Exploration of a novel prognostic risk signatures and immune checkpoint molecules in endometrial carcinoma microenvironment

Jinhui Liu
The first affiliated hospital of nanjing medical university
Sipei Nie
The first affiliated hospital of nanjing medical universty
Zhipeng Wu
The affiliated sir run run hospital of nanjing medical university
Siyue Li
The first affiliated hospital of nanjing medical university
Yi Jiang
The first affiliated hospital of nanjing medical university
Yicong Wan
The first affiliated hospital of nanjing medical university
Huangyang Meng
The first affiliated hospital of nanjing medical university
Shulin Zhou
The first affiliated hospital of nanjing medical university
Wenjun Cheng (wenjunchengdoc@163.com)
The First Affiliated Hospital of Nanjing Medical University

Research

Keywords: Endometrial carcinoma (EC), prognostic model, Tumor microenvironment, the Cancer Genome Atlas (TCGA)

Posted Date: April 1st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-18621/v1

Version of Record: A version of this preprint was published at Genomics on September 1st, 2020. See the published version at https://doi.org/10.1016/j.ygeno.2020.05.022.

Abstract

Background: Endometrial carcinoma (EC) is one of the common female malignancies. The effectiveness of immunotherapy in treating EC remains to be testified.

Methods: In our present study, we explored the underlying mechanism through tumor microenvironment (TME) and immune-related genes in EC. Based on The Cancer Genome Atlas (TCGA) database to screen immune-related differentially expressed genes (DEGs). Then, survival analysis, functional annotations and interpreted interaction network of these genes were further investigated for the underlying regulation mechanisms. Moreover, we used a conjoint Cox regression analysis to built the immune-related prognostic risk signatures and identify the candidate prognostic biomarkers. Based on the risk model, we further explored the predictive value of risk signatures by exploring the tumor-infiltrating immune cells, prognostic immune checkpoints modulators, and checkpoint inhibitors in TME.

Results: As a result, we found 799 up-regulated and 139 down-regulated immune-related and differentially expressed genes of EC. Based on the 225 prognostic DEGs, we built two risk models based on key prognostic immune-related genes, which identified respectively for overall survival and disease-free survival. Gene set enrichment analysis revealed immune-related pathways were mostly enriched in gene sets of the low-risk group, and the nomograms were constructed based on the risk models. By analyzing the immune status in low- and high- group of EC patients , we found higher immune cell infiltration and activation in low risk tumor tissues, which might cause lower tumor purity and contribute to better prognosis of EC. By analyzing the immune checkpoint molecules, we found both the expression and immunophenoscore of immune checkpoints PD-1, CTLA4, PD-L1 and PD-L2 increased significantly in low-risk group. In addition, the low-risk group were found display higher tumor mutational burden.

Conclusions: Our study suggested that high-immune TME might prevent the progression of EC, which indicated the possibility of immunotherapy to fight against EC.

Background

Endometrial carcinoma (EC) is the fourth common malignancy in the women worldwide [1], with an ever-rising incidence over past years[2]. Early diagnosis is critical for EC treatment, as evidenced by the five-year survival rate lower than 20% in advanced-stage patients but higher than 90% in early-stage patients[3]. To develop an early-diagnosis strategy, sensitive biomarkers are urgently needed.

Increasing evidence suggests that tumorigenesis is often triggered in a tumor microenvironment (TME), which is composed of extracellular matrix, blood or lymphatic vessels, fibroblasts, immune cells and inflammatory cells [4]. Recent years have seen tumor cases successfully treated by immune checkpoint therapy using cytotoxic T lymphocyte antigen-4 (CTLA-4) and programmed death-1 (PD-1), and by chimeric antigen receptor (CAR) T cells. It has been proven that managing TME might yield favorable outcomes in cancer immunotherapy [5]. TME has also been reported to play a significant role in EC

progression. Subhransu S Sahoo et al. reviewed the role of different stromal cells in different stages of EC progression, and suggested that inhibiting TME related signals might suppress metastatic EC progression[6]. Felix AS et al. demonstrated that the overexpression of VEGF was involved in tumor growth and metastasis of EC [7]. Pasanen A et al. found that PD-L1 was differentially expressed in EC tissue of different histological types, clinical stages and molecular subtypes [8]. However, the effect of immunotherapy and TME on EC remains unclear. In recent years, new genome-sequencing technologies and genomic databases allowed the massive discovery of tumor biomarkers. Increasing studies have attempted to evaluate the prognostic value of infiltrating immune and stromal cells in malignancies (hepatocellular carcinoma, renal cell carcinoma and osteosarcoma)[9–11]. In the present study (Fig. 1), we defined the immune-related genes with potentially prognostic value in EC based on TCGA database and the relationship between immune microenvironment and EC, which might provide instructions for developing new immunotherapy for EC patients.

Materials And Methods

2.1 | Collection of gene expression profiles

The gene expression profiles of 587 EC samples (552 EC samples and 35normal samples) were acquired from the TCGA database (https://tcga-data.nci.nih.gov/tcga/). Meanwhile, the corresponding clinical information of EC patients (Table 1 and Table 2), including patient age, grade, stage, status and histological type of tumor, were also downloaded from the TCGA database. Meanwhile, the immune- and stromal- scores in tumor samples were calculated by using ESTIMATE algorithm [12].

Table 1
Clinical information of included EC patients based on OS in TCGA

Characteristics	Sample(N = 544)	Percentage (%)
Age		
=60	208	38.24%
60	336	61.76%
Stage		
Stage I & Stage II	391	71.88%
Stage III & Stage IV	153	28.12%
Histological_type		
Endometrioid	408	75%
Mixed & Serous	136	25%
Grade		
G1 & G2	221	40.63%
G3 & G4	323	59.37%
Fustat		
alive	453	83.27%
dead	91	16.73%

Table 2
Clinical information of included EC patients based on DFS in TCGA

Characteristics	Sample (N = 493)	Percentage (%)
Age		
=60	201	40.77%
60	292	59.23%
Stage		
Stage I & Stage II	362	73.43%
Stage III & Stage IV	131	26.57%
Histological_type		
Endometrioid	378	76.67%
Mixed & Serous	115	23.33%
Grade		
G1 & G2	206	41.78%
G3 & G4	287	58.22%
Fustat		
Disease Free	377	76.47%
Recurred	116	23.53%

2.2 | Identification Of Immune-related Degs

According to the scores in ESTIMATE algorithm, EC samples were respectively divided into high- and low-score group to select the up- and down- regulated DEGs (|Log Fold change (FC) > 0.5|, p-value < 0.05 and FDR < 0.05). We used the “pheatmap” package in R software to generate Heatmaps of DEGs. Meanwhile, venn diagrams were plotted to screen out the overlapped DEGs in stromal-score group and immune-score group.

2.3| Identification Of Prognostic Degs

In order to identify DEGs with prognostic value in both immune-score group and stromal-score group, the survival analysis was performed into the overlapped DEGs using the “survival” package of R software.

The p-value < 0.05 was considered as significantly.

2.4 | Functional Annotation Of Prognostic Degs

To get insight of the bio-functions of the prognostic DEGs, functional annotations were performed by using "clusterProfiler" package in R. Gene Ontology (GO) term enrichment analysis, and were classified in three categories: biological processes (BP), molecular functions (MF), and cellular components (CC). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were also performed.

2.5| Construction of the protein-protein interaction (PPI) network

To investigate interactions between the coding products of these prognostic DEGs, PPI network was built by Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://www.string-db.org/), and the MCODE plug-in of Cytoscape software [13].

2.6| Construction Of Immune-related Prognostic Risk Model

To improve the reliability of the risk signature, the entire EC patients were randomly divided into train set and test set at 1:1 ratio, the train set was used for constructing the prognostic risk signature, and the test set and entire set were used for validating the predicting value. To explore candidate prognostic biomarkers of EC, a joint cox regression analysis were performed. Firstly, the univariate Cox proportional hazard regression analysis was used to identify a candidate prognostic gene in the train set when p-value was < 0.05. Then, the least absolute shrinkage and selection operator (LASSO) Cox regression analysis was performed by using “glmnet” package in R software version 3.4.3 [14]. Finally,multivariate Cox regression analysis was employed to construct the prognosis signature for predicting the prognosis in EC patients. The risk score was calculated as follows: exp gene 1 * β gene 1 + exp gene 2 * β gene 2 + exp gene 3 * β gene 3 + ... exp gene n * β gene n (the exp gene represents the gene expression level, the β gene represent the regression coefficient calculated by multivariate Cox proportional hazard regression). In addition, patients were divided into low- and high-risk groups according to the median risk score. In addition, these key prognostic genes were used to construct a risk liner model. The Kaplan-Meier curves and the ROC curves according to the risk model were plotted in train、test and entire set. In addition, a conjoint Cox regression was conducted to define the independent risk variables. Kaplan–Meier survival analysis in the clinicopathological risk stratification was executed in low- and high-risk groups of entire set.

2.7 | Building And Confirmation Of The Nomograms

We utilized the “rms” R package to construct the nomogram and present a calibrate curve which incorporated the prognosis signature and clinical factors. Then, we visualized the performance of the nomogram by presenting the predicted and actual prognosis of the nomogram in the calibration curve [15]. The 45° line represents the best prediction. The accuracy of the nomogram was examined using the consistency between the actual and the predicted outcomes.

2.8 | Gene set enrichment analysis (GSEA) analysis of key prognostic immune-related genes

GSEA (http://software.broadinstitute.org/gsea/index.jsp) was implemented to determine biological processes enriched in the gene rank. Entire EC samples in TCGA set were divided into high-risk and low-risk groups based on the prognostic models. By comparing the biological processes enriched in the two groups, we defined the underlying biological function of the key genes. The collection of annotated gene sets in c2.cp.kegg.v6.0.symbols.gmt in Molecular Signatures Database (MSigDB, http://software.broadinstitute.org/gsea/msigdb/index.jsp) was chosen as the reference gene set in GSEA software. The Nom. $p < 0.05$ were chosen as the cutoff criterion [16].

2.9 | Estimation Of Immune Cell Infiltration Fractions

CIBERSORT is a method to explore the cell components of complex tissues based on the gene expression profiles, which is highly consistent with ground truth estimations in types of cancer [17]. Normalized gene expression data were used to deduce the relative proportions of 22 types of infiltrating immune cells (encompass B cells, T cells, natural killer cells, dendritic cells, eosinophils, macrophages and neutrophils, amongst others.) via CIBERSORT algorithm, in combination with the LM22 signature matrix. Heatmap of immune cell proportions were plotted to illustrate the differential distributions of immune cells in EC cancer tissues and the matching para-cancer tissues. The Wilcoxon rank-sum test was performed to precisely assess the differential infiltrating density not only in EC cancer tissues and the matching para-cancer tissues, but also in high- and low- risk group.

2.10 | Analysis Of Immunophenoscore (ips)

Immunomodulators, effector cells, immunosuppressive cells and MHC molecules were considered as the four main components that determine the immunogenicity of tumor, which could be evaluated as IPS. IPS is calculated using a scale with a range of 0–10 based on representative cell type gene expression z-scores, where higher scores are associated with increased immunogenicity. The IPS of EC patients was acquired from The Cancer Immunome Atlas (TCIA) (https://tcia.at/home). The previous study have used IPS (including IPS-PD1/PD-L1/PD-L2, IPS-CTLA4, and IPSPD1/PD-L1/PD-L2 + CTLA4 scores) to estimate the response for immune checkpoint inhibitor for CESC patients [18].

2.10 | Evaluation of immune status between high-risk and low-risk groups stratified by prognostic model.

In order to investigate the association between immune system and the prognostic signature, the immune status of the high- and low-risk groups were further analyzed. Firstly, the immune activity in high- and low-risk groups were quantified by single-sample gene-set enrichment analysis (ssGSEA) by exploring 29 immune-related gene signatures [19]. The corresponding tumor purities was also be evaluated by using ESTIMATE algorithm [20], and the expression of HLA-genes between high- and low-risk group was further analyzed.

2.11 | Mutation Analysis

The mutation data of EC patients were obtained from the TCGA database that mentioned above. The mutation data of somatic variants were extract from Mutation Annotation Format (MAF), and analyzed by using MAF tools [21]. The tumor mutation burden (TMB) score for each EC patient were calculated by the formula as follows: TMB= (total mutation/total covered bases) × 10^6 [22].

Results

3.1| Immune status is associated with prognosis of EC

As the results uncovered by the ESTIMATE algorithm, the stromal scores were distributed from − 2224.62 to 860.43. The immune scores were ranged from 1359.51 to 3614.677. The ESTIMATE scores (composed of immune scores and stromal scores) were ranged from − 3166.98- 3990.15. By analyzing the association between these three scores and the clinical features of 587 samples, we found patients consistently show higher scores in lower grades (Fig. 2A), and other clinical factors did not show consistent statistical significance. We then analyzed the survival rate in high- and low- score groups, and found that patients with higher immune scores were significantly positively associated with overall survival (OS) (p = 0.009), but the stromal scores (p = 0.136). However, the ESTIMATE scores were still statistically significant (p = 0.045) (Fig. 2B). These results indicating that stromal and immune scores of EC were positive factors in patient prognosis.

3.2 | Acquisition Of Immune-relate Degs Of Ec

In our present study, immune-related genes were screened out by comparing different expression level in high- and low- score groups (Fig. 3A-B). There were 1093 upregulated genes and 764 downregulated genes in the high immune-score group, and 1478 upregulated genes and 243 downregulated genes in high stromal-score group. Finally, we selected out the overlapped genes in high-stromal/immune scores

groups. As the Venn diagram showed, 799 upregulated genes and 139 downregulated genes were identified as the immune-related DEGs of EC (Fig. 3C).

