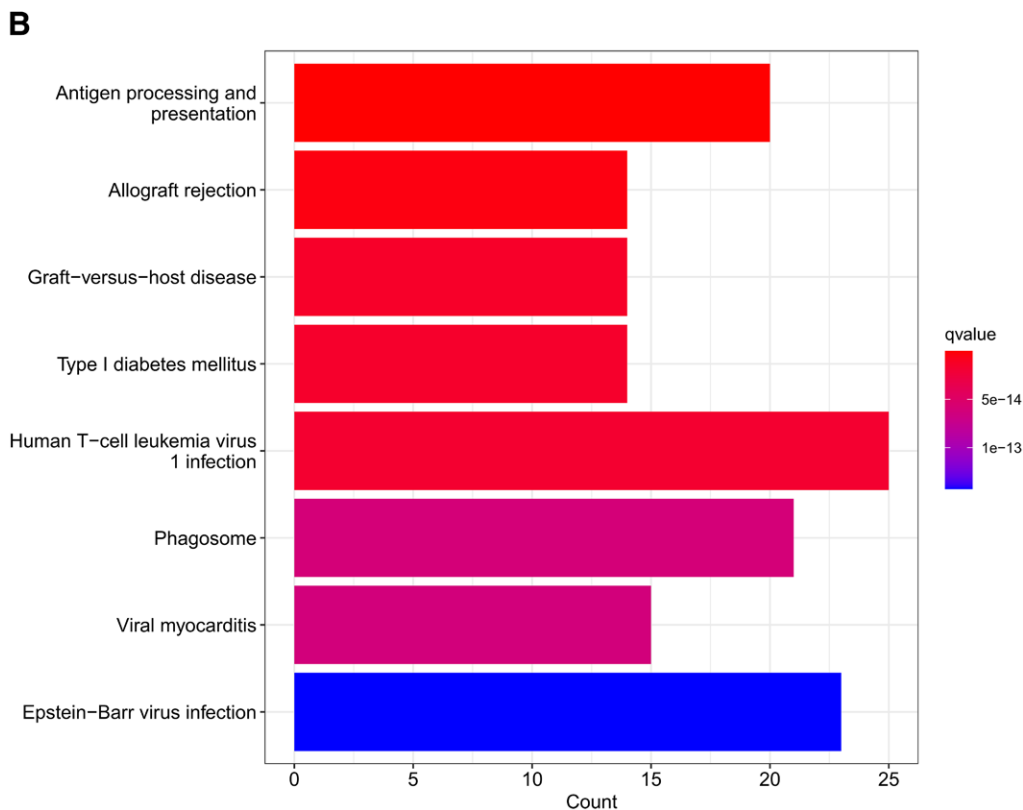
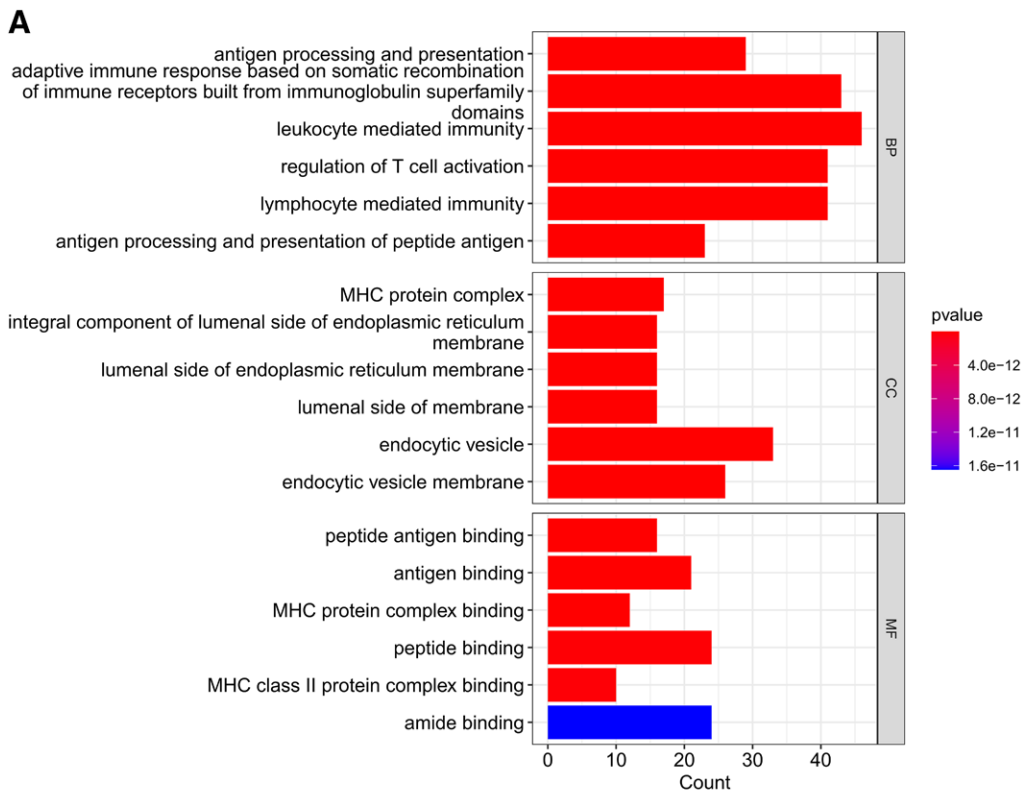


Figure 2. Function enrichment and protein–protein interaction network. (A) GO enrichment analysis of genes in the greenyellow module. (B) KEGG enrichment analysis of genes in the greenyellow module. (C) Protein–protein interaction network of genes in the greenyellow module. GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes.

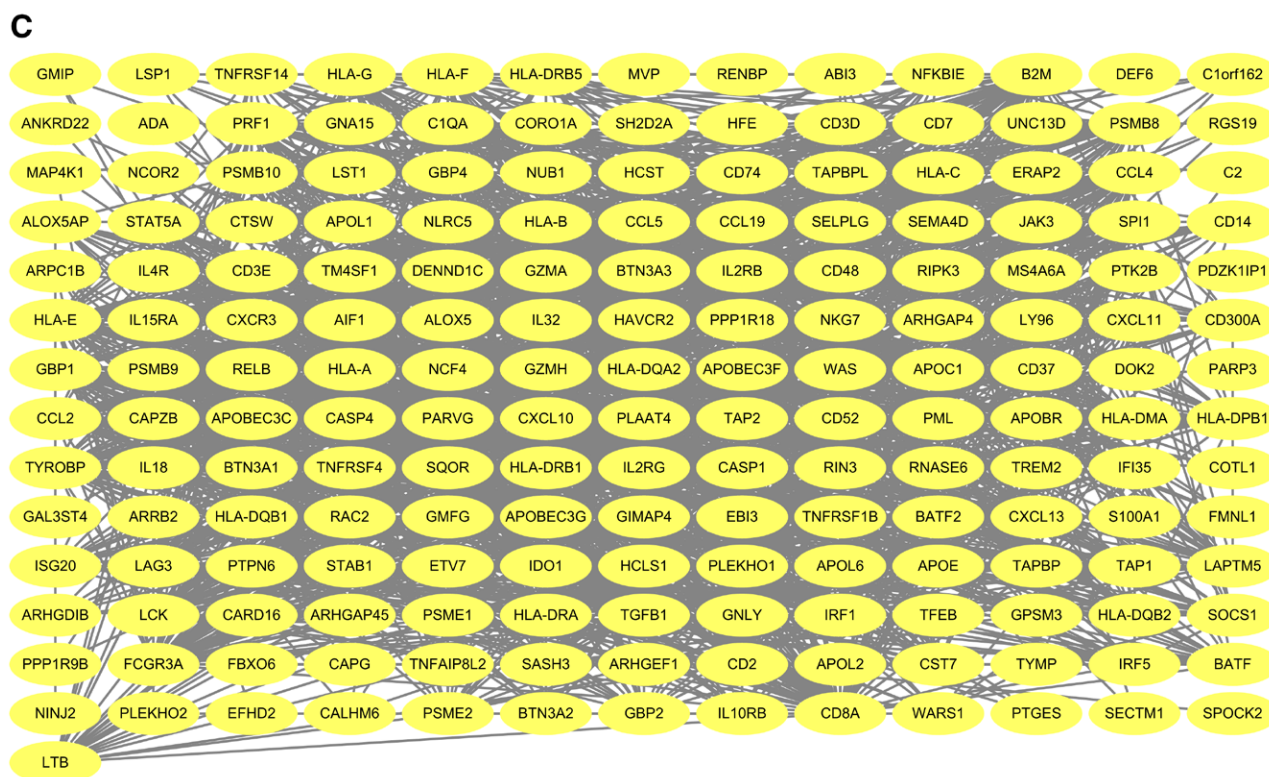


Figure 2. Continued

ranked in descending order according to the absolute value of the t -statistic. An absolute value >1 indicates a significant difference, a positive t -statistic indicates significant enrichment in the high-risk group, and a negative t -statistic indicates significant enrichment in the low-risk group.

2.11. Differential analysis of mutations in high- and low-risk groups

UCEC SNP data from the TCGA database were used to extract gene mutation profiles. Distinct waterfall plots for high- and low-risk groups were created using the “maftools” package, showing the top 20 most frequently mutated genes. Distinct waterfall plots for high- and low-risk groups were created using the “maftools” package, showing the top 20 most frequently mutated genes. In the waterfall plots, different mutation types are annotated with different colors and genes are arranged in descending order of mutation frequency. This method provides a comprehensive overview of the mutation landscape in UCEC stratified by risk level.

2.12. Immunotherapy marker analysis across risk groups

To assess the predictive value of risk scores for immunotherapy efficacy, this study obtained immunophenoscore (IPS) from the TCIA database for UCEC samples, including ips_ctla4_neg_pd1_neg, ips_ctla4_neg_pd1_pos, ips_ctla4_pos_pd1_neg and ips_ctla4_pos_pd1_pos, which represent different combinatorial forms of CTLA-4 and PD-1 inhibitors, respectively, and can respond well to immunotherapy response. Comparisons of 4 types of immune scores were conducted between high- and low-risk groups, using nonparametric tests with a significance threshold set at $P = .05$. The results were visualized using violin plots.

