A Prognostic Model Based on Immune-Related Long Non-Coding RNAs for Patients With Epithelial Ovarian Cancer

Yao Peng
Anhui Medical University
Hui Wang
Lu'an Municipal People's Hospital
Qi Huang
Second Affiliated Hospital of Anhui Medical University
Jingjing Wu
Second Affiliated Hospital of Anhui Medical University
Mingjun Zhang (zhangmjayd@126.com)
Anhui Medical University https://orcid.org/0000-0001-5212-9322

Research

Keywords: Epithelial ovarian cancer, Immune, Long non-coding RNA, Prognostic signature, TCGA

Posted Date: August 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-752966/v1

Abstract

Background: Long non-coding RNA (lncRNA), as an important regulator of gene expression, can affect a variety of physiological processes. Recent studies have shown that immune-related lncRNA play an important role in the tumor immune microenvironment and may have potential application value in the treatment and prognosis prediction of tumor patients. Epithelial ovarian cancer (EOC) is characterized by high incidence and poor prognosis. However, there are few studies on immune-related lncRNAs in EOC. In this study, we focused on the immune-related lncRNAs associated with survival in EOC.

Methods: We downloaded mRNA data for EOC patients from The Cancer Genome Atlas (TCGA) database, and mRNA data for normal ovarian tissue from the Genotype-Tissue Expression (GTEx) database, and identified differential genes through differential Expression analysis. Immune-related lncRNAs were obtained through taking intersection and co-expression analysis of differential genes and immune-related genes from the Immunology Database and Analysis Portal (ImmPort). Samples in the TCGA EOC cohort were randomly divided into training set, validation set and combination set. In the training set, Cox regression analysis and LASSO regression were used to construct an immune-related lncRNA signature. Kaplan-Meier survival analysis, time-dependent ROC curve analysis, Cox regression analysis and principal component analysis were applied to verification in the training set, training set, validation set and combination set. Further studies of pathways and immune cell infiltration were conducted through Gene Set Enrichment Analysis (GESA) and the Timer data portal.

Results: An immune-related lncRNA signature was identified in EOC, which was composed of six immune-related lncRNAs (KRT7-AS, USP30-AS1, AC011445.1, AP005205.2, DNM3OS and AC027348.1). The signature divided patients into high-risk and low-risk groups. The overall survival of the high-risk group was lower than that of the low-risk group, and was verified to be robust in both the training set and the combination set. This signature was identified as an independent prognostic biomarker. The signature was confirmed to be an independent prognostic biomarker. Principal component analysis showed the different distribution patterns of high-risk and low-risk groups. This signature may be related to immune cell infiltration (mainly macrophages) and differential expression of immune checkpoint-related molecules (PD-1, PDL1, etc.).

Conclusions: we identified and established a prognostic signature of immune-related lncRNA in EOC, which is of great value in predicting the prognosis of clinical patients and may provide a new perspective for the immunological research and individualized treatment in EOC.

Introduction

Ovarian cancer (OC) is one of the most common gynecological malignancies in the world, with high morbidity and mortality [1]. Epithelial ovarian cancer (EOC) is the most common subtype, accounting for more than 90% of all ovarian cancers [2]. Despite the rapid development of therapeutic approaches such as surgery, chemotherapy, targeted therapy and immunotherapy, most patients are diagnosed in

advanced clinical stages with poor prognosis due to the lack of specific signs and symptoms at early stage and effective clinical screening methods [3, 4]. At present, the first-line therapy for newly diagnosed EOC is complete cytoreductive surgery followed by platinum-based chemotherapy [5]. The sensitivity of first-line therapy is better in patients with stage I-IIA lesion confined to the ovary. At stage IIB-IV, the cancer cells have metastasized to the peritoneum and show resistance to first-line treatment, with a very high recurrence rate [6, 7]. Therefore, in order to improve the survival rate of EOC patients, it is very necessary to find new prognostic biomarkers and improve the prediction of prognosis.