3.3 | Functional Annotation Of Immune-related Degs

Furthermore, the GO terms and KEGG pathway analysis of the 938 DEGs were performed. As the results of GO terms analysis showed, the top five enriched terms in BP group were “GO:0042110-T cell activation, GO:0007159-leukocyte cell-cell adhesion, GO:0051249-regulation of lymphocyte activation, GO:0050863-regulation of T cell activation and GO:1903037-regulation of leukocyte cell-cell adhesion”. In the MF group, the significantly enriched terms were “GO:0005125-cytokine activity, GO:0005126-cytokine receptor binding, GO:0004896-cytokine receptor activity, GO:0008009-chemokine activity, GO:0042379-chemokine receptor binding”. “GO:0009897-external side of plasma membrane, GO:0042611-MHC protein complex, GO:0030667-secretory granule membrane, GO:0070820-tertiary granule, GO:0070821-tertiary granule membrane” were found as the top five enriched terms in CC group (Fig. 3D). Meanwhile, the results indicated by KEGG pathways analysis showed significantly enriched pathways were “Cytokine-cytokine receptor interaction, Hematopoietic cell lineage, Cell adhesion molecules (CAMs), Chemokine signaling pathway and Graft-versus-host disease” (Fig. 3E). These functional annotations were all immune-related, and suggested possible pathways or functional imbalances that cause EC to occur.

3.4 | Identification Of Prognostic Immune-related Degs

To further screened out the DEGs with potentially prognostic value for EC patients, all prognostic DEGs were subjected to survival analysis by comparing OS in high- and low- expression groups based on median expression. As a result, 225 genes were found associated with OS among the 938 intersection DEGs ($p < 0.05$), and were considered as prognostic related genes for further study. The top 9 genes were showed in Supplementary Fig. 1.

3.5 | Construction Of The Ppi Networks

The PPI networks based on 225 DEGs were obtained from STRING. The network of 225 prognostic immune-related genes contained 224 nodes and 2408 edges. The top 10 outstanding proteins in the PPI network were IL2, CD2, LCP2, CTLA4, CCR5, CCR7, CD28, GZMB, PRF1 and IL2RB (Supplementary Fig. 2A), which might play critical roles in oncogenesis of EC. Totally, eight clusters were generated in MCODE, and we selected the top 3 modules as hub-clusters by the scores calculated in MCODE (Supplementary Fig. 2B-D). Among these three hub clusters, MCODE 1, which had 55 nodes and 998 edges. The functional annotations of these three hub-clusters were further investigated in “Clusterprofiler”, and pathways enriched in these hub-clusters were all chemokine-related, which were consistent with the results above (Supplementary Fig. 3).

3.6 | Construction of prognostic models based on immune-related DEGs

We further performed survival and prognostic analysis for the 225 prognostic DEGs. First,We integrated mRNA expression profiles and clinical information so as to screen out 520 OS-related prognostic and 481 DFS-related prognostic EC samples. The above prognosis EC samples in TCGA were further investigated to define the DEGs which can be prognostic biomarkers. For OS, these samples (n = 520) were randomly split into train(n = 260) and test set (n = 260) at 1:1 ratio. Similarly, for DFS, the samples (n = 481) were randomly split into train(n = 241) and test set (n = 240) at a 1:1 ratio. The genes expression profiles and the corresponding survival information of EC patients in train sets were screened step by step through the univariate Cox regression analysis (Table S1&S2), LASSO Cox regression analysis and multivariate Cox regression analysis (Fig. 4A-D). We tried to explore this gene set from the OS and DFS respectively. Finally, thirteen genes (HLA-A, KIAA1755, EOMES, FCRL6, BATF, C1orf54, APOBR, CTSW, FOXP3, PSTPIP1, IFFO1, CD1D and ASB2) were identified as key prognostic immune-related genes in OS model (Fig. 4E), and four genes (CD1D, SCGB2A1, GPR25 and TMEM114) were identified as key prognostic immune-related genes in DFS model (Fig. 4F).The train set risk score for OS = (-0.00115134 * HLA-A) + (-2.001131455 * expression value of KIAA1755) + (-1.569001602 * expression value of EOMES) + (-0.538427733 * expression value of FCRL6) + (-0.164867433 * expression value of BATF) + (-0.145033591 * expression value of C1orf54) + (-0.556020219 * APOBR) + (-0.232926418 * CTSW) + (-+0.887749365 * FOXP3) + (1.947134037 * PSTPIP1) + (1.520488909 * IFFO1) + (0.445776601 * CD1D) + (-0.597874252 * ASB2). The train set risk score for DFS = (-0.281190502 * expression value of CD1D) + (-0.000227273 * expression value of SCGB2A1) + (-0.780965003 * expression value of GPR25) + (-0.47480137 * expression value of TMEM114).

3.7 | Validation and application of prognostic models in EC

To evaluate the prognostic value of these two risk models, we used the train, test and entire sets concurrently. Based on the OS model, we calculated and ranked the risk scores of endometrial cancer patients in the train sets, as well as marked each patient’s survival status on the dot plot for distribution displaying. Meanwhile, according to the median risk score, patients in the train set were classified into high- or low-risk groups. The EC patients in entire and test set were respectively separated into high- and low-risk group according to the train set's cut off risk score. The distribution of risk score, survival status, and the expression of thirteen genes in OS model for each patient were showed in Fig. 5A-C and Table 3. The results of Kaplan-Meier survival analysis showed the overall survival time of patients in the low-risk group was significantly longer than those in the high-risk group (Fig. 5D). Beside this, the 1-, 3- and 5-year ROC curves based on train set were plotted as Fig. 5E, and the signatures was 0.869 at 1 year, 0.897 at 3 years, and 0.876 at 5 years. The results of survival analysis and the AUC of ROC curve in test and entire sets were similar with above. As for DFS, according to prediction value of DFS model, patients were also divided into high-and low-risk groups based on the cut-off of the median risk in train set. The EC patients

in entire set were divided into high- and low-risk group according to the train set's cut off risk score The distribution of risk score, survival status, and the expression of four genes in DFS model for each patient were showed in Supplementary Fig. 4A-C and Table 4. Although the AUC value maintained between 0.6–0.7 which were slightly lower than OS model, the DFS model still could stratify patients into subgroups with different outcomes (Supplementary Fig. 4D) ($p < 0.05$). Besides, we performed subgroup analysis in age (≤ 60 and > 60), grade (G1 & G2 and G3 & G4), stage (stage Ⅰ & stage Ⅱ and stage Ⅲ & stage Ⅳ) and histological type (endometrial and mixed & serous) in OS and DFS model, and we found that patients with high-risk scores were inclined to have shorter OS in all the subgroups ($p < 0.05$) (Supplementary Fig. 5). In addition, the heatmap showed the association of prognostic gene expression and clinical information, and significant differences were found between the high- and low-risk groups associated with grade, histological type and stage (Supplementary Fig. 6).

Table 3

Univariate and multivariate Cox regression analysis in each group with

Variables	Univariate analysis			Multivariate analysis		
	HR	95%CI	p	HR	95%CI	p
TCGA training group						
Age (≤ 60 vs 60)	1.54	0.80–2.98	0.20	1.06	0.52–2.15	0.88
Stage(stage I & stage II vs stage III & stage IV)	3.86	2.07–7.20	0.00	2.46	1.24–4.86	0.01
Histological type (endometrial vs mixed & serous)	4.19	2.25–7.79	0.00	2.41	1.09–5.33	0.03
Grade (G1&G2 vs G3&G4)	3.42	1.58–7.41	0.00	1.77	0.72–4.38	0.21
Risk score(high/low)	1.00	1.00–1.00	0.00	1.00	1.00–1.00	0.00
TCGA testing group						
Age (≤ 60 vs 60)	2..07	1.05–4.06	0.03	1.92	0.95–3.91	0.07
Stage(stage I & stage II vs stage III & stage IV)	4.67	2.62–8.34	0.00	3.84	2.10–7.13	0.00
Histological type (endometrial vs mixed & serous)	2.36	1.34–4.18	0.00	1.28	0.66–2.48	0.47
Grade (G1&G2 vs G3&G4)	3.32	1.55–7.09	0.00	1.86	0.79–4.39	0.16
Risk score (high/low)	1.00	1.00–1.00	0.00	1.00	1.00–1.00	0.00
TCGA entire group						
Age (≤ 60 vs 60)	1.78	1.11–2.84	0.02	1.53	0.93–2.49	0.09
Stage(stage I & stage II vs stage III & stage IV)	4.12	2.70–6.28	0.00	3.22	2.06–5.04	0.00
Histological type (endometrial vs mixed & serous)	3.04	2.00-4.62	0.00	1.53	0.93–2.50	0.09

OS of EC patients.

Bold values indicate P < 0.05.

Variables	Univariate analysis			Multivariate analysis		
	HR	95%CI	p	HR	95%CI	p
Grade (G1&G2 vs G3&G4)	3.40	1.98–5.84	0.00	1.96	1.06–3.61	0.03
Risk score(high/low)	1.00	1.00–1.00	0.00	1.00	1.00–1.00	0.00
OS of EC patients.						
Bold values indicate P < 0.05.						

Table 4

Univariate and multivariate Cox regression analysis in each group with

Variables	Univariate analysis			Multivariate analysis		
	HR	95%CI	p	HR	95%CI	p
TCGA training group						
Age (≤ 60 vs 60)	1.41	0.83–2.42	0.21	1.61	0.88–2.94	0.12
Stage(stage I & stage II vs stage III & stage IV)	2.83	1.68–4.74	0.00	2.44	1.39–4.29	0.00
Histological type (endometrial vs mixed & serous)	2.09	1.23–3.55	0.01	0.75	0.39–1.46	0.40
Grade (G1&G2 vs G3&G4)	3.30	1.71–6.37	0.00	1.76	0.81–3.80	0.15
Risk score(high/low)	2.67	1.71–4.17	0.00	2.22	1.33–3.68	0.00
TCGA testing group						
Age (≤ 60 vs 60)	1.57	0.84–2.91	0.15	1.32	0.69–2.54	0.41
Stage(stage I & stage II vs stage III & stage IV)	2.62	1.50–4.57	0.00	2.15	1.21–3.81	0.01
Histological type (endometrial vs mixed & serous)	2.34	1.32–4.16	0.00	2.05	1.00-4.23	0.05
Grade (G1&G2 vs G3&G4)	1.09	0.62–1.93	0.76	0.43	0.20–0.90	0.03
Risk score (high/low)	2.16	1.38–3.39	0.00	2.11	1.27–3.50	0.00
TCGA entire group						
Age (≤ 60 vs 60)	1.45	0.97–2.16	0.07	1.28	0.84–1.97	0.26
Stage(stage I & stage II vs stage III & stage IV)	2.76	1.89–4.03	0.00	2.32	1.56–3.45	0.00
Histological type (endometrial vs mixed & serous)	2.23	1.51–3.29	0.00	1.28	0.80–2.06	0.31

DFS of EC patients.

Bold values indicate P < 0.05.

Variables	Univariate analysis			Multivariate analysis		
	HR	95%CI	p	HR	95%CI	p
Grade (G1&G2 vs G3&G4)	1.87	1.23–2.85	0.00	0.85	0.51–1.42	0.53
Risk score(high/low)	2.41	1.76–3.30	0.00	2.17	1.52–3.09	0.00
DFS of EC patients.						
Bold values indicate P < 0.05.						

Since risk score has such a distinctive distinction, can it become an independent predictor? By conjoint univariate and multivariate analyses in train,test and entire sets, we found that risk score (both in OS and DFS model) might be independent risk factors in EC patients (Tables 2 and 3). Meanwhile, by drawing the ROC curve, we found that risk score has better predictive ability than other clinical factors in 1-,3-, and 5-year (Fig. 6A). The AUC of ROC curve increased after combining risk score with other clinical factors (Fig. 6B), which suggested that the risk score might contribute to be an independent risk factor for patients. The nomogram integrating risk score, age, stage, grade and histological type was introduced (Fig. 6C). In the calibration plots, the actual survival and nomogram-predicted probability survival showed amazing concordance in 1-, 3- and 5-year OS (Fig. 6D-F). The outcome of DFS were similar with OS (Supplementary Fig. 7). Taking together, these findings indicated that compared with nomograms built with a single prognostic factor, the nomogram built might be well nomogram for predicting OS and DFS for patients with EC, which might help clinical management.

3.8 | Gene Set Enrichment Analyses

Furthermore, the transcript message of patients stratified by risk score into high- and low- risk subgroups were analyzed by GSEA. In the OS model, representative KEGG pathways in low- risk patients were “allograft rejection” “autoimmune thyroid disease”, “cytokine-cytokine receptor interaction”, “graft versus host disease” and “intestinal immune network for IgA production”. Besides, representative KEGG pathways in high- risk patients were “basal transcription factor”, “cell cycle”, “nucleotide excision repair”, “rna degradation”, and “splicesome” (Fig. 7A). While in the DFS model, in addition to those pathways mentioned in the OS model, “DNA replication”, “homologous recombination” and “mismatch repair” pathways were added in the high-risk group. No matter in OS model or DFS model, pathways enriched in the low-risk group were immune-related; pathways enriched in the high-risk group were not immune-related but frequently dysregulated pathways in cancer (Fig. 7B).