2.13. Immunohistochemistry of tissue microarrays

Tissue microarrays (TMA) were obtained from WEIAOBIO Company (Shanghai, China; TMA No. ZL-UteS961), which included 96 wells representing 48 cases, with each well containing samples from both cancerous and adjusted noncancerous tissues. The TMAs comprised 43 cases of endometrioid adenocarcinomas. Detailed parameters of the TMA and specific information for each point are provided in Supplementary Material S1, Supplemental Digital Content, <http://links.lww.com/MD/O103>. Immunohistochemical (IHC) staining was conducted on the TMAs as follows: after deparaffinization, antigen retrieval, and blocking of endogenous peroxidase activity, the TMAs were incubated with 3 % BSA for 30 minutes at room temperature. Primary antibodies were applied in a wet box at 4 °C overnight, followed by secondary antibody incubation at room temperature for 50 minutes. Detailed antibody parameters and dilution ratios are listed in Supplementary Material S2, Supplemental Digital Content, <http://links.lww.com/MD/O104>.

DAB staining was performed after antibody incubation. The color development was monitored microscopically to achieve a brownish-yellow staining pattern. Once optimal color development was achieved, cell nuclei were counterstained with hematoxylin. The microarray was then dehydrated and sealed to finalize the staining process. The specific IHC procedure and conditions are detailed in Supplementary Material S3, Supplemental Digital Content, <http://links.lww.com/MD/O105> while information on reagents and consumables is provided in Supplementary Material S4, Supplemental Digital Content, <http://links.lww.com/MD/O106>. The methodology and scanning results are described in Supplementary Material S5, Supplemental Digital Content, <http://links.lww.com/MD/O107> and Supplementary Material S6, Supplemental Digital Content, <http://links.lww.com/MD/O108>. The ratio of CD8⁺ positive cells and the basic leucine zipper activating transcription factor-like transcription factor (BATF) H-score were

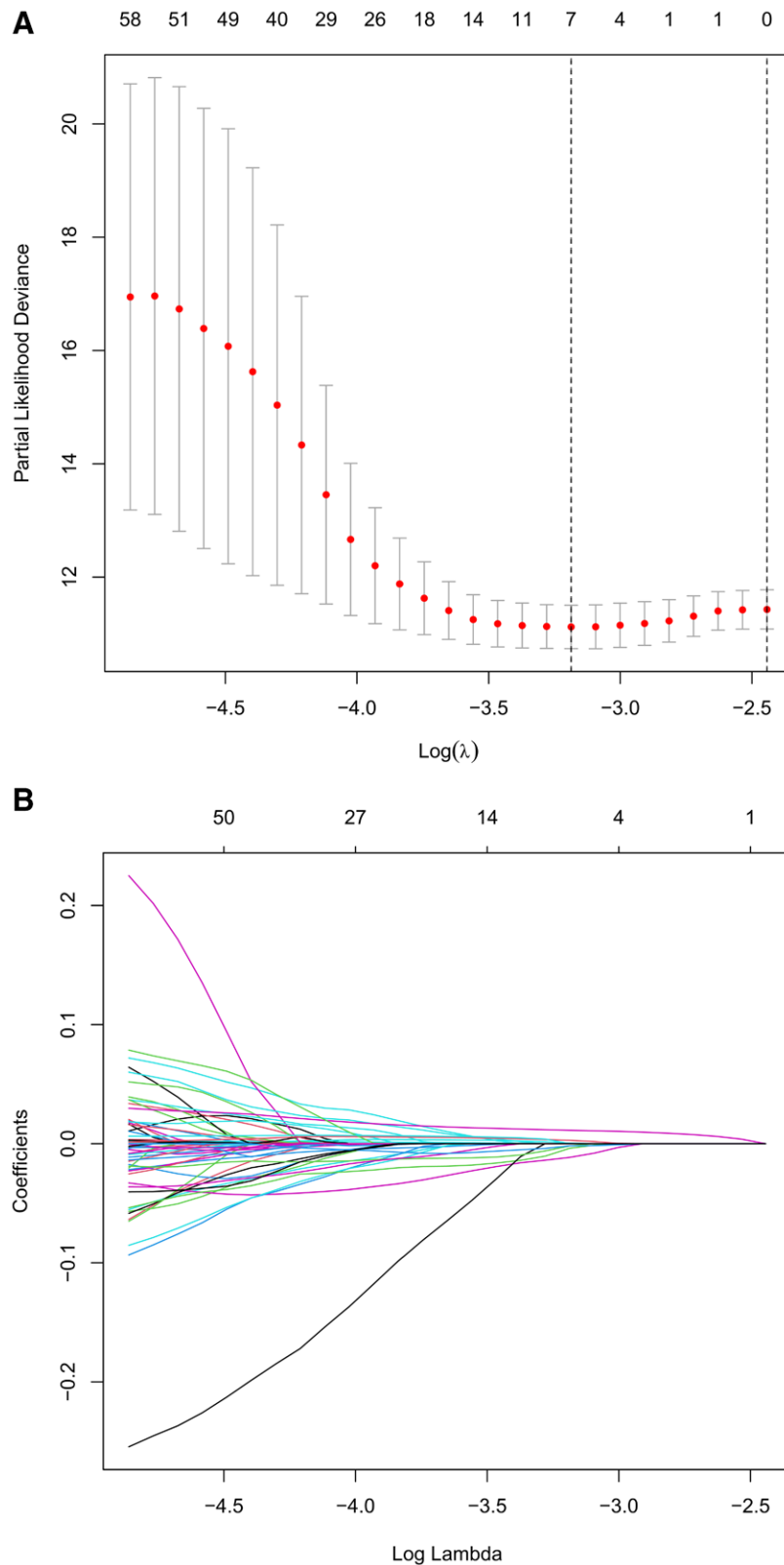


Figure 3. Lasso regression constructs the prognostic model composed of CD8⁺ T cell infiltration-related genes. (A) the best λ value ($\lambda = .02153083$) were obtain from the 10-fold cross-validation method. (B) LASSO coefficient of prognosis-related genes in greenyellow module.

calculated from each well. Endometrioid adenocarcinoma tissues samples were used in this study. Spearman correlation analysis and scatter plotting were performed using OriginLab 2021 software.

2.14. Statistical methods

All statistical calculations were performed using R software (version 4.2.1). Statistical significance was set at $P < .05$.

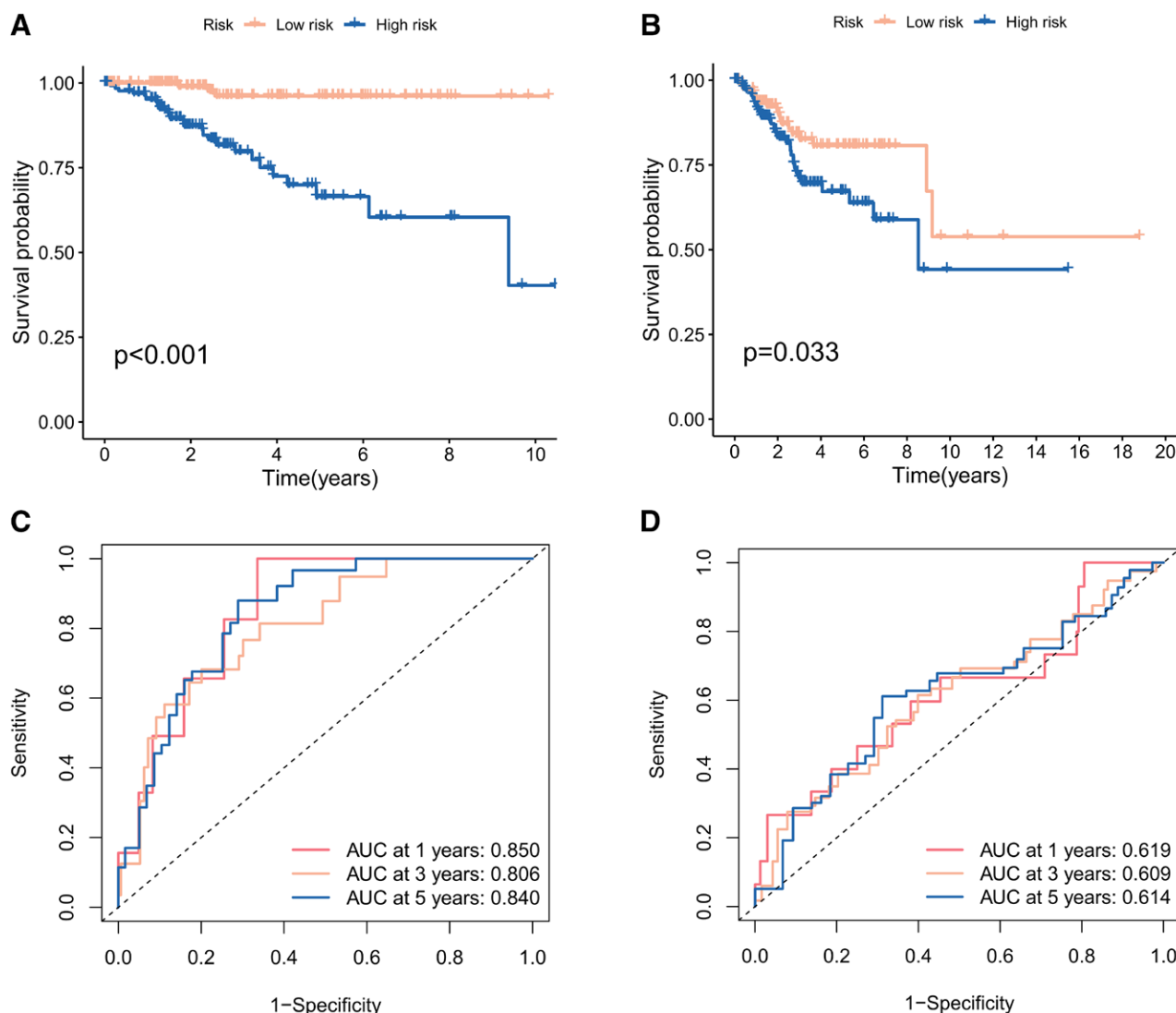


Figure 4. The predictive performance of the model. (A) Kaplan–Meier for TCGA training. (B) Kaplan–Meier for TCGA testing. (C) Survival ROC for training dataset. (D) Survival ROC for internal testing dataset. TCGA = The Cancer Genome Atlas.