Long non-coding RNA (lncRNA) is a kind of non-coding RNA with a length of more than 200 nucleotides. The expression level of lncRNAs was relatively low in tissues, but they were widely distributed in various organs such as brain, lung, heart and ovary [8]. Despite the lack of protein coding function, the lncRNAs are involved in various types of gene regulation, including epigenetic, transcriptional or post-transcriptional regulation. These regulations are closely related to the occurrence, development and prognosis of tumors and other diseases, and play important physiological roles in cell proliferation, apoptosis, metastasis, invasion and migration [9–11]. Expression profiles of lncRNAs were different in different types of tumors, as well as in normal tissues, benign tumors and malignant tumors. These differential expression patterns may promote or inhibit tumor genesis and progression. For example, it has been found that HOTAIR lncRNA overexpression correlates with aggressive colorectal cancer, ovarian cancer, etc., while MEG3 lncRNA may play an inhibitory role in esophageal cancer [12, 13]. One of the important characteristics of malignant tumors is the ability to escape immune surveillance, such as avoiding recognition by down-regulating the expression of MHC-I molecules and forming an immunosuppressive tumor microenvironment (TME) to avoid killing [14]. Immune-related lncRNA (IR-lncR) play an important role in tumor microenvironment, and as important regulators of gene expression, they are involved in tumor genesis, development and prognosis [15]. Currently, immunotherapy is widely used in clinical practice which can regulate TME [16]. Although the benefits of immunotherapy are significant, there are still many patients with low sensitivity and high resistance to immunotherapy. Therefore, it is of great significance to develop some new and more sensitive prognostic biomarkers and anti-tumor targets.

In this study, we obtained high-throughput sequencing data of EOC from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/, March 2021) data portal, identified and used the differential expression of IR-lncRs to build the model to predict the prognosis of patients with EOC. And the model was verified. Finally, we evaluated the function of the predictive model for immune infiltration and demonstrated the expression of several popular immune checkpoint-related genes.

Materials And Methods

Data collection and identification of differentially expressed genes (DEGs)

We downloaded mRNA transcriptome data from 379 patients with EOC and clinicopathological information from 587 patients with EOC from the TCGA. At the same time, mRNA transcriptome data of 88 normal ovarian tissues from the Genotype-Tissue Expression (GTEx, https://www.gtexportal.org/home/index.html) database were included in the analysis. We normalized the mRNA data of 379 EOC patients and 88 normal ovarian tissues by fragment per million exon model. A list of 2483 immune-related genes (IRGs) was downloaded from the Immunology Database and Analysis Portal (ImmPort, https://immport.niaid.nih.gov) (Attachment 1). The above data are available to the public, so it is not subject to review by an ethics committee. The mRNA values of TCGA and GTEx portal were consolidated into a matrix file using Perl (https://www.perl.org/), and then differential expression analysis between serous ovarian cancer and normal tissue was performed using the "Limma package" in R (version 4.0.3) to identify the DEGs. The filter criteria were FDR less than 0.05 and |log2(FC)| greater than 1. According to The Gene Coding Plan (https://www.gencodegenes.org/), lncRNA profiles were extracted from the mRNA expression profiles of TCGA.

Mining differentially expressed IRGs and IR-lncRs

The differentially expressed IRGs (DE-IRGs) were extracted from the overlap of IRGs and DEGs. COR|>0.4 and P < 0.001 were set as the cut-off value, and R package "Limma" was used to analyze the co-expression of DE-IRGs and lncRNAs in TCGA, and the differentially expressed IR-lncRs were obtained for subsequent analysis.

Construction of IR-lncR signature based on differential expression

After excluding patients with incomplete prognostic information, 374 EOC patients with complete OS information were finally included. They were randomly divided into training set (n = 236) and validation set (n = 138) according to 5:3, and all patients were regarded as a combination set (n = 374). The IR-lncR prediction model was constructed with the data from the training set, and verified in the validation set and the combination set. First, univariate Cox regression analysis was performed on the training set to screen out IR-lncRs (P < 0.05) that were significantly correlated with overall survival (OS). Next, the Least Absolute Shrinkage and Selection Operator (LASSO) regression model was used to perform a multivariate Cox proportional hazard regression analysis. After 1000 times of cross-validation, the lambda value with the smallest error was finally determined, and the model was refitted with the best lambda value. Using multivariate Cox regression to establish a prognostic risk assessment model, we obtained the risk score calculation formula as follows: RiskScore = lncRNA1 expression value × β1 + lncRNA2 expression value × β2 +…… + lncRNAn expression value × βn. Where β represented the regression coefficient calculated by the multivariate Cox regression model. EOC patients in the training, validation, and combination set were divided into high-risk and low-risk groups according to the median risk score. Finally, R packages "Pheatmap", "SurvMiner", "Survival" and "Survivalroc" were used to evaluate the accuracy of the model in the training, validation, and combination set, respectively. And we drew risk heatmap, risk curve and survival state chart. Kaplan-Meier method was used to generate survival curves for the high-risk and low-

risk groups. The time-dependent receiver operating characteristic (ROC) curves were drawn and the area under the curve (AUC) was calculated at 3-year and 5-year.