3.9 | Estimation Of Immune Cell Type Fractions

Through the above investigation, we verified the relationship between endometrial cancer and immunity utilizing RNA-seq data. So which kind of immune cells play the main role? By using CIBERSORT, we estimated the differences in the immune infiltration of 22 immune cell types between low- and high-risk groups. Firstly, bar plot was employed to represent the proportion of 22 immune cells in each sample (Supplementary Fig. 8A). The heat map of 22 immune cells in each sample was also performed in Supplementary Fig. 8B. Secondly, various immunocyte subpopulation fractions showed a low to moderate correlation (Supplementary Fig. 8C). Thirdly, we compared the differential infiltration of 22 immune cells between normal tissues and EC tissues. The results showed that relative to para-carcinoma tissues, higher proportions of native B cells, resting CD4 + memory T cells, gamma delta T cells monocytes and resting Mast cells could be detected in cancer tissues. While plasma cells, follicular helper T cells, regulatory T cell (Tregs) and M0 macrophages showed lower proportions in cancer tissues (Supplementary Fig. 8D). In this paper, we analyzed the tumor infiltrating cells between high- and low-risk subgroups. The results showed that Macrophages M0, T cells regulatory (Tregs), T cells CD8, Mast cells resting and Neutrophils infiltrated differently in OS model (Fig. 8), and Dendritic cells resting, T cells regulatory (Tregs), Dendritic cells activated, T cells CD8, Macrophages M0, Plasma cells and NK cells resting infiltrated differently in DFS model (Supplementary Fig. 9). What's more, patients with different infiltrating conditions of Tregs and gamma delta T cells may possess with different clinical outcomes (Supplementary Fig. 10A).

3.10 | The prognostic value of immune checkpoint modulators and the correlation with Tumor-infiltrating immune cells in endometrial cancer.

Immune checkpoint proteins play important roles in the immune response and have multiple mutual interactions with immune cells. 17 molecules (including CD27, CD40, CD70, CD270, B7-H3, B7-H4, IDO1, PD-1, PD-L1, PD-L2, TIM-3, TIGIT, CTLA4, CD86, ICOS, LAG3, and CD58) were reported to be crucial immune modulators [23]. Firstly, we analyzed the prognosis of these molecules, among which 11 molecules showed statistical significance ($p < 0.05$) in EC (Supplementary Fig. 10B). Then, the correlations between 17 immune checkpoint modulators and 22 immune cells were evaluated. As illustrated in Fig. 9A, immunomodulators showed mostly positive correlation with "$CD8^+$ T cells", "activated memory CD4 + T cells" and "M1 Macrophages", and showed negative correlation with "resting memory CD4 + T cells", "M0 Macrophages" and "activated Dendritic cells". And we found that most of these immunomodulators were significantly increased in high-immune (Fig. 9B) or high-stromal score group (Fig. 9C) except B7-H3 and B7-H4.

3.11 | The immune-related risk signature and response to immune checkpoint inhibitor

Combining previous research, we found not only the four scores mentioned above but also the expression of PD-1, CTLA4, PD-L1 and PD-L2 increased significantly in low-risk score subgroup in OS (Fig. 10). The results of DFS were showed in Supplementary Fig. 11.

3.12 | Evaluation the immune status between low-risk and high-risk groups

Except for the IPS used for the estimation of patients' potential to be placed on immune checkpoint inhibitor, we also evaluate the relationship between tumor purity and the immune-related risk signature. The ssGSEA was performed to investigate sthe expression profiles of the 29 immune signature gene sets respectively in high- and low-risk group. The immune status of the low-risk and high-risk samples showed a certain degree of heterogeneity in OS (Fig. 11A) and DFS model (Fig. 11B). The low-risk subgroup showed not only more immune activities, but also significantly lower in tumor purity. This could be caused by the higher infiltrated immune and stromal status in low-risk group. Tumor purity of low- and high-risk groups were performed by violin plots (Fig. 11C-D). For HLA related genes play a crucial roles in immune regulation, we further analyzed and found the expression of key HLA genes wes significantly higher in the low-risk group by comparing that in high-risk group (Fig. 11E-F).

3.13 | The immune related risk signature and mutation profile

Gene mutations are an important cause of tumorigenesis and development. Hence, we evaluated tumor mutational burden (TMB) of patients in low- and high-risk group with somatic mutation data. In OS model, the high-risk group had somatic mutations in the following order: PTEN > PIK3CA > TP53 > ARID1A > TTN > PIK3R1 > KMT2D > CCNNB1 > CSMD3 > RYR2 (Fig. 12A) and in the low-risk group: PTEN > PIK3CA > ARID1A > TTN > PIK3R1 > CTCF > TP53 > KMT2D > MUC16 > ZFHX3 (Fig. 12B). In DFS model,The high-risk group had somatic mutations in the following order: TP53 > PTEN > PIK3CA > ARID1A > TTN > KMT2D > CHD4 > CSMD3 > ZFHX3 > MUC16 (Fig. 12E) and in the low-risk group: PTEN > ARID1A > PIK3CA > TTN > PIK3R1 > CTCF > CTNNB1 > MUC16 > MUC5B > KMT2D (Fig. 12F). Compared with high-risk patients, TMB was higher in low-risk patients both in OS and DFS model. However, the results was not statistically significant in two models. In OS model, $p = 0.169$ (Fig. 12C), and in DFS, $p = 0.953$ (Fig. 12G). Furthermore, we found that TMB was not associated with OS ($p = 0.096$) (Fig. 12D) but associated with DFS ($p = 0.009$) (Fig. 12H).

Discussions

Immunotherapy has shown therapeutic efficacy in the treatment of some cancers, like hematological malignancies and melanoma [24] [25], but its efficacy in EC remains untested [26]. Previous studies have showed that the immune system is active in the endometrium of both normal and EC individuals [27]. To

understand the relationship between EC and TME, we used the ESTIMATE algorithm to estimate the risk scores of immune cells and stromal cells. Then the ESTIMATE score based on the combination of both types of cells was calculated. We next explored the association between these scores and the clinical features of 587 samples. We found that the lower scores were represented in the high-grade tumor, which suggested that the immune molecules in TME might be associated with EC tumor differentiation. Jones Nathaniel L et al, also suggested that high grade tumors appeared to be more immunogenic than low grade tumors and might preferentially benefit from immunotherapy [28]. By survival analysis, we found both the immune-cell and ESTIMATE scores were positively correlated with overall survival rate, but the score of stromal cells was not significantly with the survival time. It suggested that immune cells might play more crucial roles in EC microenvironment. However, there is a lack of evidence to confirm the function of immune cells in EC.

Based on the expression profiles of EC-related genes and the scores of immune and stromal cells, we identified 799 up- and 139 down-regulated immune-related DEGs. Through GO and KEGG pathway analyses, we explored the underlying mechanism of these DEGs in EC. These DEGs were mostly involved in processes like immune cell activation and cell-cell adhesion. As the results showed, the enrichment terms and KEGG pathways were mostly involved in immune regulation, especially related to the interaction of cytokines and cytokine receptors. The roles of cytokine and cytokine receptor have been gradually identified in cancers, but never in EC. Liao Shengbin et al, suggested that cytokine IL-24 suppressed tumor growing by inducing cell apoptosis and inhibiting angiogenesis in EC [29]. We then used 225 prognostic immune-related DEGs to construct a PPI network, and some genes with high connectivity were screened out. IL-2 can maintain the function of T cells and has been used in immunotherapies for human cancers, like melanoma [30, 31]. CTLA4, a negative regulator of T-cell activation, was approved as the first immune checkpoint in the immunotherapy of melanoma [32]. The hub genes we found in PPI network might also play a vital role in EC. In addition, the top three significant modules in PPI network were identified, and their function annotations can help us better understand the pathogenesis of EC. We then constructed two risk liner models respectively for OS and DFS based on the 225 prognostic immune-related genes. To validate the prognostic value of two models, the entire set was randomly divided into a trainset and a test set at 1:1 ratio. Based on the train set, thirteen genes (HLA-A, KIAA1755, EOMES, FCRL6, BATF, C1orf54, APOBR, CTSW, FOXP3, PSTPIP1, IFFO1, CD1D and ASB2) in OS model, and four genes (CD1D, SCGB2A1, GPR25 and TMEM114) in DFS model were identified as key prognostic immune-related genes. The results of survival analysis showed low-risk patients had better prognosis in each set in OS- and DFS-risk model. Moreover, the AUC of the ROC curve suggested both the two risk models were efficient in survival prognosis. Besides, we performed stratification analysis in the high- and low-risk groups in each model. We found that patients with low-risk scores were inclined to longer survival in all the subgroups. In addition, we observed the high- and low-risk groups showed significantly different grade, histological type and stage. By univariate and multivariate analyses of the risk score and other clinical risk factors in entire sets, we found that risk score might act as an independent risk factor for prognosis of EC patients both in OS- and DFS- risk model. The ROC curve also suggested that risk score was more reliable than other factors in predicting the EC prognosis than other clinical risk factors in both

OS- and DFS- risk models. And we built a nomogram incorporating the risk score, age, stage, grade and histological type to visualize the prognosis of EC patients. Taking together, both OS- and DFS-risk models might perform well in predicting the prognosis of EC patients.

Among these candidate genes in OS- and DFS-risk model, HLA-A was reported to be associated with non-Hodgkin lymphoma [33]. KIAA1755 was found as an effective biomarker for early diagnosis, prognosis, and treatment response in ovarian cancer [34]. EOMES was related with NK cell properties, and EOMES was considered as an immunotherapeutic target in chimeric AgR T cell anti-cancer therapy [35]. Increased expression of FCRL6 was identified in the microenvironment of MHC-II + tumors, meaning FCRL6 may act as an immunotherapy target [36]. BATF, a transcription factor, was found promote cell proliferation in non-small cell lung cancer [37]. BATF can improve T cell persistence and effector function to improve the efficacy of adoptive cell therapy for cancers [38]. CTSW was considered as a tumor-suppressor highly specific to immune cells, and its expression was positively associated with the prognosis of breast cancer [39]. FOXP3-expressing Treg cells have been widely studied in tumor microenvironment, often functioning as immunosuppressor in cancers, including hepatocellular carcinoma and gastric cancer [40, 41]. Previous study suggested that chromosome containing the PSTPIP1 gene was absent in prostate cancer [42]. IFFO1 was associated with chromosome translocation during tumorigenesis [43]. High expression of ASB2 worsened the outcome of colorectal cancer through enhancing Th2 function and diminishing type 1-antitumor immune response [44]. SCGB2A1 was found differentially expressed in histological types of ovarian cancers and might represent an attractive target for the immunotherapy in recurrent tumor with chemotherapy resistance [45]. CD1D was identified as a key gene in both signatures. CD1D, a kind of widely expressed lipid antigen presenting molecule, modulates the development of CD1d-restricted invariant natural killer T (iNKT) cells, and the performance of M1-macrophages and effector Th1 cells in tumor microenvironment [46]. The progression of chronic lymphocytic leukemia was associated with high CD1D expression [47]. Scarce research has explored the roles of CD1D and iNKT cells in EC. Other genes (C1orf54, APOBR, GPR25 and TMEM114) are also ignored in research on EC mechanisms.

GSEA showed the pathways enriched in the low-risk group, including “allograft rejection” “autoimmune thyroid disease”, “cytokine-cytokine receptor interaction”, “graft versus host disease” and “intestinal immune network for IgA production”. Though pathways enriched in the high-risk group were not immune-related, most of them were frequently dysregulated, such as “cell cycle”, “nucleotide excision repair”, “rna degradation”, and “splicesome”. The pathway analysis indicated that infiltrating immune cells activated the immune response and promoted the progression of EC.

Then, we focused on the relationship between infiltrating immune cells and EC. We compared the infiltration of 22 immune cells between high- and low-risk patients and investigated their relationship with EC prognosis. The results showed that the infiltration of Treg cells, CD8 + T cells was higher in low-risk group ,while the Macrophages M0 showed lower infiltration in the low-risk groups in both OS and DFS models. What’s more, it is worth noticing that a higher level of T (γδT) cells was indicative of a shorter survival time. Conversely, the level of Treg cells was lower in cancer tissue and high-risk group, and the

higher infiltration of Treg cells represented the longer survival time. Previous study suggested the high intraepithelial CD8 + T cell count represents an independent favorable prognostic factor in grade 3 endometrial carcinoma [48, 49]. Microsatellite instability-high endometrial cancers with increased tumor-associated CD8 + T lymphocytes might benefit from this therapy [50]. Treg cells have been characterized in kinds of cancers, always suppressing tumor-specific immune responses and promoting tumor progression. Blocking T cells (Tregs) into TME is a promising strategy for tumor immunotherapy [51, 52]. Γδ T cells, for their MHC-unrestricted cytotoxicity, were used in several clinical trials to treat cancer and distinguish transformed cells and normal cells [53, 54]. Our results suggested that Γδ T cells and Treg cells were crucial to the development of EC, which enables them possible targets in EC immunotherapies. Among the 11 immune molecules in EC, most of them were significantly increased in high-immune or high-stromal score group, suggesting that they might play a crucial role in EC TME. Hence, by exploring the response of the immune checkpoint inhibitors, we further found that the expression and IPS of immune checkpoints PD-1, CTLA4, PD-L1 and PD-L2 increased significantly in low-risk subgroup. In the previous study, anti-PD-1/PD-L1 antibodies therapy has shown efficacy in treating endometrioid endometrial cancers with microsatellite instability[55], but clinical evidences supporting the effectiveness of anti-CTLA4 therapy in EC have not been reported so far. Combining the previous studies and our present findings, we speculate that EC patients with low-risk score might possess better response to checkpoint inhibitors for PD-1, CTLA4, PD-L1 and PD-L2. Then, ssGSEA and ESTIMATE algorithm were used to evaluate the overall immune status and tumor purity in EC TME. The results of ssGAEA uncovered the overall immune activity of low-risk groups was higher than that of the high-risk group. The analysis of tumor purity also suggested the high-risk group possessed higher tumor purity, suggestive of the lower infiltration of immune and stromal cells. HLA- related genes play a key role in immune system, so we analyzed the HLA-related genes expression in high- and low- risk group, finding that HLA- related genes were significantly higher in the low-risk group. These finding also proved higher immune status was associated with tumor prognosis of EC. T cells can recognize neoantigens via HLA molecules on the tumor cell surface, which provides an opportunity to initiate specific and effective anti-cancer immune responses through an appropriate immunotherapy [56].