3. Results

3.1. Weighted gene co-expression networks analysis

The optimal soft threshold was determined to be 5 (Fig. 1A), and hierarchical clustering revealed 18 gene modules (Fig. 1B). These modules are: black (588 genes), blue (993 genes), brown (846 genes), cyan (125 genes), green (744 genes), green-yellow (225 genes), gray (527 genes), grey60 (71 genes), light cyan (88 genes), magenta (268 genes), midnight blue (93 genes), pink (304 genes), purple (257 genes), red (697 genes), salmon (126 genes), tan (171 genes), turquoise (3114 genes), yellow (763 genes). Among these, the green-yellow module showed the highest correlation with CD8⁺ T cell infiltration ($\text{cor} = .43, P = 3e-27$) (Fig. 1C and D).

3.2. Enrichment analysis of Hub-module

To understand the functional significance of the genes in the green-yellow module, we conducted GO and KEGG enrichment analyses and constructed a protein interaction network. GO enrichment analysis revealed that genes in this module were predominantly involved in leukocyte-mediated immunity, lymphocyte-mediated immunity, regulation of T cell activation, and other related pathways (Fig. 2A). KEGG enrichment analysis indicated that these genes were primarily associated with

antigen processing and presentation and human T-cell leukemia virus infection (Fig. 2B). The protein interaction network analysis is shown in Figure 2C. These pathways are crucial for immune responses and can influence tumorigenesis and antitumor immunity, suggesting that the genes in this module play a role in modulating immune cell activity within the TME.

3.3. Selection of prognostic genes and construction of predictive model

We constructed prognostic models using the training set (Fig. 3A and B). The model formula is:

$$\text{Risk score} = \text{ARPC1B} \times (.005821) + \text{BATF} \times (-.241) + \text{CCL2} \times (.01723724) + \text{COTL1} \times (.010098).$$

This model incorporates 4 genes with significant prognostic associations. In this formula, each gene is assigned a coefficient: a negative coefficient indicates a negative correlation, a positive coefficient indicates a positive correlation, and the absolute value of the coefficient reflects the strength of the correlation. The risk score for each sample is computed as the sum of the products of gene expressions and their respective coefficients. Survival analyses of both the training and test sets demonstrated that the OS of the high-risk group was significantly higher than that of the low-risk group (Fig. 4A and B). Additionally, ROC

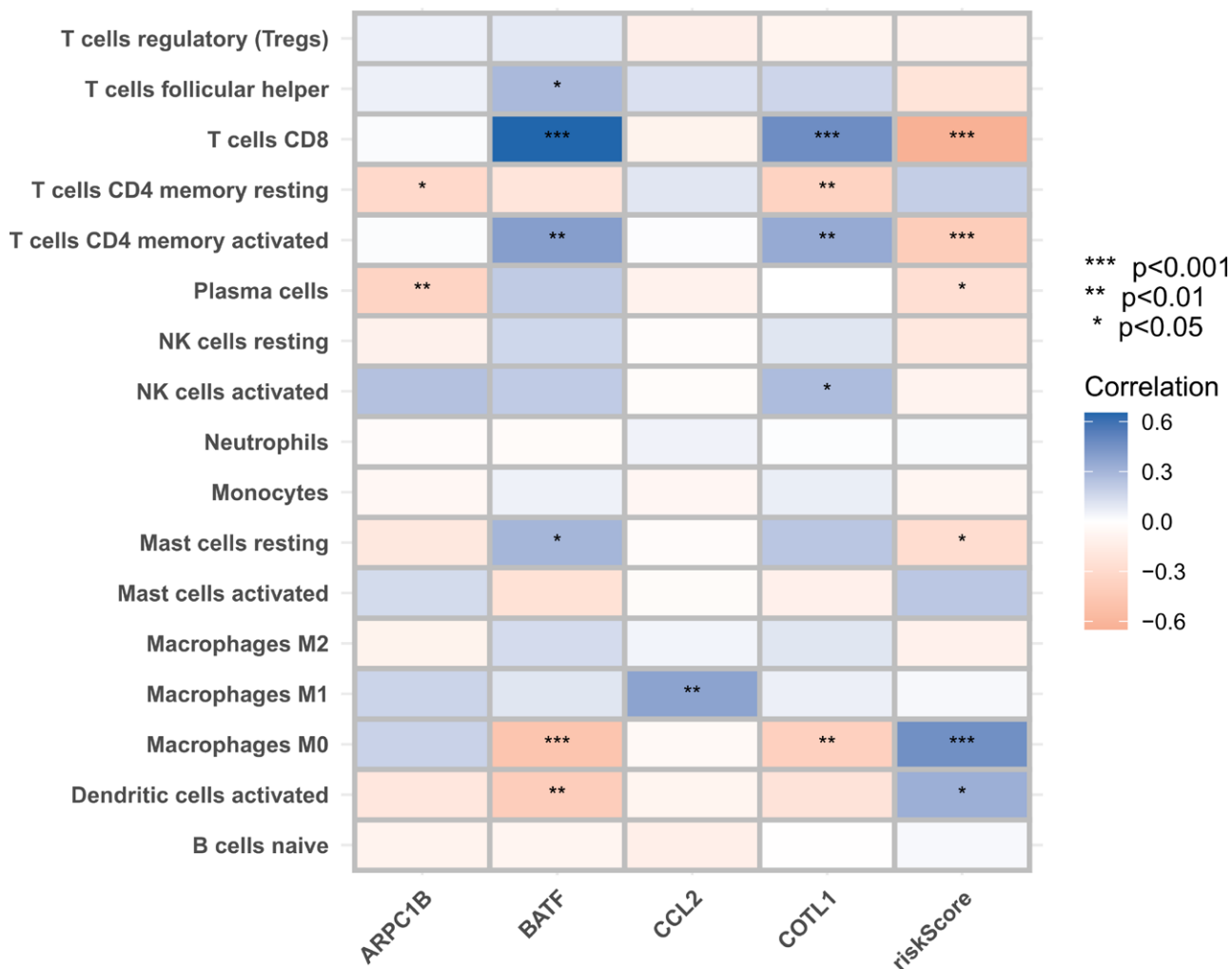


Figure 5. Correlation of risk score with immune infiltration.

curves for the training and test sets showed that the risk scores had strong predictive efficacy (Fig. 4C and D).

3.4. Predictive value of models for immune cell infiltration into EC tissues

The tumor immune microenvironment significantly influences tumor development, survival outcomes, and response to clinical treatments. We conducted a correlation analysis between the risk score and the relative proportions of immune cells in EC tissues. The analysis revealed a significant positive correlation between the risk score and M0 Macrophages ($R = .46, P < .01$) and a significant negative correlation with CD8⁺ T cells ($R = -.65, P < .01$) and activated CD4 + memory T cells ($R = -.42, P < .01$) (Fig. 5). These results suggest that poor prognosis in the high-risk group may be associated with a more suppressive immune environment. This insight is crucial for developing personalized and targeted therapeutic strategies for high-risk patients.

3.5. Relationship between the model and tumorigenic mutation

TMB is a key indicator of sensitivity to immunotherapies. We first assessed the correlation between the risk score and TMB. Our results indicated that TMB was lower in the high-risk group compared to the low-risk group ($P = .0064$)

(Fig. 6A) and that risk scores had a negative correlation with TMB ($R = -.12, P = .0061$) (Fig. 6B). These results suggest that high-risk patients may have a reduced response to immunotherapy. Further analysis of gene mutations in both risk groups revealed distinct differences in mutation frequencies. Specifically, TP53 mutation frequency was significantly higher in the high-risk group, whereas PTEN and ARIDA1A mutation frequencies were higher in the low-risk group (Fig. 6D). These findings offer valuable insights into potential immunotherapy resistance mechanisms in high-risk groups and highlight specific gene mutations that may contribute to this resistance.

3.6. Relationship between the model and MSI

MSI is another predictor of immunotherapy sensitivity. We analyzed MSI differences between high- and low-risk groups using TCGA tumor samples. The analysis showed no significant difference in MSI between the 2 groups (Fig. 6C).

3.7. Relationship between the model and the immunotherapy efficacy

To compare immunotherapy effectiveness between high- and low-risk groups, we used the TCIA database to calculate and compare various immune phenomenon scores. The results indicated that tumors in the high-risk group may exhibit poorer

immunoreactivity and a reduced potential response to immunotherapy ($P < .01$; Fig. 7).

3.8. Relationship between the model and chemotherapy resistance

We predicted chemotherapy sensitivity for each tumor sample based on drug sensitivity data from the GDSC database. The

results showed that the risk score was significantly correlated with sensitivity to bosutinib, bortezomib, BIRB.0796, and other drugs (Fig. 8).