Based on the risk signatures, we further explored the TMB of immune-related genes. TMB is a significant factor factor affecting responses to immunotherapy, and patients with endometrial cancer have a higher median TMB than patients with cervical and ovarian cancer, indicating that immune checkpoint blockade might be an effective treatment [57]. Previous study have demonstrated that immunotherapy was more effective in treating microsatellite-stable/TMB-high EC tumors than microsatellite instability-high EC cancer [58]. In our present study, TMB of genes PTAN, PIK3CA, ARID1A, TTN and PIK3R1 was higher in low-risk EC patients. PTEN, a major tumor-suppressor protein, can serve as a suppressor in EC. The deletion or mutations of PTEN may favor tumor cell growth [59, 60]. PIK3CA is a gene most frequently mutated in human cancers, like EC [61]. Previous study have suggested that gene PIK3R1 mutated in at least 5% of glioblastoma multiforme [62]. The mutations of gene TTN was rarely reported in the

progression of cancers. The mutation of these immune related genes might play a crucial role in immunoregulation of EC, which deserves further validation.

However, there are also some limitations in our study. First, our present study was completely researched based on TCGA databases and the results should be validated by external database and further experiments. Second, the predictive value of the immune-related prognostic model should be experimentally tested by a large number of EC samples. Third, the biological functions of immune-related genes based on OS and DFS models need to be examined by a series of cell function assay.

Conclusions

In conclusion, our current study screened out significant immune-related genes in EC and built prognostic risk signatures. According to the models based on immune-related genes, we explored the potential roles of TME molecules in EC, including the immune cells, immune checkpoint modulators and immune checkpoint inhibitors. In summary, our study proved that high-immune TME might prevent the progression of EC, which indicated the possibility of using immunotherapy to fight against EC.

Abbreviations

EC
Endometrial carcinoma
TME
tumor microenvironment
TCGA
The Cancer Genome Atlas
Tregs
Blocking T cells
GO
gene ontology
KEGG
Kyoto Encyclopedia of Genes and Genomes
PPI
protein–protein interaction
STRING
Search Tool for the Retrieval of Interacting Genes Database
MCODE
Molecular Complex Detection
ROC
Receiver operating characteristic
iNKT
invariant natural killer T

CAMs
Cell adhesion molecules
MAF
Mutation Annotation Format
TCIA
The Cancer Immunome Atlas
GSEA
Gene set enrichment analysis

Declarations

Funding

National Natural Science Foundation (81472442, 81872119) and Jiangsu province medical innovation team (CXTDA2017008) supported our study.

Conflicts of interest

The authors have no conflicts of interest to declare.

Availability of data and material

All data, models, and code generated or used during the study appear in the submitted article.

Authors' contributions

Authors Wenjun Cheng and Jinhui Liu designed the project. Authors Jinhui Liu, Sipei Nie and Zhipeng Wu contributed on data analysis and prepared the main manuscript. All authors reviewed the manuscript.

Ethical Approval and Consent to participate

Not applicable

Consent for publication

Written informed consent for publication was obtained from all participants.

Acknowledgements

Not applicable

References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin. 2015. 65(1): 5-29.

[2] Moore K, Brewer MA. Endometrial Cancer: Is This a New Disease. Am Soc Clin Oncol Educ Book. 2017. 37: 435-442.

[3] Njoku K, Chiasserini D, Whetton AD, Crosbie EJ. Proteomic Biomarkers for the Detection of Endometrial Cancer. Cancers (Basel). 2019. 11(10).

[4] Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med. 2013. 19(11): 1423-37.

[5] Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. J Clin Invest. 2015. 125(9): 3335-7.

[6] Sahoo SS, Zhang XD, Hondermarck H, Tanwar PS. The Emerging Role of the Microenvironment in Endometrial Cancer. Cancers (Basel). 2018. 10(11).

[7] Felix AS, Weissfeld J, Edwards R, Linkov F. Future directions in the field of endometrial cancer research: the need to investigate the tumor microenvironment. Eur J Gynaecol Oncol. 2010. 31(2): 139-44.

[8] Pasanen A, Ahvenainen T, Pellinen T, Vahteristo P, Loukovaara M, Bützow R. PD-L1 Expression in Endometrial Carcinoma Cells and Intratumoral Immune Cells: Differences Across Histologic and TCGA-based Molecular Subgroups. Am J Surg Pathol. 2020. 44(2): 174-181.

[9] Wang Z, Zhu J, Liu Y, et al. Development and validation of a novel immune-related prognostic model in hepatocellular carcinoma. J Transl Med. 2020. 18(1): 67.

[10] Zeng Q, Zhang W, Li X, Lai J, Li Z. Bioinformatic identification of renal cell carcinoma microenvironment-associated biomarkers with therapeutic and prognostic value. Life Sci. 2020. 243: 117273.

[11] Zhang C, Zheng JH, Lin ZH, et al. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. Aging (Albany NY). 2020. 12.

[12] Chakraborty H, Hossain A. R package to estimate intracluster correlation coefficient with confidence interval for binary data. Comput Methods Programs Biomed. 2018. 155: 85-92.

[13] Bandettini WP, Kellman P, Mancini C, et al. MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. J Cardiovasc Magn Reson. 2012. 14: 83.

[14] Qu L, Wang ZL, Chen Q, et al. Prognostic Value of a Long Non-coding RNA Signature in Localized Clear Cell Renal Cell Carcinoma. Eur Urol. 2018. 74(6): 756-763.

[15] Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol. 2008. 26(8): 1364-70.

[16] Zhu X, Tian X, Yu C, et al. A long non-coding RNA signature to improve prognosis prediction of gastric cancer. Mol Cancer. 2016. 15(1): 60.

[17] Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol. 2018. 1711: 243-259.

[18] Charoentong P, Finotello F, Angelova M, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. 2017. 18(1): 248-262.

[19] He Y, Jiang Z, Chen C, Wang X. Classification of triple-negative breast cancers based on Immunogenomic profiling. J Exp Clin Cancer Res. 2018. 37(1): 327.

[20] Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013. 4: 2612.

[21] Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018. 28(11): 1747-1756.

[22] Robinson DR, Wu YM, Lonigro RJ, et al. Integrative clinical genomics of metastatic cancer. Nature. 2017. 548(7667): 297-303.

[23] Lim S, Phillips JB, Madeira da Silva L, et al. Interplay between Immune Checkpoint Proteins and Cellular Metabolism. Cancer Res. 2017. 77(6): 1245-1249.

[24] Nakazawa Y. [Chimeric antigen receptor T-cell therapy for hematological malignancies]. Rinsho Ketsueki. 2019. 60(9): 1351-1357.

[25] Rodríguez-Cerdeira C, Carnero Gregorio M, López-Barcenas A, et al. Advances in Immunotherapy for Melanoma: A Comprehensive Review. Mediators Inflamm. 2017. 2017: 3264217.

[26] 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: 109-116.

[27] Agrawal U, Kumari N, Mishra AK, et al. Immune signature of urothelial cancer associated with grade, recurrence, and invasion. Urol Oncol. 2016. 34(9): 418.e17-26.

[28] Jones NL, Xiu J, Rocconi RP, Herzog TJ, Winer IS. Immune checkpoint expression, microsatellite instability, and mutational burden: Identifying immune biomarker phenotypes in uterine cancer. Gynecol Oncol. 2020. 156(2): 393-399.

[29] Liao S, Yang Y, Chen S, et al. IL-24 inhibits endometrial cancer cell proliferation by promoting apoptosis through the mitochondrial intrinsic signaling pathway. Biomed Pharmacother. 2020. 124: 109831.

[30] Rosenberg SA. IL-2: the first effective immunotherapy for human cancer. J Immunol. 2014. 192(12): 5451-8.

[31] Marabondo S, Kaufman HL. High-dose interleukin-2 (IL-2) for the treatment of melanoma: safety considerations and future directions. Expert Opin Drug Saf. 2017. 16(12): 1347-1357.

[32] Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017. 541(7637): 321-330.

[33] Zhong C, Gragert L, Maiers M, et al. The association between HLA and non-Hodgkin lymphoma subtypes, among a transplant-indicated population. Leuk Lymphoma. 2019. 60(12): 2899-2908.

[34] Antony F, Deantonio C, Cotella D, et al. High-throughput assessment of the antibody profile in ovarian cancer ascitic fluids. Oncoimmunology. 2019. 8(9): e1614856.

[35] Sasawatari S, Okamoto Y, Kumanogoh A, Toyofuku T. Blockade of N-Glycosylation Promotes Antitumor Immune Response of T Cells. J Immunol. 2020. 204(5): 1373-1385.

[36] Johnson DB, Nixon MJ, Wang Y, et al. Tumor-specific MHC-II expression drives a unique pattern of resistance to immunotherapy via LAG-3/FCRL6 engagement. JCI Insight. 2018. 3(24).

[37] Feng Y, Pan L, Zhang B, Huang H, Ma H. BATF acts as an oncogene in non-small cell lung cancer. Oncol Lett. 2020. 19(1): 205-210.

[38] Wei J, Long L, Zheng W, et al. Targeting REGNASE-1 programs long-lived effector T cells for cancer therapy. Nature. 2019. 576(7787): 471-476.

[39] Zhang Y, Manjunath M, Yan J, et al. The Cancer-Associated Genetic Variant Rs3903072 Modulates Immune Cells in the Tumor Microenvironment. Front Genet. 2019. 10: 754.

[40] Li X, Zhang Y, Ma W, et al. Enhanced glucose metabolism mediated by CD147 contributes to immunosuppression in hepatocellular carcinoma. Cancer Immunol Immunother. 2020 .

[41] Ma K, Li X, Lv J, et al. Correlations between CD4+ FoxP3+ Treg and expression of FoxM1 and Ki-67 in gastric cancer patients. Asia Pac J Clin Oncol. 2020 .

[42] Strefford JC, Lane TM, Hill A, et al. Molecular characterisation of the t(1;15)(p22;q22) translocation in the prostate cancer cell line LNCaP. Cytogenet Genome Res. 2006. 112(1-2): 45-52.

[43] Li W, Bai X, Li J, et al. The nucleoskeleton protein IFFO1 immobilizes broken DNA and suppresses chromosome translocation during tumorigenesis. Nat Cell Biol. 2019. 21(10): 1273-1285.

[44] Spinner CA, Lamsoul I, Métais A, Febrissy C, Moog-Lutz C, Lutz PG. The E3 Ubiquitin Ligase Asb2α in T Helper 2 Cells Negatively Regulates Antitumor Immunity in Colorectal Cancer. Cancer Immunol Res. 2019. 7(8): 1332-1344.

[45] Bellone S, Tassi R, Betti M, et al. Mammaglobin B (SCGB2A1) is a novel tumour antigen highly differentially expressed in all major histological types of ovarian cancer: implications for ovarian cancer immunotherapy. Br J Cancer. 2013. 109(2): 462-71.

[46] Paul S, Chhatar S, Mishra A, Lal G. Natural killer T cell activation increases iNOS+CD206- M1 macrophage and controls the growth of solid tumor. J Immunother Cancer. 2019. 7(1): 208.

[47] Gorini F, Azzimonti L, Delfanti G, et al. Invariant NKT cells contribute to chronic lymphocytic leukemia surveillance and prognosis. Blood. 2017. 129(26): 3440-3451.

[48] Hu J, Sun J. MUC16 mutations improve patients' prognosis by enhancing the infiltration and antitumor immunity of cytotoxic T lymphocytes in the endometrial cancer microenvironment. Oncoimmunology. 2018. 7(10): e1487914.

[49] Vagios S, Yiannou P, Giannikaki E, et al. The impact of programmed cell death-ligand 1 (PD-L1) and CD8 expression in grade 3 endometrial carcinomas. Int J Clin Oncol. 2019. 24(11): 1419-1428.

[50] Crumley S, Kurnit K, Hudgens C, Fellman B, Tetzlaff MT, Broaddus R. Identification of a subset of microsatellite-stable endometrial carcinoma with high PD-L1 and CD8+ lymphocytes. Mod Pathol. 2019. 32(3): 396-404.

[51] Chaudhary B, Abd Al Samid M, al-Ramadi BK, Elkord E. Phenotypic alterations, clinical impact and therapeutic potential of regulatory T cells in cancer. Expert Opin Biol Ther. 2014. 14(7): 931-45.

[52] Wang R, Feng W, Wang H, et al. Blocking migration of regulatory T cells to leukemic hematopoietic microenvironment delays disease progression in mouse leukemia model. Cancer Lett. 2020. 469: 151-161.

[53] Deniger DC, Maiti SN, Mi T, et al. Activating and propagating polyclonal gamma delta T cells with broad specificity for malignancies. Clin Cancer Res. 2014. 20(22): 5708-19.

[54] Paul S, Lal G. Regulatory and effector functions of gamma-delta (γδ) T cells and their therapeutic potential in adoptive cellular therapy for cancer. Int J Cancer. 2016. 139(5): 976-85.

[55] Le DT, Uram JN, Wang H, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. N Engl J Med. 2015. 372(26): 2509-20.

[56] Roerden M, Nelde A, Walz JS. Neoantigens in Hematological Malignancies-Ultimate Targets for Immunotherapy. Front Immunol. 2019. 10: 3004.

[57] Wang M, Fan W, Ye M, et al. Molecular profiles and tumor mutational burden analysis in Chinese patients with gynecologic cancers. Sci Rep. 2018. 8(1): 8990.

[58] Goodman AM, Sokol ES, Frampton GM, Lippman SM, Kurzrock R. Microsatellite-Stable Tumors with High Mutational Burden Benefit from Immunotherapy. Cancer Immunol Res. 2019. 7(10): 1570-1573.

[59] Cancer Genome Atlas Research Network, Kandoth C, Schultz N, et al. Integrated genomic characterization of endometrial carcinoma. Nature. 2013. 497(7447): 67-73.

[60] Pulido R, Mingo J, Gaafar A, et al. Precise Immunodetection of PTEN Protein in Human Neoplasia. Cold Spring Harb Perspect Med. 2019. 9(12).

[61] Myers AP, Filiaci VL, Zhang Y, et al. Tumor mutational analysis of GOG248, a phase II study of temsirolimus or temsirolimus and alternating megestrol acetate and tamoxifen for advanced endometrial cancer (EC): An NRG Oncology/Gynecologic Oncology Group study. Gynecol Oncol. 2016. 141(1): 43-8.

[62] Zhang M, Yang D, Gold B. Origin of mutations in genes associated with human glioblastoma multiform cancer: random polymerase errors versus deamination. Heliyon. 2019. 5(3): e01265.

Supplementary Files Legend

Supplementary Figure 1. Top 9 potential prognostic immune-related genes.

Supplementary Figure 2. Top 30 hub genes in protein-protein interaction (PPI) network, as well as top 3 modules were constructed. (A) Top 30 hub genes. (B-D) Top 3 hub modules were identified by Cytoscape plug-in MCODE.

Supplementary Figure 3. The functional annotations of top three hub-clusters.

Supplementary Figure 4. Risk score analysis, time-dependent ROC analysis and Kaplan–Meier analysis for the validation of prognostic model in train, test and entire sets based on DFS. (A) Rank of risk score and distribution of groups. Patients with EC were divided into low- and high-risk subgroups based on the median value of the risk score calculated. (B) The survival status and survival time of patients with EC ranked by risk score. In (A) and (B), Green dots represent for patients with a low level of risk score and red dots represent for patients with a high level of risk score. (C) Heatmap of included genes in DFS model. Patients were divided into two groups according to risk score. Kaplan–Meier curve (D) and time-dependent ROC analysis (E) of risk score in TCGA cohort.

Supplementary Figure 5. Stratified analysis of clinical factors. The stratified analysis of patients' age (≤60 and >60), stage Ⅰ & Ⅱ, grade (G1&G2 and G3&G4) and histological type.

Supplementary Figure 6. Heatmap showed the expression of the OS (A) and DFS (B) immune-related prognosis genes in high- and low-risk patients in the TCGA dataset. The distribution of clinicopathological characteristics was compared between the high-risk and low-risk groups. *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$

Supplementary Figure 7. The nomogram to predict 1-, 3- or 5-year DFS in EC based on DFS model.

(A) Time-dependent ROC analysis to evaluate the prognostic value of risk score and clinical factors (age, stage, grade and histological type). (B) Time-dependent ROC analysis based on risk score and merged clinical factors. (C) Nomogram for predicting DFS of EC. There were five factors containing age, gender, stage, histological type and risk score in the nomogram. Each of them generates points according to the line drawn upward. And the total points of the seven components of an individual patient lie on "Total Points" axis which corresponds to the probability of 1-year, 3-year and 5-year DFS rate plotted on the two axes below. The calibration plots for predicting patient 1-(D), 3-(E) or 5-(F) year DFS.

Supplementary Figure 8. Comprehensive assessment the fractions of 22 immune cell subtypes in EC using CIBERSORT algorithm. (A) Summary of estimated fractions of 22 immune cell subtypes. Each Bar chart exhibited the cell proportions of each patient and various colors represent the 22 immune cells with annotations aside. (B) Differential abundance of immune cells in EC and adjacent normal tissues shown by heatmap plot. (C) The Pearson correlation coefficient between 22 immune cells. The color from blue to red means the value vary from -1 to 1. (D)The violin plot shows the distinct compositions of 22 immune cell infiltrates normal tissues (blue) and EC tissues (red) by the Wilcoxon rank-sum test.

Supplementary Figure 9. The relationship of DFS immune-related prognosis risk score with the immune cells.

Supplementary Figure 10. Kaplan–Meier survival curves of immunomodulators and immune cells (A) The Kaplan–Meier survival curves showed that patients with different infiltrating conditions of Tregs and gamma delta T cells may possess with different clinical outcomes. (B) 11 immune modulators showed statistical significance with patient’s survival. The yellow line indicates samples with highly expressed genes (above best-separation value), and the green line designates the samples with lowly expressed genes (below best-separation value)

Supplementary Figure 11. IPS and immunotherapy gene expression analysis. (A) The association between IPS and the risk score based on DFS in EC patients. (B) The gene expression of PD1, CTLA4, PD-L1, and PD-L2 in low-risk and high- risk subgroups.

Figures

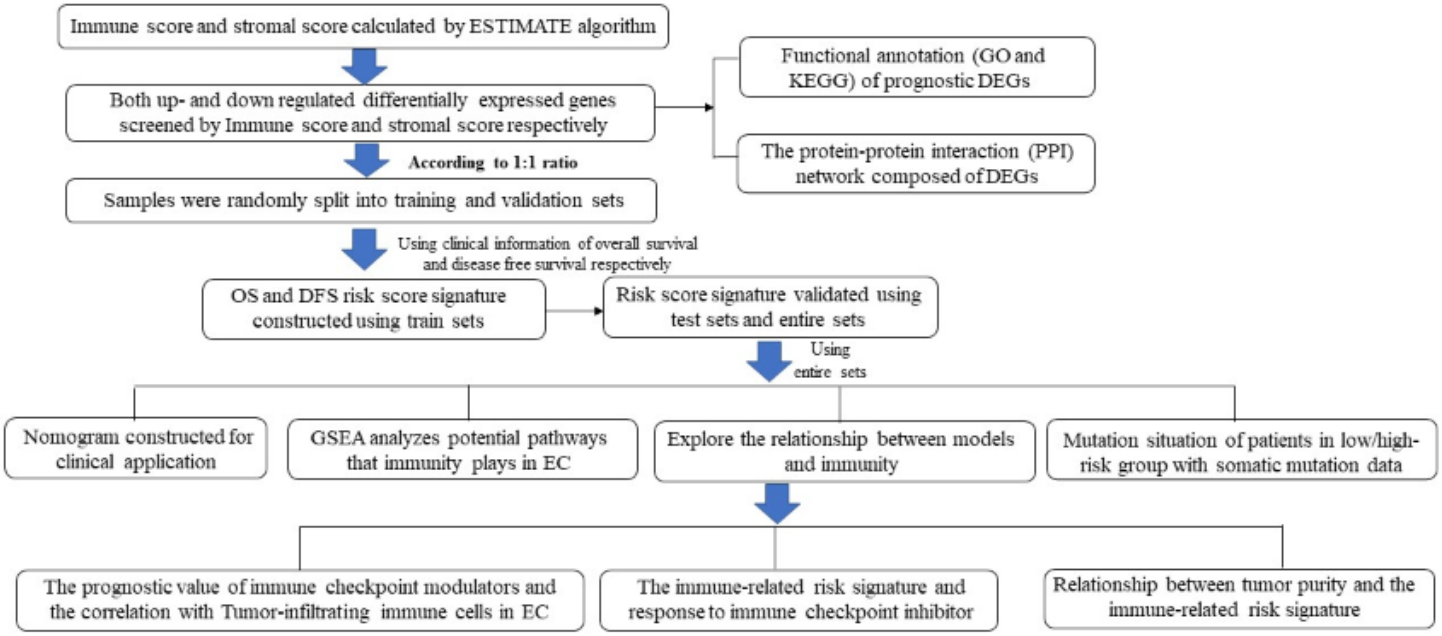

Figure 1

Flow chart of data preparation, processing, analysis and validation in this study.

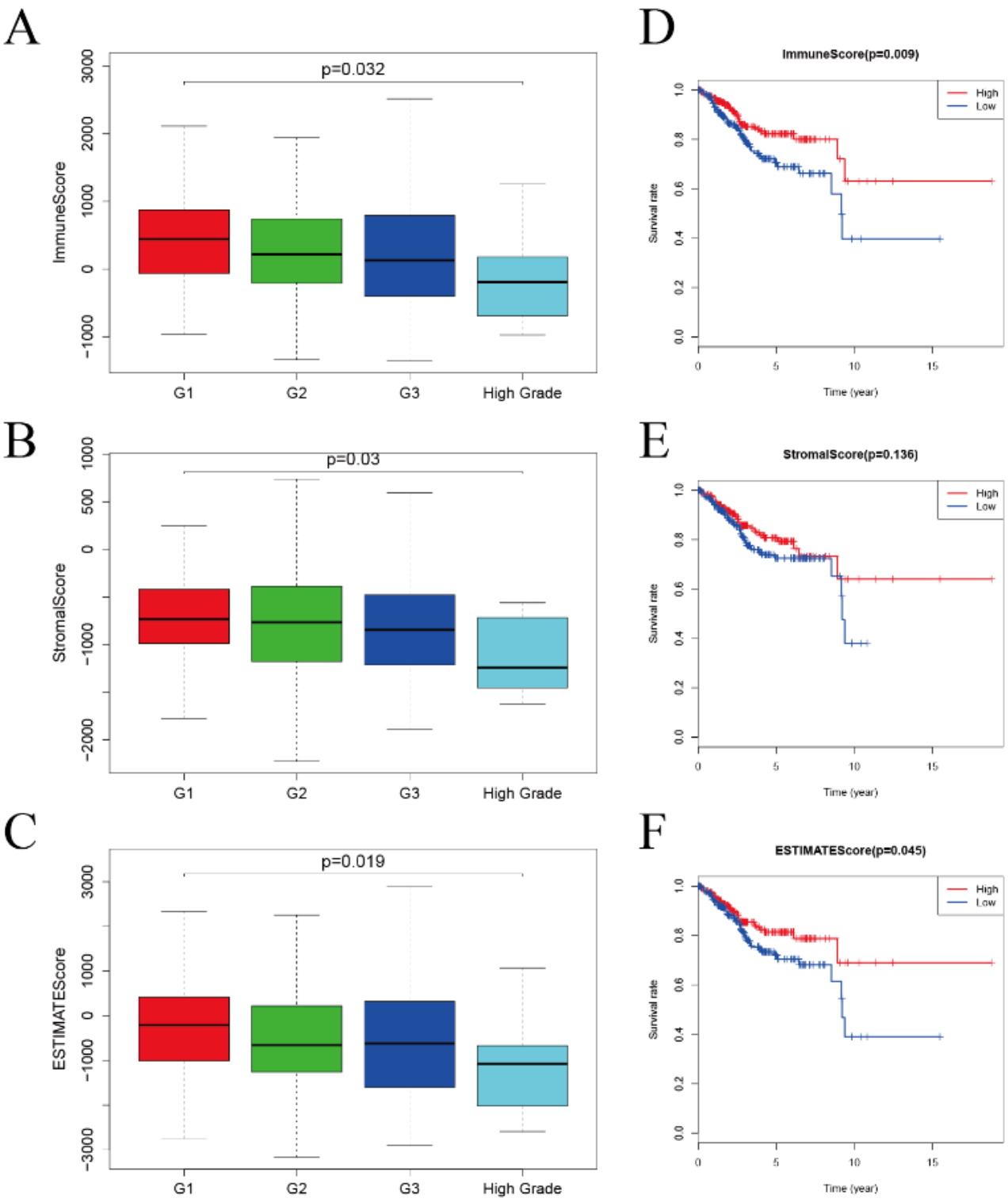

Figure 4

Immune, stromal and ESTIMATE scores were significantly associated with EC patients' grades and prognosis. The AVONA result of Immune (A), stromal (B) and ESTIMATE (C) score in different grades of EC patients. The analysis of patients' overall survival based on Immune (D), stromal (E) and ESTIMATE (F) score.