3.9. GSVA analysis of models

To investigate the molecular mechanisms underlying the model's effect on EC, we performed GSVA to analyze differentially

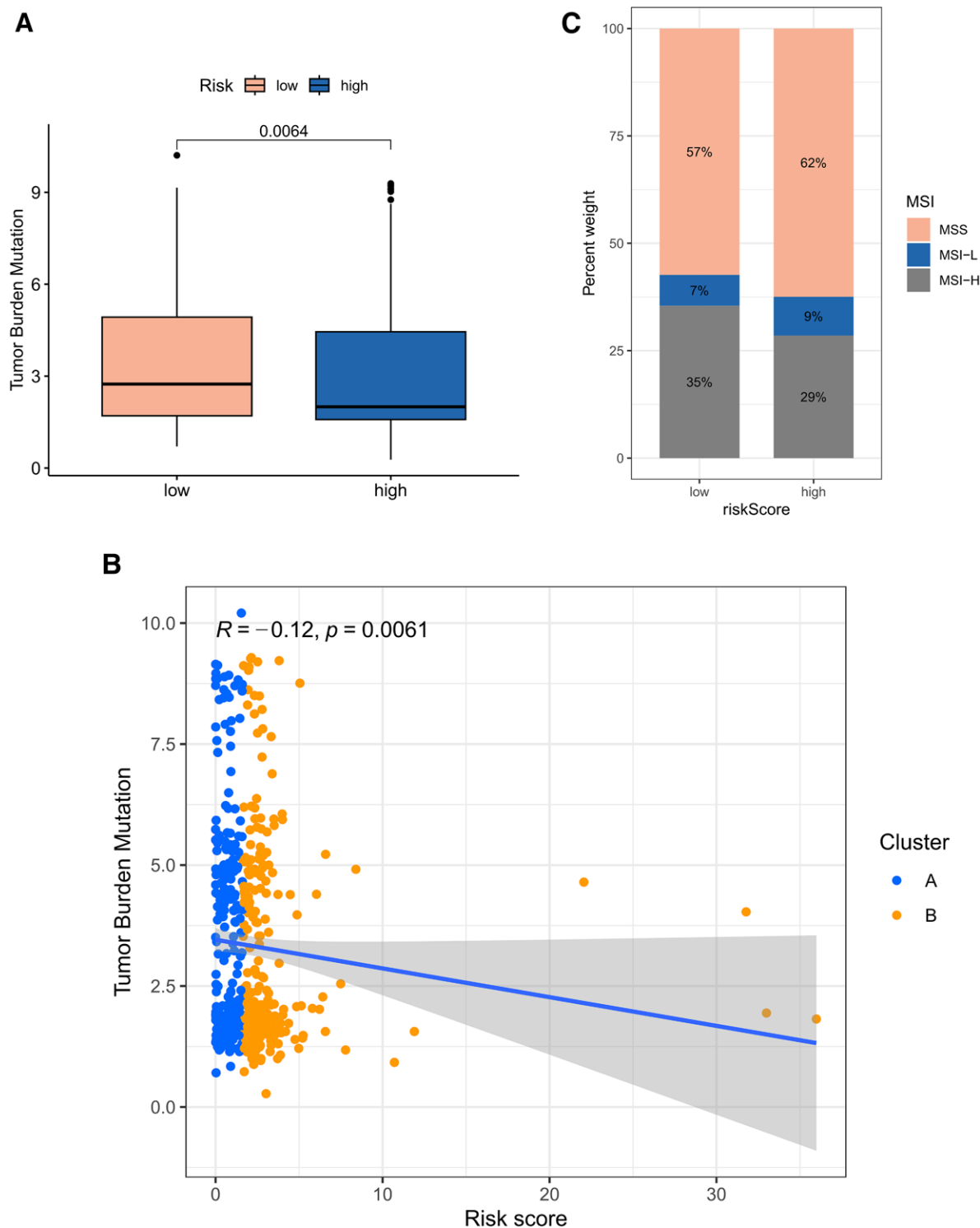


Figure 6. Mutational status of genes and microsatellite status in high- and low-risk groups. (A) Differences in TMB between high and low risk groups. (B) Correlation of risk scores with TMB. (C) Differences in microsatellite status between high and low risk groups. (D) The mutation profiles of high- and low-risk groups. TMB = tumor mutational burden.

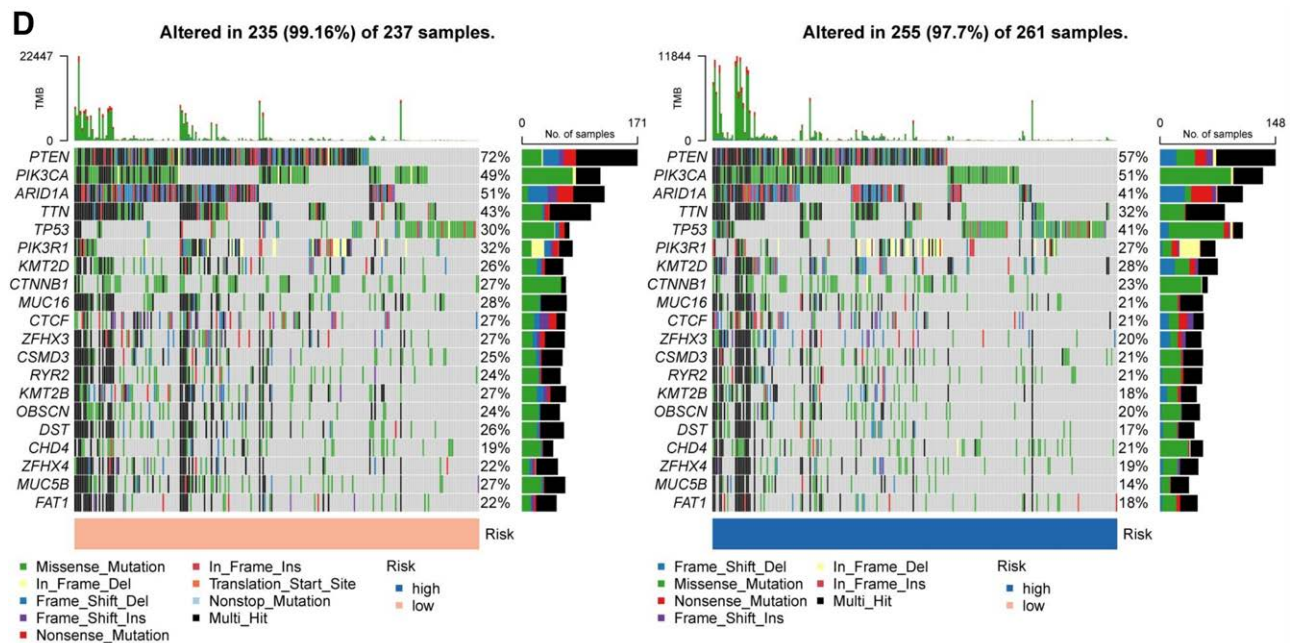


Figure 6. Continued

enriched pathways between high- and low-risk groups. Key pathways identified include TGF_BETA_SIGNALING, UV_RESPONSE_DN, HEDGEHOG_SIGNALING, ALLOGRAFT_REJECTION (Fig. 9).

3.10. Immunohistochemistry of tissue microarrays

Given that BATF had the highest absolute coefficient in the model, we performed IHC for CD8 and BATF in EC TMA to examine the relationship between BATF expression level and CD8⁺ T cell infiltration. Figure 10A. We found a significant positive correlation between BATF H-score and the CD8⁺ cell ratio ($R = .13348$, $P = .00136$ Fig. 10B). This indicates that BATF expression in EC tissues correlates with CD8⁺ T cell infiltration and may play a significant role in antitumor immunity. Higher BATF expression is associated with increased CD8⁺ T cell infiltration and enhanced antitumor immunity activity.

4. Discussion

Treating recurrent and metastatic EC presents significant challenges, and current methods for early identification of high-risk cases are limited. To address these gaps, several prognostic models incorporating genes related to various biological processes have been developed.^[10–12] The TIME is crucial for tumor progression, with CD8⁺ T cells playing a key role in tumor immunity.

Previous research has developed prognostic models for other cancers, such as lung squamous cell carcinoma, using genes related to CD8⁺ T-cell infiltration.^[39] Our study builds on this approach by developing a prognostic model for EC based on genes associated with CD8 + T-cell infiltration. Our model differentiates itself by employing WGCNA to identify relevant genes, enhancing precision compared to direct analyses and minimizing the inclusion of irrelevant genes. Shi et al^[40] also used machine learning methods to identify a few key genes and constructed a model. Our study shares some methodological similarities with theirs, but what sets us apart is our emphasis on using WGCNA for initial screening. Our model calculates risk scores for EC patients, classifying them into high- and low-risk groups. This model proves valuable for early identification of high-risk patients and can guide

clinicians in tailoring treatment strategies. Validation using the TCGA-UCEC test set demonstrated robust predictive efficacy for EC prognosis.

The genes incorporated in our model associated with CD8⁺ T-cell infiltration, suggesting that risk scores may reflect immune infiltration patterns in EC. Our analysis revealed that higher risk scores correlated with increased infiltration of Macrophages M0 cells and decreased infiltration of CD8⁺ T cells and activated CD4⁺ memory T cells. This pattern indicates that high-risk scores are linked to a suppressive immune microenvironment. Enrichment analysis further identified key pathways, including TGF- β signaling, UV response, Hedgehog signaling, and allograft rejection, with TGF- β and Hedgehog signaling being notably associated with antitumor immunity.

Hedgehog-GLI signaling, active during embryonic development and in adult stem cells,^[41] is implicated in tumor proliferation, invasion, and metastasis. Activated Hedgehog-GLI signaling contributes to a suppressive tumor microenvironment by inhibiting CD8⁺ T-cell recruitment and enhancing the expression of immunosuppressive cytokines, such as CCL21, CCL9, and CXCL69, along with PD-L1.^[42–46] Similarly, TGF- β signaling impacts various cellular processes and is known to suppress CD8⁺ T-cell activity and induce immune tolerance, further contributing to a suppressive tumor microenvironment.^[47–53]

Our TMA analysis of EC demonstrated a significant positive correlation between BATF expression, having the highest coefficient value in our model, and CD8⁺ T cell infiltration. And BATF has a negative coefficient value in the model. This suggests that a high-risk score correlates with a suppressive immune microenvironment. Notably, patients with high-risk scores exhibited lower responses to immunotherapy, as confirmed by TMB and TCIA analyses. This underscores the potential of our model in guiding clinicians to consider alternative or combination therapies to overcome immune suppression and enhance patient outcomes.

Interestingly, our study found no association between risk scores and MSI-H status. Although MSI-H is a known predictor of immunotherapy efficacy, it is not always a reliable marker for predicting responses.^[54–59] Therefore, our risk model may provide valuable insights for clinical decision-making in immunotherapy for EC.