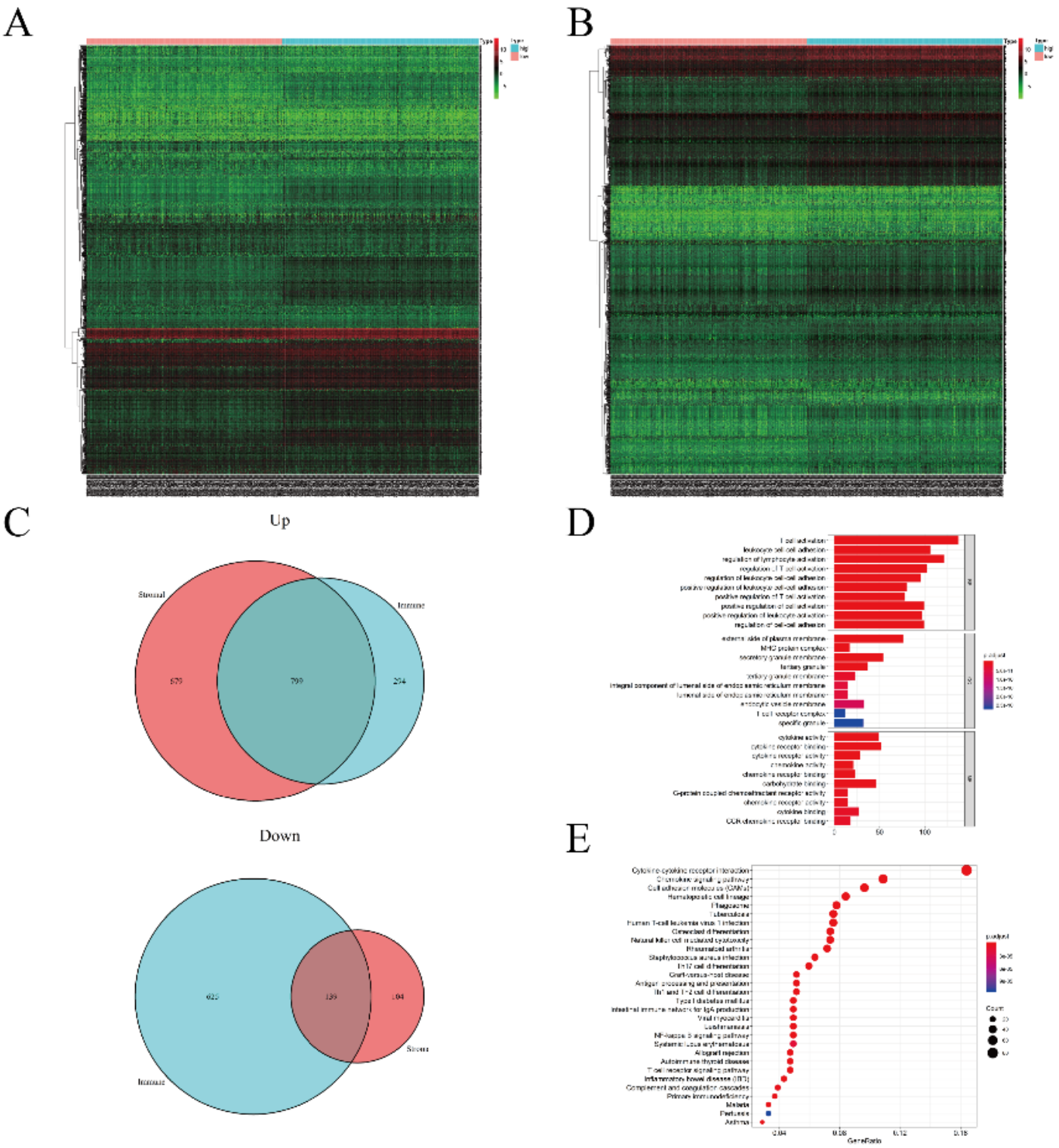

Figure 5

Comparison of gene expression profile with immune and stromal scores of EC. Heatmap of significantly differentially expressed genes based on immune (A) and stromal (B) scores. Red indicates genes with higher expression level and green indicates genes with lower expression. (C) Venn diagram analysis of aberrantly expressed genes based on immune and stromal scores. (D) GO analysis of immune-related

genes. (E) KEGG analysis of immune-related genes. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

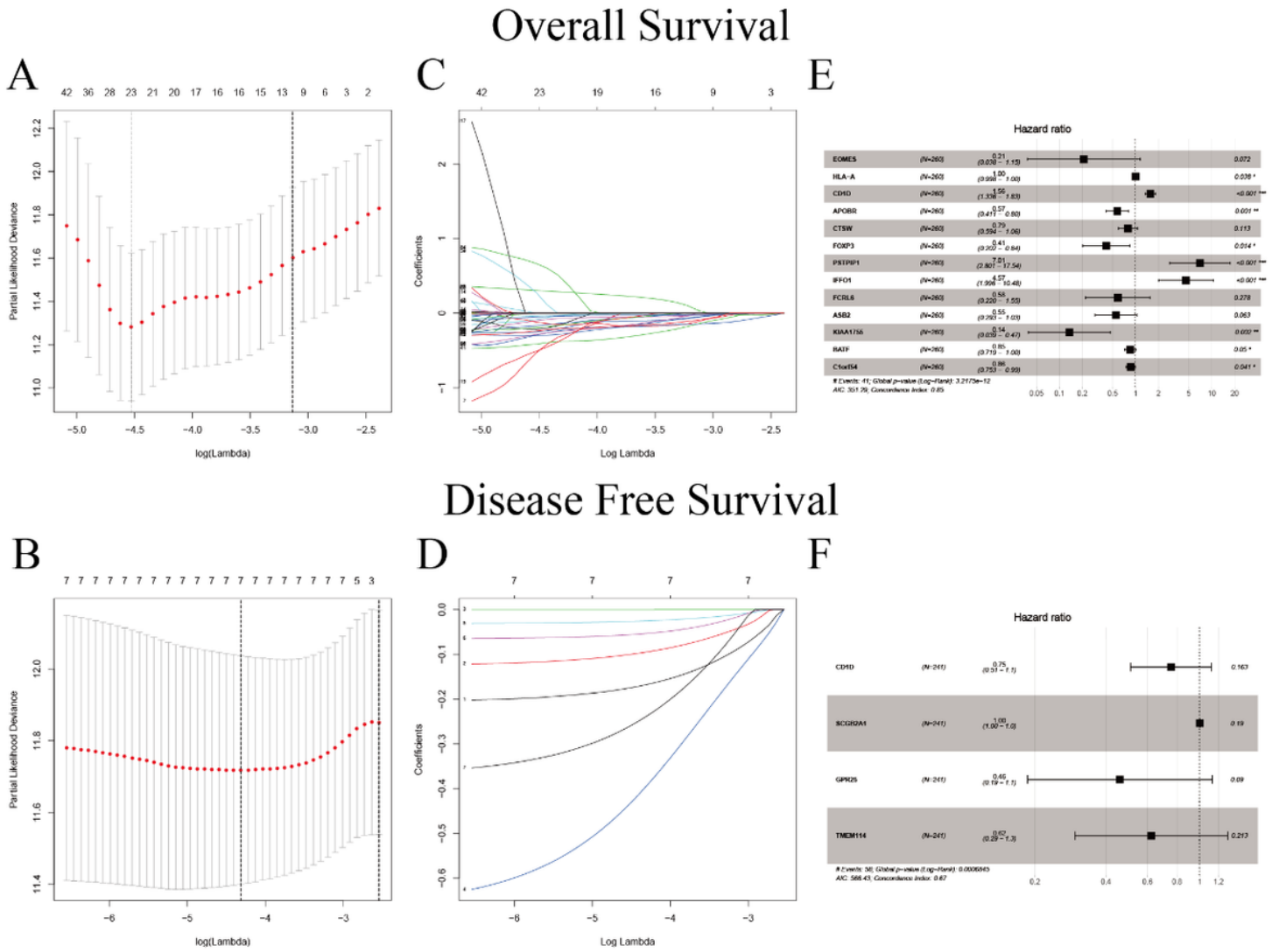

Figure 7

Construction of prognostic models based on immune-related DEGs. LASSO Cox regression analysis and multivariate Cox regression analysis based on OS and DFS (A-D). Forest plots presenting the multivariate Cox proportional hazards regression analysis of prognostic immune-related genes in OS (E) and DFS (F) model.

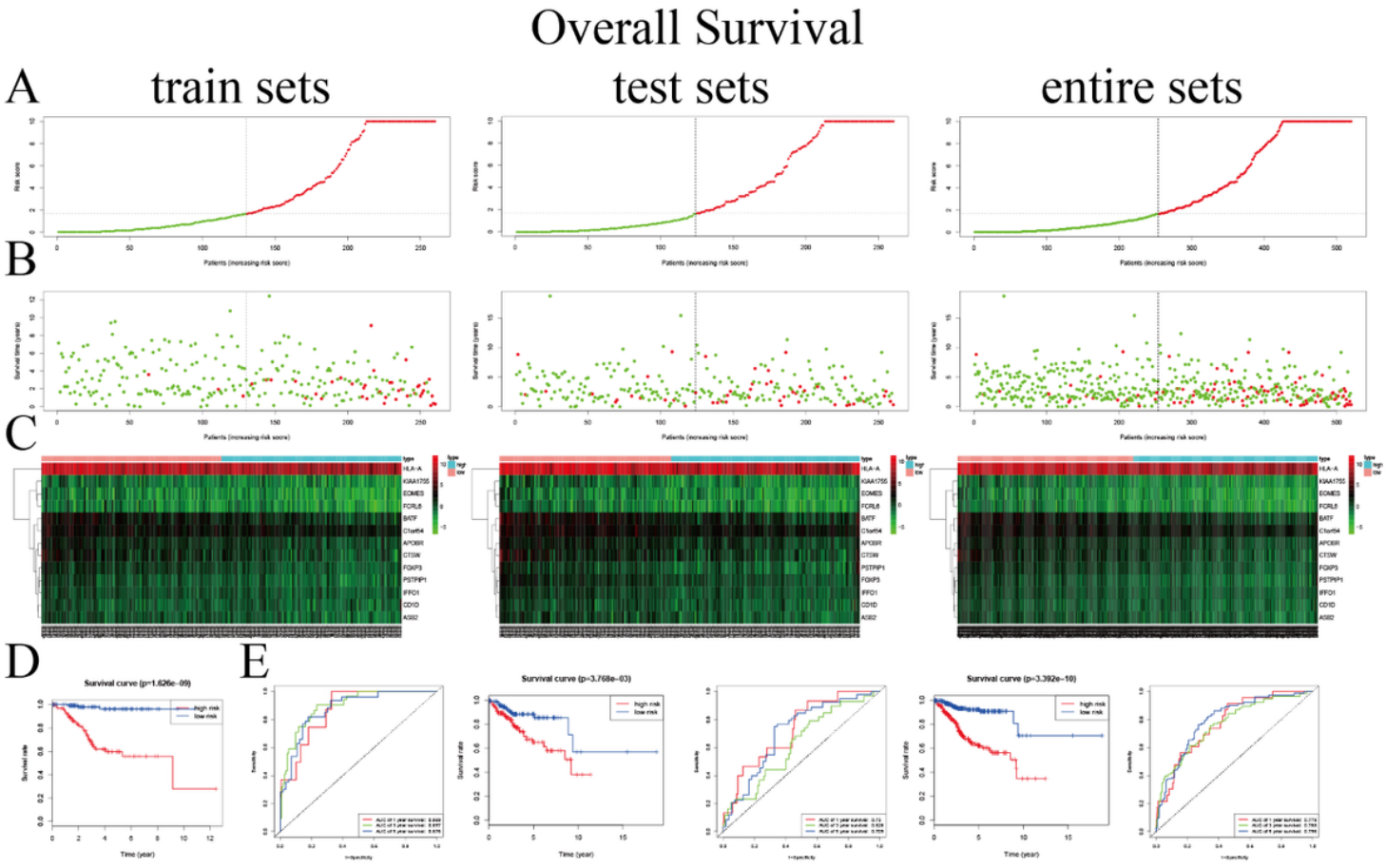

Figure 10

Risk score analysis, time-dependent ROC analysis and Kaplan–Meier analysis for the validation of prognostic model in train, test and entire sets based on OS. (A) Rank of risk score and distribution of groups. Patients with EC were divided into low- and high-risk subgroups based on the median value of the risk score calculated. (B) The survival status and survival time of patients with EC ranked by risk score. In (A) and (B), Green dots represent for patients with a low level of risk score and red dots represent for patients with a high level of risk score. (C) Heatmap of included genes in OS model. Patients were divided into two groups according to risk score. Kaplan–Meier curve (D) and time-dependent ROC analysis (E) of risk score in TCGA cohort.