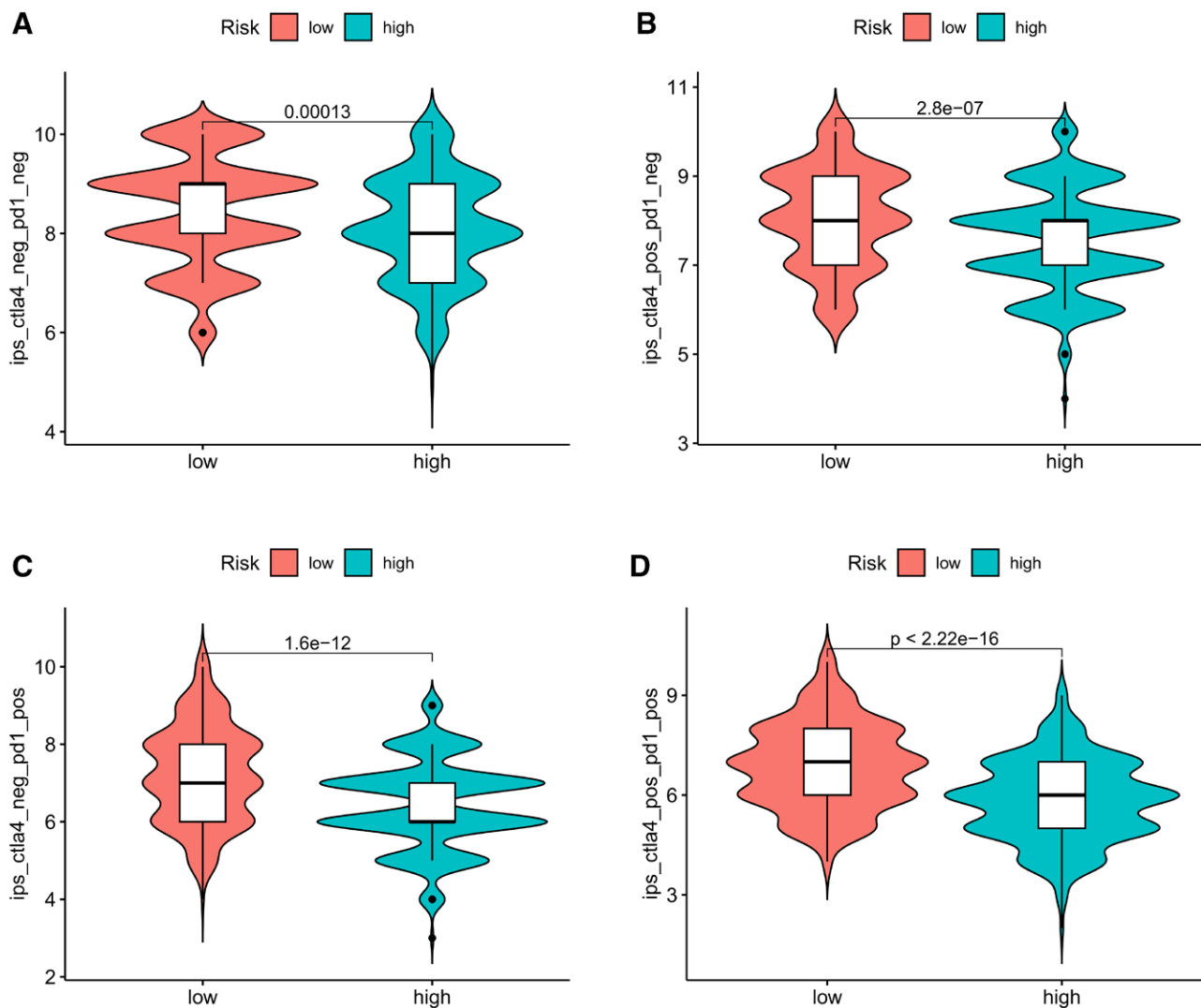


Figure 7. Differences in immune checkpoint expression in patients in high- and low-risk groups (expressed in terms of immunization phenomenon scores). (A) CTLA4 (-), PD1 (-), (B) CTLA4 (+), PD1 (-). (C) CTLA4 (-), PD1 (+), (D) CTLA4 (+), PD1 (+).

Our prognostic model includes 4 genes: BATF, ARPC1B, CCL2, and COTL1. Among these, BATF exhibited the highest coefficient. BATF, encoded by the BATF gene, is crucial for CD8⁺ T cell activation, influencing their development, differentiation, and function.^[60–62] Previous research has demonstrated that BATF knockdown in CD8⁺ T cells leads to defects in proliferation, altered mRNA expression of other transcription factors, and abnormal cytokine production.^[62] Our study found that BATF expression in EC tissues was significantly positively correlated with CD8⁺ T-cell infiltration, as observed through TMA IHC. This suggests BATF plays a pivotal role in antitumor immunity in EC. In our model, the negative coefficient for BATF implies that lower BATF expression is associated with higher risk scores. This finding, combined with TMA and immunotherapy data, suggests that low BATF expression may contribute to an inhibitory immune microenvironment, resulting in poor prognosis and resistance to immunotherapy. Further research is needed to explore this mechanism.

The other genes in our model also play significant roles in antitumor immunity across various cancers. Actin-related protein 2/3 complex 1 is a key component of the actin-related protein 2/3 (Arp2/3) complex, crucial for forming the actin cytoskeleton,^[63,64] which is essential for

immune cell functions, including those of CD8⁺ T cells.^[65,66] ARPC1B mediates CD8⁺ T cell proliferation, migration, and cytotoxicity.^[67] Our study found that ARPC1B deletion leads to reduced CD8⁺ T cells and decreased cytotoxicity,^[68] contributing to immune-related diseases.^[68–73] Additionally, ARPC1B expression correlates with metastasis and prognosis in several cancers, such as prostate cancer, glioblastoma, ovarian cancer, and oral squamous cell carcinoma.^[74–77] However, its role in EC remains unexplored.

Coactosin-like 1 (COTL1), also known as coactosin-like protein, is an actin-binding protein that regulates the actin cytoskeleton by binding to α - and β -actin and F-actin.^[78] COTL1 acts as a T-cell activator by competing with cofilin for binding to F-actin, thus promoting lamellipodial protrusion and attenuating cofilin-mediated F-actin depolymerization.^[79] Although COTL1 has been identified as a cancer-associated antigen,^[80,81] its role in EC and antitumor immunity remains to be elucidated. The C-C motif chemokine ligand 2 (CCL2) is a key mediator of immune cell recruitment during microbial infections and tissue damage. It is often overexpressed in cancer cells and the tumor microenvironment, where high CCL2 levels are associated with aggressive malignancies, increased metastasis, and poorer outcomes across various cancers.^[82] CCL2 is a potent

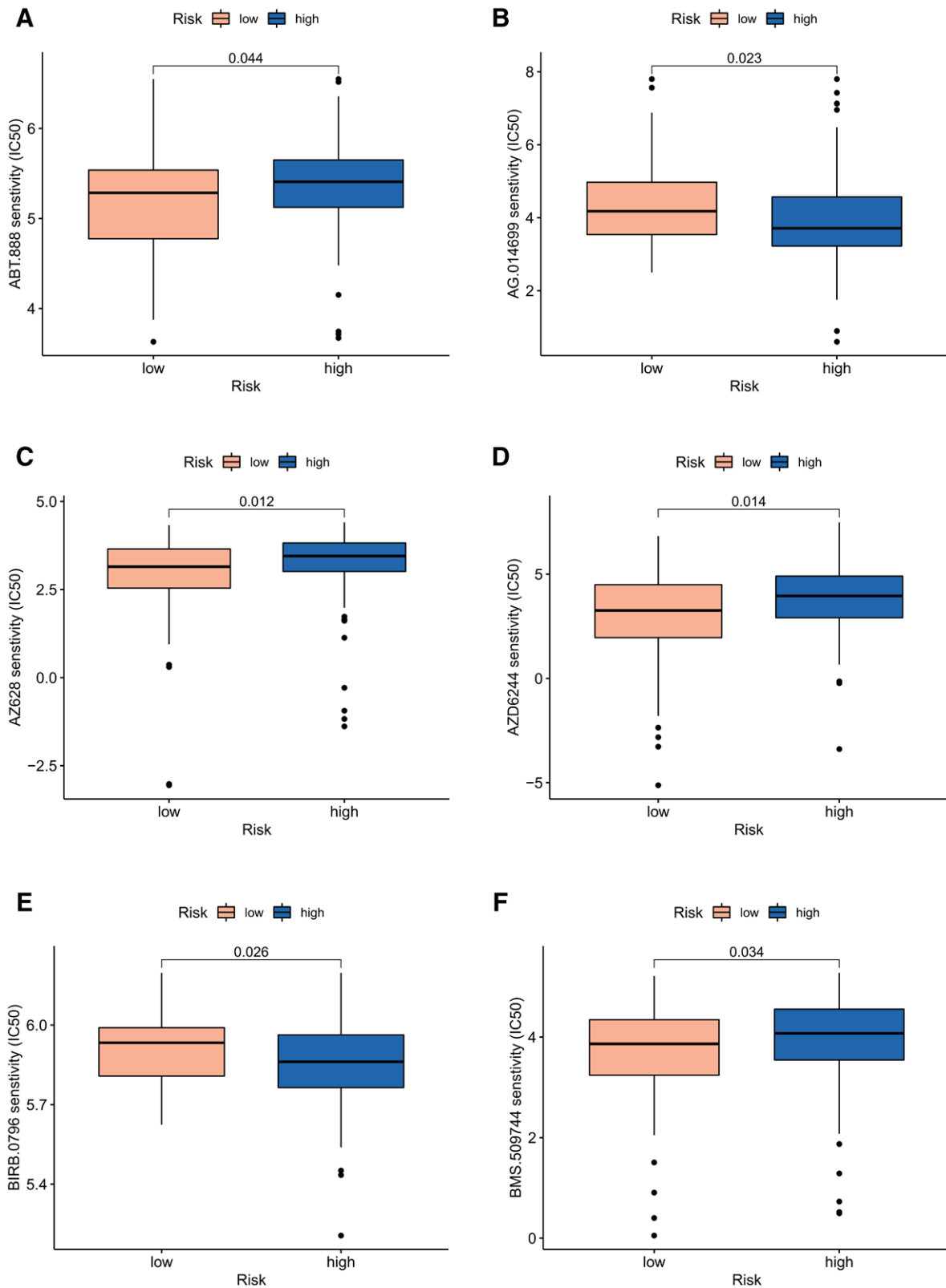


Figure 8. Differences in chemotherapy sensitivity between high- and low-risk groups, based on data from the GDSC database. (A) ABT.888, (B) AG.014699, (C) AZ628, (D) AZD6244, (E) BIRB.0796, (F) BMS.509744, (G) BMS.536924, (H) Bortezomib, (I) Bosutinib, (J) Camptothecin.

chemoattractant for macrophages and plays a role in maintaining a suppressive immune microenvironment.^[83] Our study is the first to suggest potential roles for ARPC1B, CCL2, and COTL1 in EC progression and antitumor immunity, warranting further investigation.

The innovation of this study lies in its development of a prognostic model for EC based on CD8⁺ T cell infiltration using co-expression module screening and machine learning. The model effectively identifies high-risk EC patients and reflects the activity of the immune microenvironment and

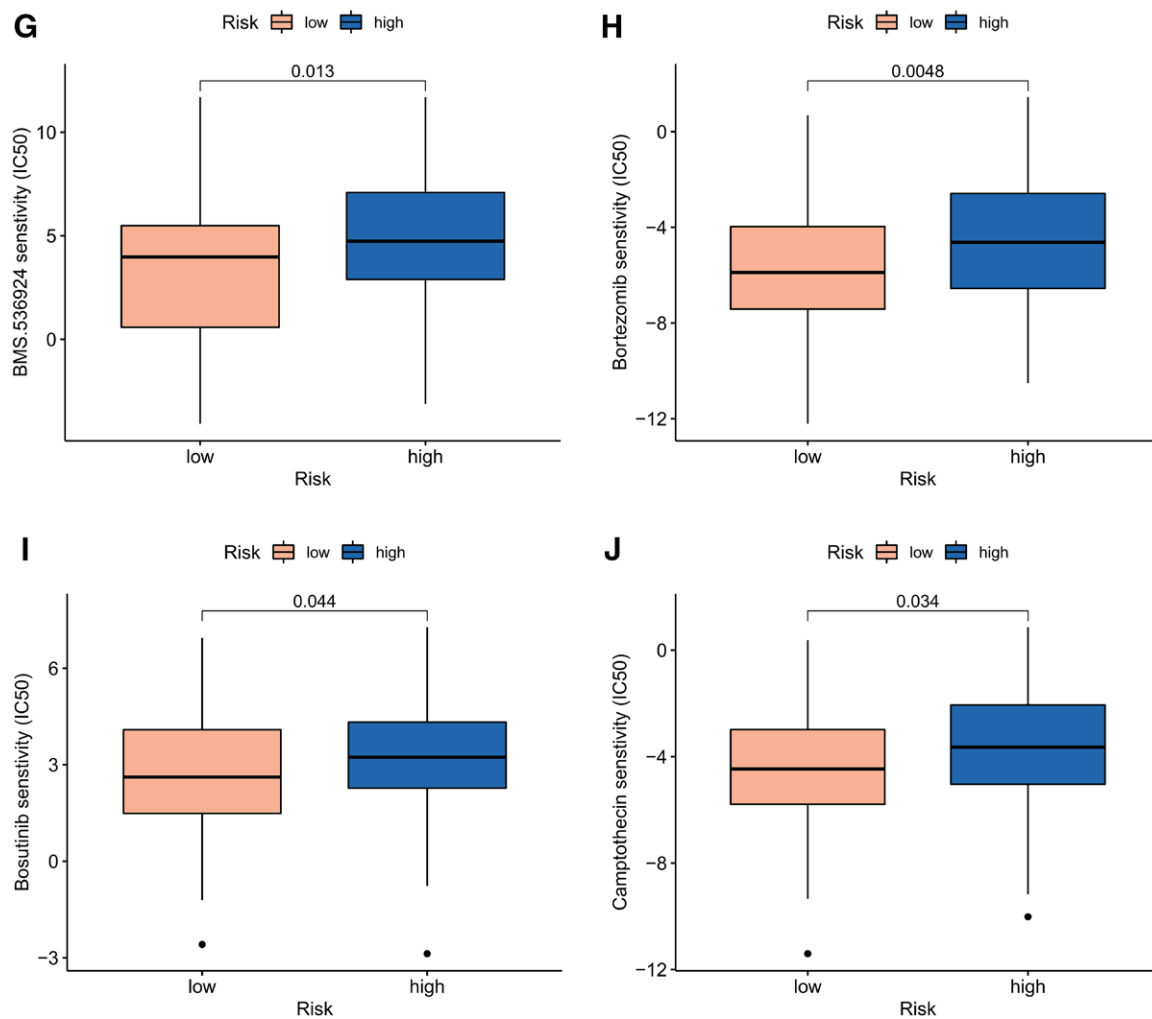


Figure 8. Continued

potential response to immunotherapy. It offers clinicians a tool for early identification of high-risk patients and prediction of immunotherapy resistance, supporting more personalized treatment strategies with potential clinical applications. Additionally, our study is the first to explore the relationship between BATF expression and CD8⁺ T cell infiltration using TMA, which maintains sample processing consistency and provides more accurate results compared to dispersed tissue sections or cell lines. This approach highlights BATF's significant role in EC antitumor immunity and its impact on prognosis and immunotherapy efficacy. Future studies should delve deeper into BATF's role in EC and consider it as a potential target to overcome immunotherapy resistance. Our model also identifies ARPC1B, CCL2, and COTL1 as potentially important in EC progression and antitumor immunity, necessitating further research.

However, this study has some limitations. First, data were primarily obtained from public databases, introducing potential selection and information biases despite a large sample size. Second, while the model performed well in training and test sets, its clinical accuracy requires validation through prospective studies with larger samples. Third, the multidimensional nature of the study increased data processing complexity and the risk of errors, which must be addressed in larger clinical studies. Finally, although we identified potential roles for BATF, ARPC1B, CCL2, and COTL1 in EC, their mechanisms of action need further exploration.

5. Conclusion

We employed WCGNA and machine learning techniques to develop a prognostic model for EC based on CD8⁺ T cell infiltration. This model is valuable for early identification of EC patients at high risk of recurrence and metastasis and provides predictive insights into the efficacy of immunotherapy. By aiding clinicians in formulating more targeted and individualized treatment strategies, this model has the potential to improve patient outcomes. Additionally, our findings suggest that the genes BATF, ARPC1B, CCL2, and COTL1 may play significant roles in EC progression and immune infiltration. Specifically, BATF emerges as a promising target for future EC therapies.

Acknowledgments

We would like to thank Editage (www.editage.cn) for English language editing.

Author contributions

Conceptualization: Chao Chen.

Investigation: Lipeng Pei.

Project administration: Wei Ren.

Supervision: Wei Ren.

Writing – original draft: Chao Chen.

Writing – review & editing: Jingli Sun.

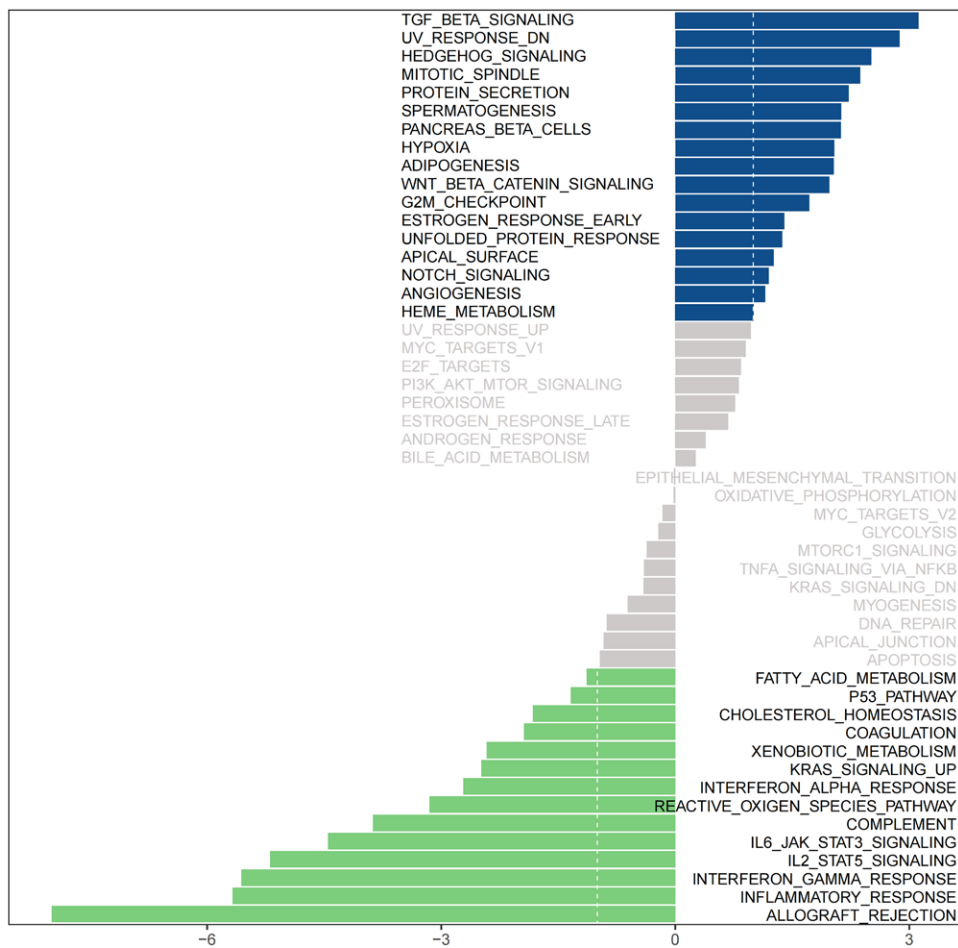


Figure 9. GSEA enrichment analysis of differential genes in high- and low-risk groups. GSEA = Gene Set Variation Analysis.