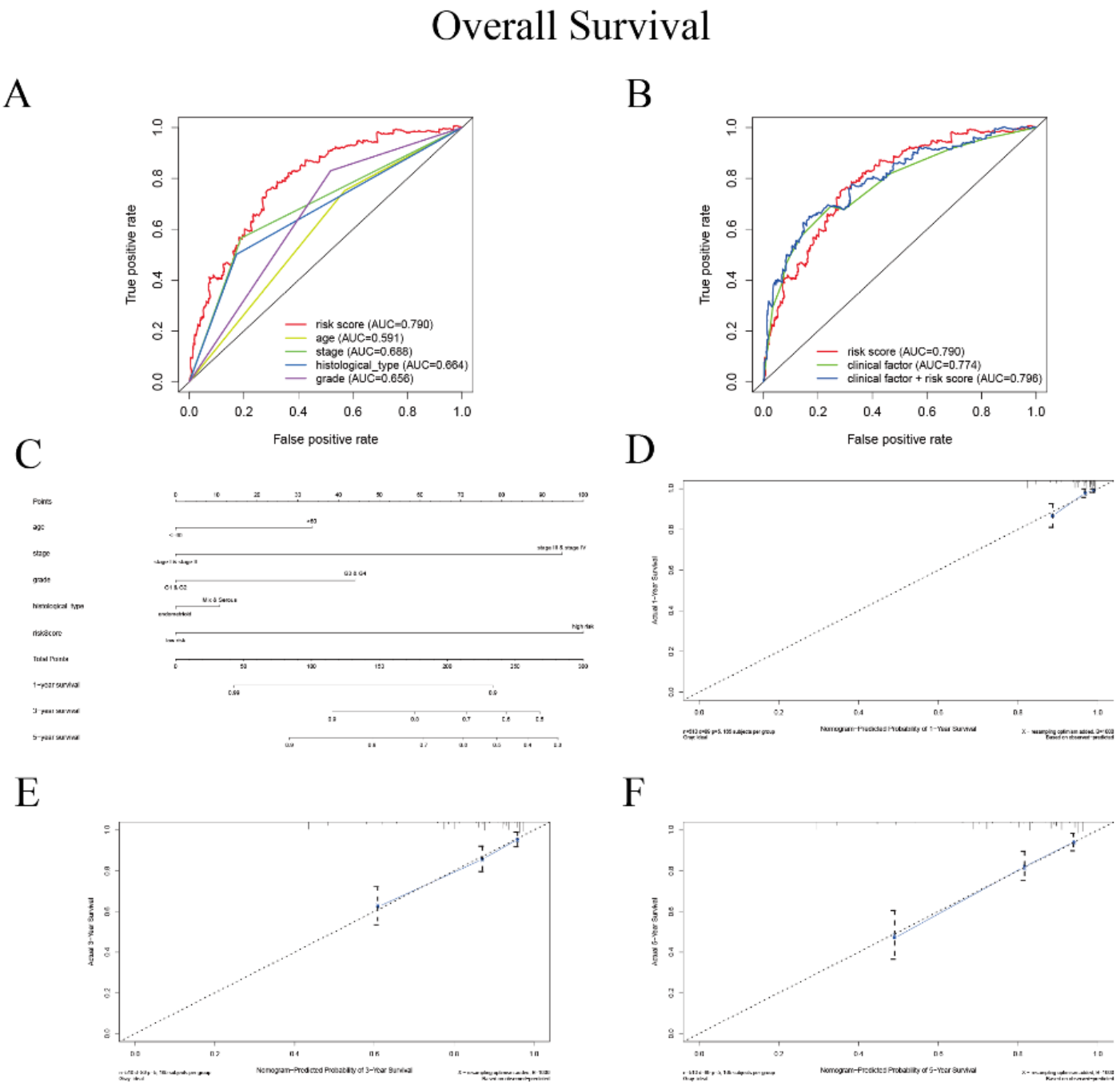

Figure 12

The nomogram to predict 1-, 3- or 5-year OS in EC based on OS model. (A) Time-dependent ROC analysis to evaluate the prognostic value of risk score and clinical factors (age, stage, grade and histological type). (B) Time-dependent ROC analysis based on risk score and merged clinical factors. (C) Nomogram for predicting OS of EC. There were five factors containing age, gender, stage, histological type and risk score in the nomogram. Each of them generates points according to the line drawn upward. And the total points of the seven components of an individual patient lie on "Total Points" axis which corresponds to the probability of 1-year, 3-year and 5-year OS rate plotted on the two axes below. The calibration plots for predicting patient 1-(D), 3-(E) or 5-(F) year OS.

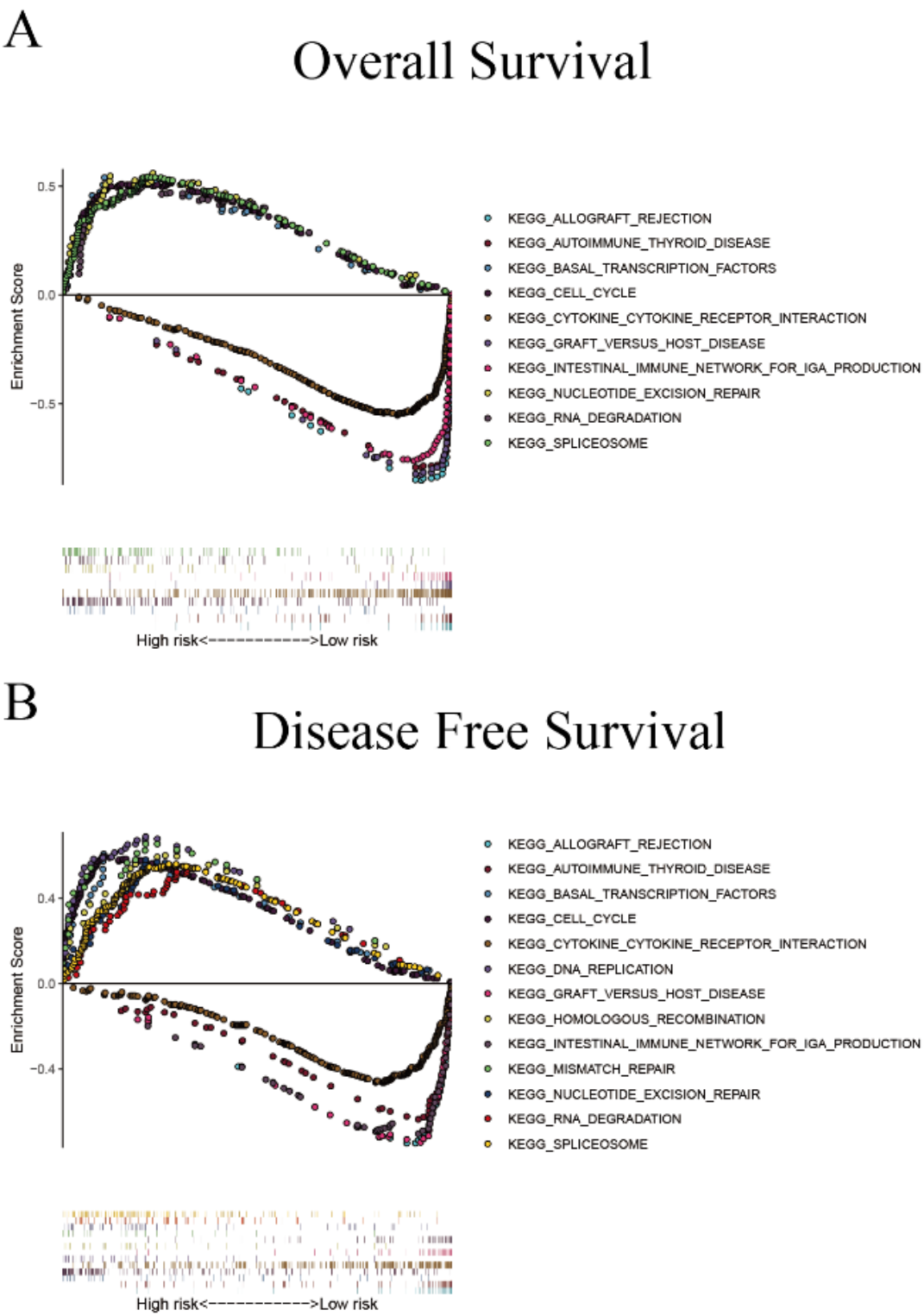

Figure 14

Gene set enrichment analysis showed the significantly enriched KEGG pathways in TCGA based on OS (A) and DFS(B).

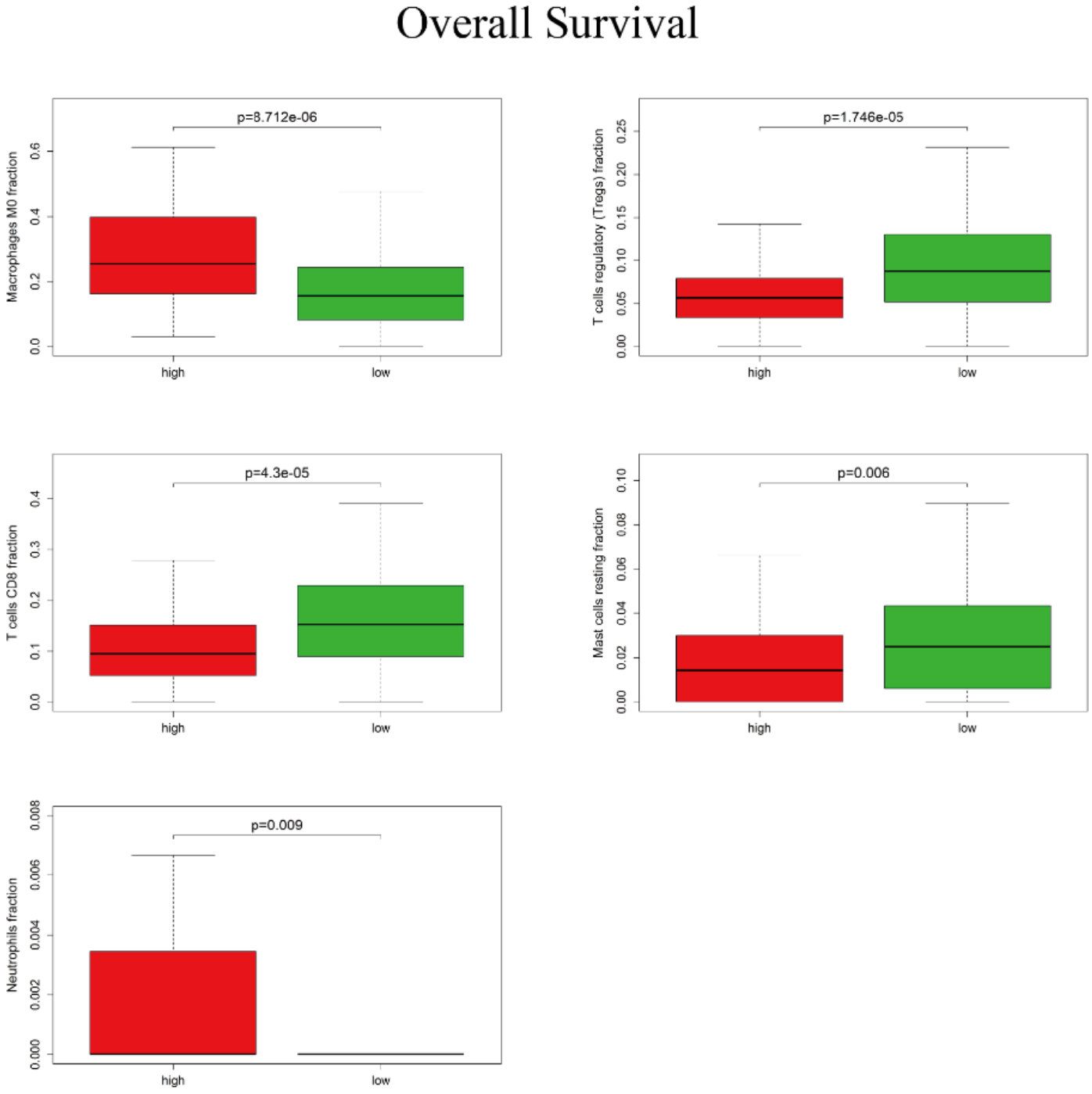

Figure 15

The relationship of OS immune-related prognosis risk score with the immune cells.

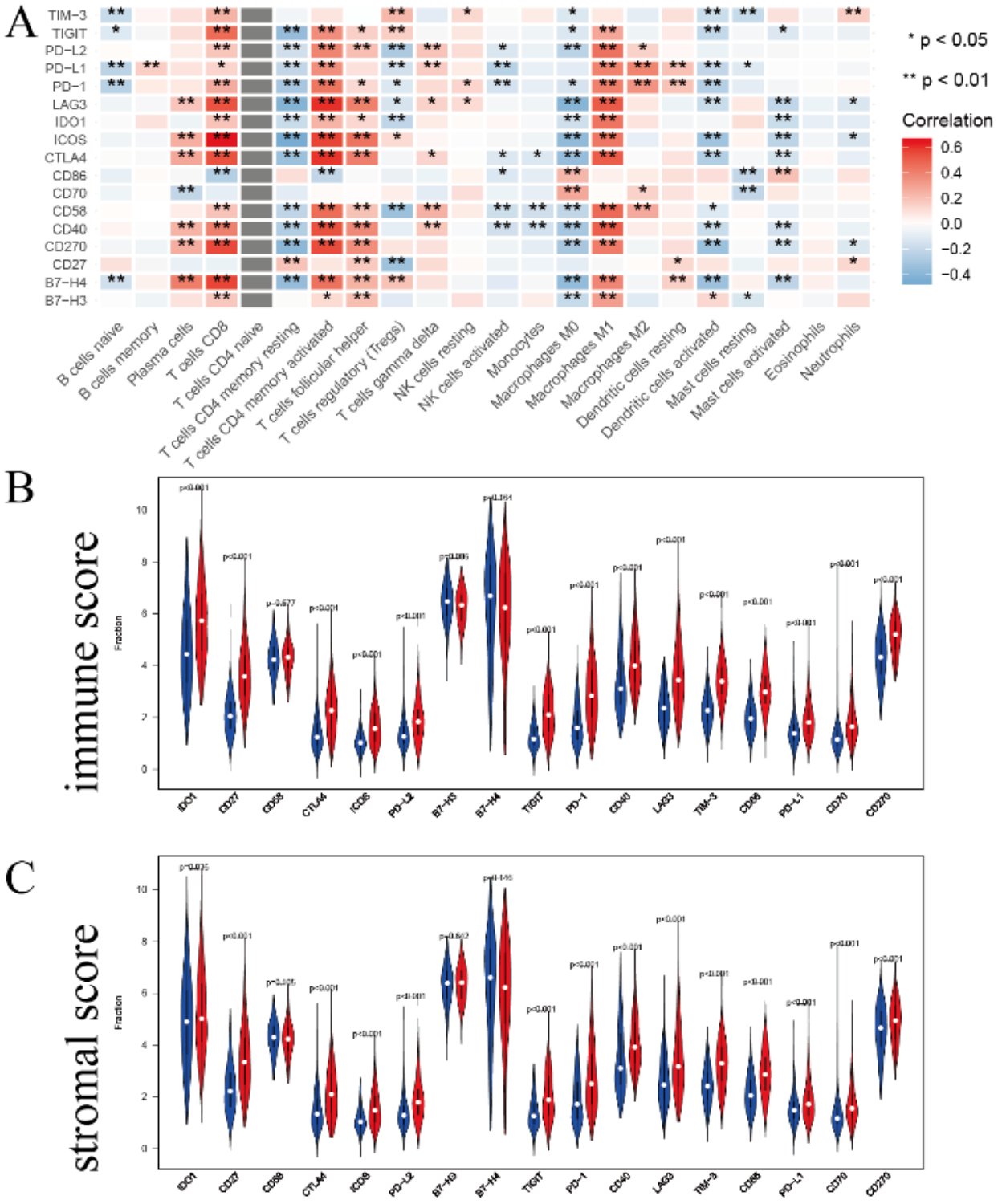

Figure 18

The relationship of immunomodulators with immune cells, immune and stromal score. (A) Correlations between immunomodulators and the distribution of immune cell infiltration. The distribution of immunomodulators in high-/low- immune (B) and stromal (C) score subgroup.