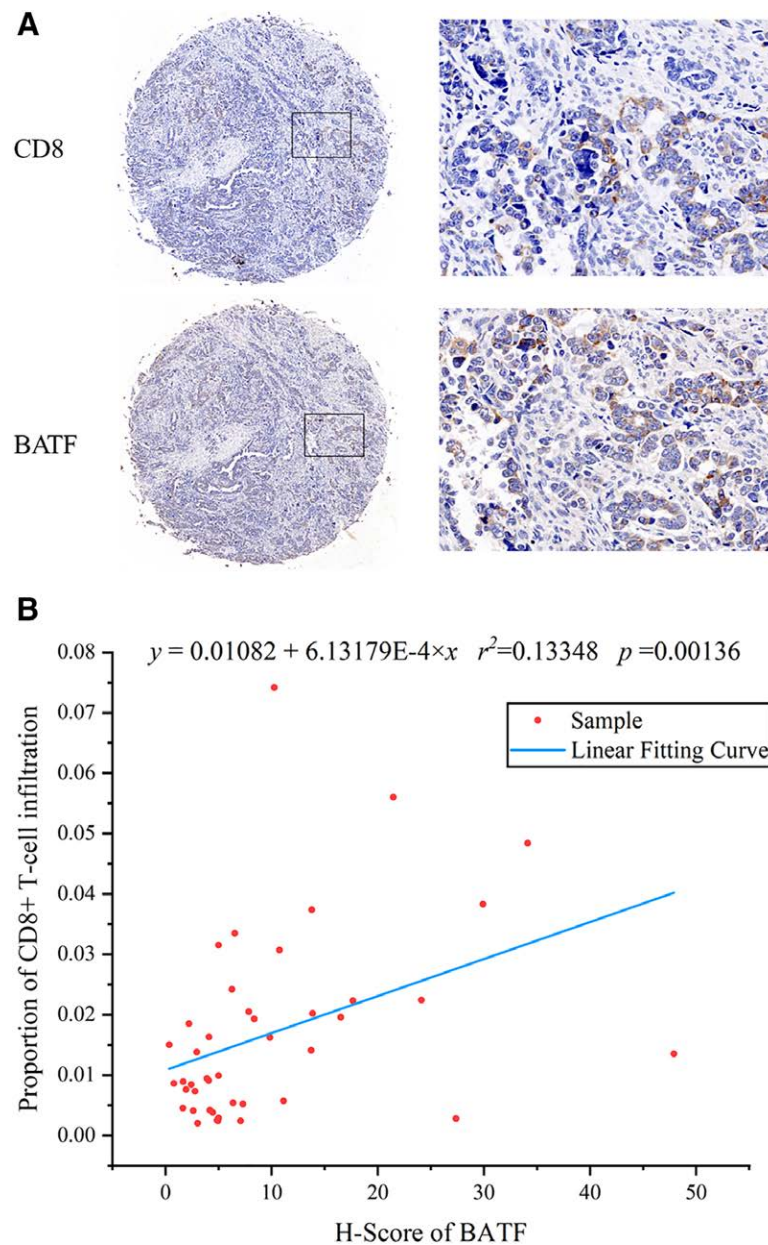


Figure 10. Expression level of BATF and proportion of CD8⁺ T cell infiltration in TMA of EC. (A) Representative graphs of CD8 and BATF staining in TMA (brownish-yellow is positive). (B) Scatterplot of correlation between H-score of BATF and proportion of CD8⁺ T cell infiltration in TMA. BATF = activating transcription factor-like transcription factor, EC = endometrial cancer, TMA = tissue microarrays.

References

- [1] Ventriglia J, Paciolla I, Pisano C, et al. Immunotherapy in ovarian, endometrial and cervical cancer: State of the art and future perspectives. *Cancer Treat Rev*. 2017;59:109–16.
- [2] Islami F, Ward EM, Sung H, et al. annual report to the nation on the status of cancer, Part 1: national cancer statistics. *J Natl Cancer Inst*. 2021;113:1648–69.
- [3] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin*. 2013;63:11–30.
- [4] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70:7–30.
- [5] Lortet-Tieulent J, Ferlay J, Bray F, Jemal A. International patterns and trends in endometrial cancer incidence, 1978–2013. *J Natl Cancer Inst*. 2018;110:354–61.
- [6] Pfeiffer RM, Webb-Vargas Y, Wheeler W, Gail MH. Proportion of U.S. trends in breast cancer incidence attributable to long-term changes in risk factor distributions. *Cancer Epidemiol Biomarkers Prev*. 2018;27:1214–22.
- [7] Clarke MA, Devesa SS, Harvey SV, Wentzensen N. Hysterectomy-corrected uterine corpus cancer incidence trends and differences in relative survival reveal racial disparities and rising rates of nonendometrioid cancers. *J Clin Oncol*. 2019;37:1895–908.
- [8] Jemal A, Ward EM, Johnson CJ, et al. Annual Report to the Nation on the Status of Cancer, 1975–2014, featuring survival. *J Natl Cancer Inst*. 2017;109:djx030.
- [9] Lu KH, Broaddus RR. Endometrial cancer. *N Engl J Med*. 2020;383:2053–64.
- [10] Wang X, Dai C, Ye M, Wang J, Lin W, Li R. Prognostic value of an autophagy-related long-noncoding-RNA signature for endometrial cancer. *Aging (Albany NY)*. 2021;13:5104–19.
- [11] Jiang Y, Chen J, Ling J, et al. Construction of a Glycolysis-related long noncoding RNA signature for predicting survival in endometrial cancer. *J Cancer*. 2021;12:1431–44.
- [12] Liu J, Li S, Feng G, et al. Nine glycolysis-related gene signature predicting the survival of patients with endometrial adenocarcinoma. *Cancer Cell Int*. 2020;20:183.

- [13] Grizzi F, Basso G, Borroni EM, et al. Evolving notions on immune response in colorectal cancer and their implications for biomarker development. *Inflamm Res*. 2018;67:375–89.
- [14] Sanmamed MF, Chen L. A paradigm shift in cancer immunotherapy: from enhancement to normalization. *Cell*. 2018;175:313–26.
- [15] Forde PM, Chaft JE, Pardoll DM. Neoadjuvant PD-1 Blockade in Resectable Lung Cancer. *N Engl J Med*. 2018;379:e14.
- [16] Yang L, Lin PC. Mechanisms that drive inflammatory tumor micro-environment, tumor heterogeneity, and metastatic progression. *Semin Cancer Biol*. 2017;47:185–95.
- [17] Elinav E, Nowarski R, Thaïss CA, Hu B, Jin C, Flavell RA. Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms. *Nat Rev Cancer*. 2013;13:759–71.
- [18] Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer*. 2016;16:275–87.
- [19] Nagarsheth N, Wicha MS, Zou W. Chemokines in the cancer micro-environment and their relevance in cancer immunotherapy. *Nat Rev Immunol*. 2017;17:559–72.
- [20] Wu T, Dai Y. Tumor microenvironment and therapeutic response. *Cancer Lett*. 2017;387:61–8.
- [21] Meads MB, Gatenby RA, Dalton WS. Environment-mediated drug resistance: a major contributor to minimal residual disease. *Nat Rev Cancer*. 2009;9:665–74.
- [22] Degos C, Heinemann M, Barrou J, et al. Endometrial tumor micro-environment alters human NK cell recruitment, and resident NK cell phenotype and function. *Front Immunol*. 2019;10:877.
- [23] Jhunjhunwala S, Hammer C, Delamarre L. Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. *Nat Rev Cancer*. 2021;21:298–312.
- [24] Kondratiev S, Sabo E, Yakirevich E, Lavie O, Resnick MB. Intratumoral CD8+ T lymphocytes as a prognostic factor of survival in endometrial carcinoma. *Clin Cancer Res*. 2004;10:4450–6.
- [25] Workel H, Komdeur F, Wouters M, et al. CD103 defines intraepithelial CD8+ PD1+ tumour-infiltrating lymphocytes of prognostic significance in endometrial adenocarcinoma. *Eur J Cancer*. 2016;60:1–11.
- [26] Jong R, Leffers N, Boezen H, et al. Presence of tumor-infiltrating lymphocytes is an independent prognostic factor in type I and II endometrial cancer. *Gynecol Oncol*. 2019;114:105–10.
- [27] Smith BJ, Silva-Costa LC, Martins-de-Souza D. Human disease biomarker panels through systems biology. *Biophys Rev*. 2021;13:1179–90.
- [28] Mitra S, Das S, Chakrabarti J. Systems biology of cancer biomarker detection. *Cancer Biomark*. 2013;13:201–13.
- [29] Liao Y, Wang Y, Cheng Y, Huang C, Fan X. Weighted gene co-expression network analysis of features that control cancer stem cells reveals prognostic biomarkers in lung adenocarcinoma. *Front Genet*. 2020;11:311.
- [30] Zhong J, Shi S, Peng W, et al. Weighted Gene Co-Expression Network Analysis (WGCNA) reveals the functions of syndecan-1 to regulate immune infiltration by influenced T cells in glioma. *Front Genet*. 2022;13:792443.
- [31] Momeni K, Ghorbian S, Ahmadpour E, Sharifi R. Unraveling the complexity: understanding the deconvolutions of RNA-seq data. *Translat Med Commun*. 2023;8:0–0.
- [32] Chen B, Khodadoust MS, Liu CL, et al. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. 2018;0:243–59.
- [33] Palmeri M, Mehnert JM, Silk AW, et al. Real-world application of tumor mutational burden-high (TMB-high) and microsatellite instability (MSI) confirms their utility as immunotherapy biomarkers. *ESMO Open*. 2022;7:100336–100336.
- [34] Zou D, Xu T. Construction and validation of a colon cancer prognostic model based on tumor mutation burden-related genes. *Sci Rep*. 2024;14:0–0.
- [35] Bergom HE, Sena LA, Day A, et al. Divergent immune microenvironments in two tumor nodules from a patient with mismatch repair-deficient prostate cancer. *npj Genomic Med*. 2024;9:0–0.
- [36] Yang W, Lightfoot H, Bignell GR, et al. Genomics of Drug Sensitivity in Cancer (GDSC): A resource for biomarker discovery in cancer cells. *Eur J Cancer*. 2016;69:S82–S82.
- [37] Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinf*. 2013;14:0–0.
- [38] Zhao K, Rhee SY. Interpreting omics data with pathway enrichment analysis. *Trends Genet*. 2023;39:308–19.
- [39] Tang M, Li Y, Luo X, et al. Identification of Biomarkers Related to CD8(+) T Cell Infiltration with Gene co-expression network in lung squamous cell carcinoma. *Front Cell Dev Biol*. 2021;9:606106.
- [40] Shi W, Chen Z, Liu H, et al. COL11A1 as a novel biomarker for breast cancer with machine learning and immunohistochemistry validation. *Front Immunol*. 2022;13:937125.
- [41] Ingham PW, McMahon AP. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev*. 2001;15:3059–87.
- [42] Hanna A, Shevde LA. Hedgehog signaling: modulation of cancer properties and tumor microenvironment. *Mol Cancer*. 2016;15:24.
- [43] Grund-Gröschke S, Stockmaier G, Aberger F. Hedgehog/GLI signaling in tumor immunity - new therapeutic opportunities and clinical implications. *Cell Commun Signal*. 2019;17:172.
- [44] Giammona A, Crivaro E, Stecca B. emerging roles of hedgehog signaling in cancer immunity. *Int J Mol Sci*. 2023;24:1321.
- [45] Kogerman P, Grimm T, Kogerman L, et al. Mammalian suppressor-of-fused modulates nuclear-cytoplasmic shuttling of Gli-1. *Nat Cell Biol*. 1999;1:312–9.
- [46] Onishi H, Fujimura A, Oyama Y, et al. Hedgehog signaling regulates PDL-1 expression in cancer cells to induce anti-tumor activity by activated lymphocytes. *Cell Immunol*. 2016;310:199–204.
- [47] Bai X, Yi M, Jiao Y, Chu Q, Wu K. Blocking TGF- β signaling to enhance the efficacy of immune checkpoint inhibitor. *Oncotargets Ther*. 2019;12:9527–38.
- [48] Gupta A, Budhu S, Merghoub T. One checkpoint may hide another: inhibiting the TGF β signaling pathway enhances immune checkpoint blockade. *Hepatobiliary Surg Nutr*. 2019;8:289–94.
- [49] Haque S, Morris JC. Transforming growth factor- β : A therapeutic target for cancer. *Hum Vaccin Immunother*. 2017;13:1741–50.
- [50] Hendriks ME, Bolarinwa OA, Nelissen HE, et al. Costs of cardiovascular disease prevention care and scenarios for cost saving: a micro-costing study from rural Nigeria. *J Hypertens*. 2015;33:376–684.
- [51] Battle E, Massagué J. Transforming Growth Factor- β Signaling in Immunity and Cancer. *Immunity*. 2019;50:924–40.
- [52] Yu Y, Feng XH. TGF- β signaling in cell fate control and cancer. *Curr Opin Cell Biol*. 2019;61:56–63.
- [53] Larson C, Oronsky B, Carter CA, et al. TGF-beta: a master immune regulator. *Expert Opin Ther Targets*. 2020;24:427–38.
- [54] Crosbie EJ, Kitson SJ, McAlpine JN, Mukhopadhyay A, Powell ME, Singh N. Endometrial cancer. *Lancet*. 2022;399:1412–28.
- [55] Brooks RA, Fleming GF, Lastra RR, et al. Current recommendations and recent progress in endometrial cancer. *CA Cancer J Clin*. 2019;69:258–79.
- [56] Eskander RN, Sill MW, Beffa L, et al. Pembrolizumab plus chemotherapy in advanced endometrial cancer. *N Engl J Med*. 2023;388:2159–70.
- [57] Harris E. Pembrolizumab immunotherapy improved endometrial cancer survival. *JAMA*. 2023;329:1341.
- [58] Mirza MR, Chase DM, Slomovitz BM, et al. Dostarlimab for primary advanced or recurrent endometrial cancer. *N Engl J Med*. 2023;388:2145–58.
- [59] Makker V, Colombo N, Casado Herráez A, et al. Lenvatinib plus pembrolizumab for advanced endometrial cancer. *N Engl J Med*. 2022;386:437–48.
- [60] Betz BC, Jordan-Williams KL, Wang C, et al. Batf coordinates multiple aspects of B and T cell function required for normal antibody responses. *J Exp Med*. 2010;207:933–42.
- [61] Ise W, Kohyama M, Schraml BU, et al. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nat Immunol*. 2011;12:536–43.
- [62] Kurachi M, Barnitz RA, Yosef N, et al. The transcription factor BATF operates as an essential differentiation checkpoint in early effector CD8+ T cells. *Nat Immunol*. 2014;15:373–83.
- [63] Campellone KG, Welch MD. A nucleator arms race: cellular control of actin assembly. *Nat Rev Mol Cell Biol*. 2010;11:237–51.
- [64] Goley ED, Rammohan A, Znameroski EA, Firat-Karalar EN, Sept D, Welch MD. An actin-filament-binding interface on the Arp2/3 complex is critical for nucleation and branch stability. *Proc Natl Acad Sci U S A*. 2010;107:8159–64.
- [65] Pollard TD, Cooper JA. Actin, a central player in cell shape and movement. *Science*. 2009;326:1208–12.
- [66] Randzavola LO, Strege K, Juzans M, et al. Loss of ARPC1B impairs cytotoxic T lymphocyte maintenance and cytolytic activity. *J Clin Invest*. 2019;129:5600–14.
- [67] Abella JV, Galloni C, Pernier J, et al. Isoform diversity in the Arp2/3 complex determines actin filament dynamics. *Nat Cell Biol*. 2016;18:76–86.
- [68] Kahr WH, Pluthero FG, Elkadri A, et al. Loss of the Arp2/3 complex component ARPC1B causes platelet abnormalities and predisposes to inflammatory disease. *Nat Commun*. 2017;8:14816.

- [69] Kuijpers TW, Tool ATJ, van der Bijl I, et al. Combined immunodeficiency with severe inflammation and allergy caused by ARPC1B deficiency. *J Allergy Clin Immunol.* 2017;140:273–7.e10.
- [70] Volpi S, Cicalese MP, Tuijnburg P, et al. A combined immunodeficiency with severe infections, inflammation, and allergy caused by ARPC1B deficiency. *J Allergy Clin Immunol.* 2019;143:2296–9.
- [71] Somech R, Lev A, Lee YN, et al. Disruption of thrombocyte and T lymphocyte development by a mutation in ARPC1B. *J Immunol.* 2017;199:4036–45.
- [72] Brigida I, Zoccolillo M, Cicalese MP, et al. T-cell defects in patients with ARPC1B germline mutations account for combined immunodeficiency. *Blood.* 2018;132:2362–74.
- [73] Papadatou I, Marinakis N, Botsa E, et al. Case report: a novel synonymous ARPC1B gene mutation causes a syndrome of combined immunodeficiency, asthma, and allergy with significant intrafamilial clinical heterogeneity. *Front Immunol.* 2021;12:634313.
- [74] Huang J, Zhou H, Tan C, Mo S, Liu T, Kuang Y. The overexpression of actin related protein 2/3 complex subunit 1B (ARPC1B) promotes the ovarian cancer progression via activation of the Wnt/ β -catenin signaling pathway. *Front Immunol.* 2023;14:1182677.
- [75] Auzair LB, Vincent-Chong VK, Ghani WM, et al. Caveolin 1 (Cav-1) and actin-related protein 2/3 complex, subunit 1B (ARPC1B) expressions as prognostic indicators for oral squamous cell carcinoma (OSCC). *Eur Arch Otorhinolaryngol.* 2016;273:1885–93.
- [76] Gamallat Y, Zaaluk H, Kish EK, et al. ARPC1B is associated with lethal prostate cancer and its inhibition decreases cell invasion and migration in vitro. *Int J Mol Sci.* 2022;23:1476.
- [77] Liu T, Zhu C, Chen X, et al. Dual role of ARPC1B in regulating the network between tumor-associated macrophages and tumor cells in glioblastoma. *Oncoimmunology.* 2022;11:2031499.
- [78] Provost P, Doucet J, Stock A, Gerisch G, Samuelsson B, Rådmark O. Coactosin-like protein, a human F-actin-binding protein: critical role of lysine-75. *Biochem J.* 2001;359(Pt 2):255–63.
- [79] Kim J, Shapiro MJ, Bamidele AO, et al. Coactosin-like 1 antagonizes cofilin to promote lamellipodial protrusion at the immune synapse. *PLoS One.* 2014;9:e85090.
- [80] Nakatsura T, Senju S, Ito M, Nishimura Y, Itoh K. Cellular and humoral immune responses to a human pancreatic cancer antigen, coactosin-like protein, originally defined by the SEREX method. *Eur J Immunol.* 2002;32:826–36.
- [81] Guo S, Yang P, Jiang X, et al. Genetic and epigenetic silencing of miRcoRNA-506-3p enhances COTL1 oncogene expression to foster non-small lung cancer progression. *Oncotarget.* 2017;8:644–57.
- [82] O'Connor T, Heikenwalder M. CCL2 in the tumor microenvironment. *Adv Exp Med Biol.* 2021;1302:1–14.
- [83] Jin J, Lin J, Xu A, et al. CCL2: an important mediator between tumor cells and host cells in tumor microenvironment. *Front Oncol.* 2021;11:722916.