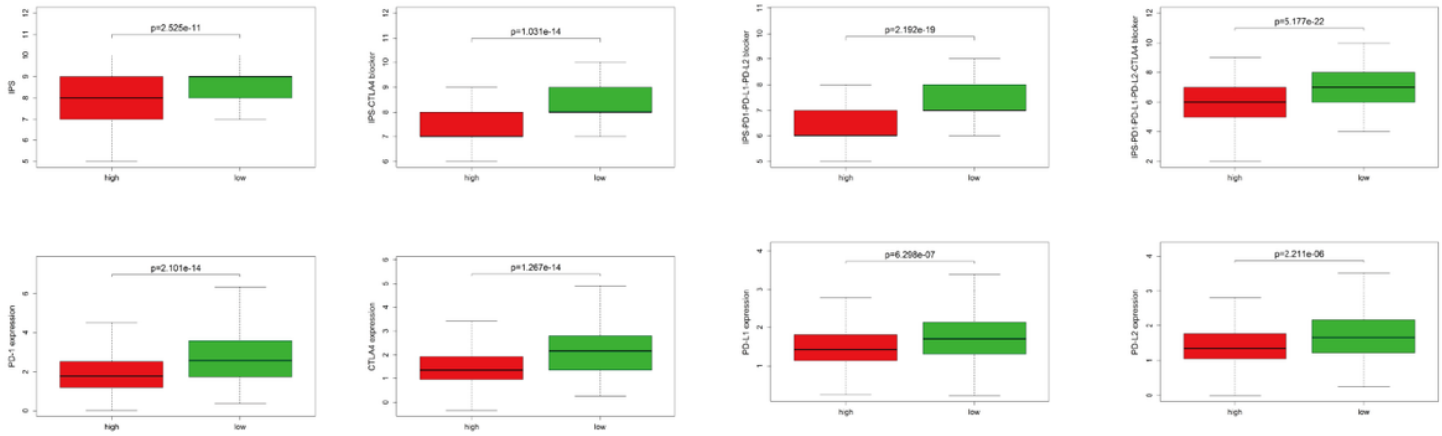

Figure 20

IPS and immunotherapy gene expression analysis. (A) The association between IPS and the risk score based on OS in EC patients. (B) The gene expression of PD1, CTLA4, PD-L1, and PD-L2 in low-risk and high- risk subgroups.

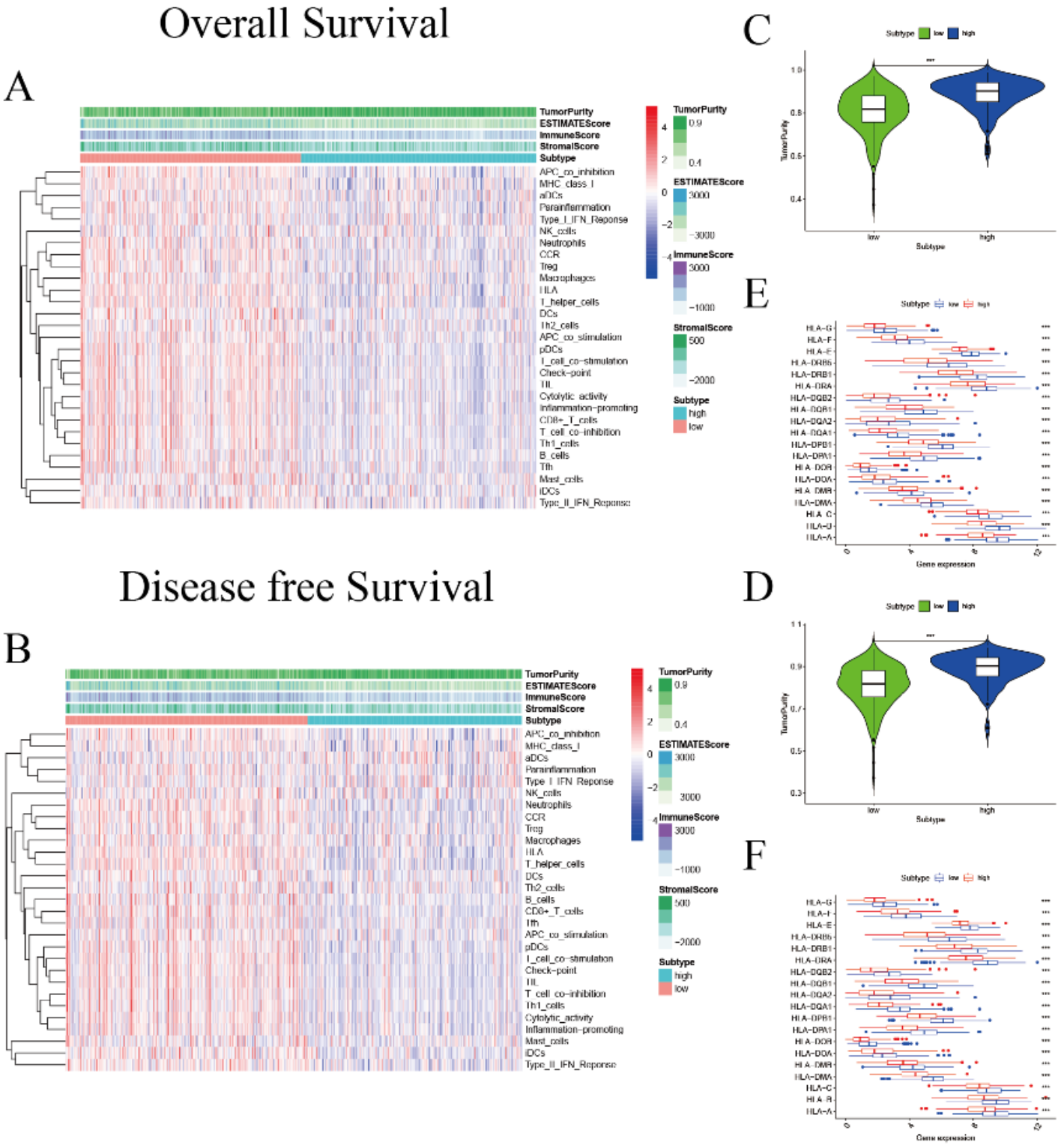

Figure 22

The low-risk and high-risk samples show different tumor purity. The immune status (A-B), tumor purity (C-D) and the distribution of HLA related genes (E-F) of the low-risk and high-risk samples in OS and DFS model.

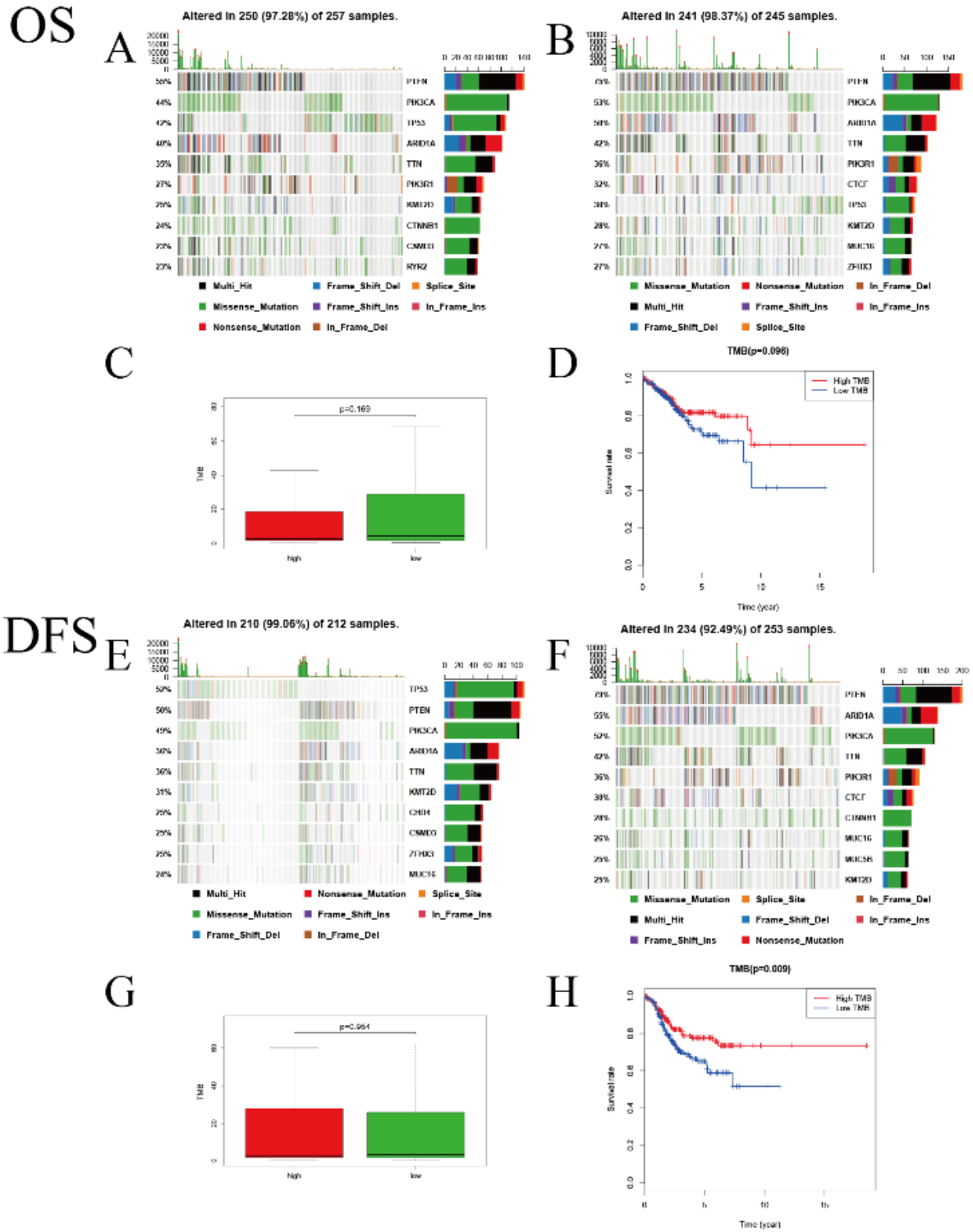

Figure 23

Immune cells infiltration was associated with genetic mutations. Based on OS model: Oncoplot displaying the somatic landscape of EC with high risk groups (A) and low risk group (B), the genes are sorted according to their mutation frequency. (C) The relationship between the risk signature and TMB. (D) The association of TMB and OS in TCGA. Based on DFS model oncoplot displaying the somatic landscape of EC with high risk groups (E) and low risk groups (F) , the genes are sorted according to their

mutation frequency. (G) The relationship between the risk signature and TMB. (H) The association of TMB and DFS in TCGA.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigure8.tif
- tableS1.docx
- SupplementaryFigure11.tif
- SupplementaryFigure6.tif
- SupplementaryFigure5.tif
- SupplementaryFigure6.tif
- SupplementaryFigure4.tif
- SupplementaryFigure2.tif
- SupplementaryFigure4.tif
- SupplementaryFigure1.tif
- SupplementaryFigure2.tif
- tableS2.docx
- SupplementaryFigure1.tif
- SupplementaryFigure8.tif
- tableS2.docx
- tableS1.docx
- SupplementaryFigure9.tif
- SupplementaryFigure3.tif
- SupplementaryFigure9.tif
- SupplementaryFigure3.tif
- SupplementaryFigure7.tif
- SupplementaryFigure10.tif
- SupplementaryFigure7.tif
- SupplementaryFigure10.tif
- SupplementaryFigure5.tif
- SupplementaryFigure11.tif