Independent prognostic analysis and construction of the normogram

To determine the predictive effect of the constructed IR-lncR model on prognosis, we excluded patients who lacked detailed clinicopathological information, including age, histological grade, and FIGO stage. We used the R packages "Survival" and "Forestplot" to incorporate risk scores and clinical indicators into univariate and multivariate Cox regression analyses. We used the "Rms" package to construct a nomogram to predict 1-, 3-, and 5-year survival in patients with EOC in conjunction with risk scores and clinical indicators. The calibration curve of the nomogram was used to evaluate the accuracy of the prediction effect. We also used ROC curves to compare a nomogram containing only one independent prognostic factor with a nomogram containing all independent prognostic factors. In addition, R package "Vioplot" was used to visualize the expressions of each lncRNA in the signature between normal ovarian tissue and ovarian cancer tissue, and R package "corrplot" and Pearson correlation test were used to illustrate the interactions between each lncRNA.

Gene Set Enrichment Analysis (GSEA)

GSEA (http://software.broadinstitute.org/gsea/index.jsp) is used to understand the expression status of certain genes in specific functional gene sets. According to the prediction model, TCGA EOC samples were divided into high-risk and low-risk groups. C7.all.v7.3. Symbols. GMT from the Molecular Signature Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/index.jsp) was selected as the reference gene bank. GSEA was used to identify differences in biological function between the two groups.

Evaluation of immune status and immune cell infiltration based on predictive model

The R packages "Limma" and "Scatterplot3D" were used to perform principal component analysis (PCA) on the gene expression profiles of the TCGA EOC cohort, and the immune status and expression pattern of the subgroup samples were visualized. We downloaded the level of immune infiltration in EOC patients from the Tumor Immune Estimation Resource (Timer, https://cistrome.shinyapps.io/timer/) data portal to evaluate the correlation between IR-lncR and immune cell infiltration. At the same time, in order to clearly show the distribution of immune checkpoint-related genes expression in the prediction model, we used the R packages "Limma" and "Beeswarm" to draw the box map, and analyzed the differential expression of eight immune checkpoint related genes between low-risk and high-risk groups.

Results

This research was carried out according to the process shown in Fig. 1.

Identification of DE-IRGs and IR-lncRs in EOC

We identified a total of 7255 DEGs between tumors and normal tissues, including 3790 up-regulated genes and 3465 down-regulated genes (Fig. 2A and 2C). By intersecting these DEGs with the 2,483 IRGs from ImmPort, we found 339 up-regulated DE-IRGs and 157 down-regulated DE-IRGs in EOC (Fig. 2B and 2D). Through the common expression analysis, we identified 421 differentially expressed IR-lncRs (| COR | > 0.4, p < 0.001) (Attachment 2).

Construction of IR-lncR signature in TCGA EOC cohort

A total of 374 EOC patients with complete OS information were included from the TCGA for the follow-up study. In order to increase the credibility of the study, we divided the whole data set into a training set and a verification set, and defined the whole data set as a combination set. The expression profiles of 421 IR-lncRs in the training set was used to construct a prognostic prediction model. Univariate Cox regression analysis was performed on the expression profiles of lncRNAs in the training set, and 32 IR-lncRs were significantly correlated with OS (P < 0.05) (Table 1). LASSO regression was performed on these IR-lncRs to prevent over-fitting of the model, and the prediction accuracy was estimated through 1000 cross validations (Fig. 3A-B). Next, we identified six key IR-lncRs using multivariate Cox regression analysis (Fig. 3C). These six lncRNAs were used in the prognostic model construction, and they were respectively KRT7-AS, USP30-AS1, AC011445.1, AP005205.2, DNM3OS and AC027348.1, and the corresponding coefficients were also given (Table 2). The final risk score calculation formula was as follows: Risk score = expression value of KRT7-AS * 0.2079 + expression value of USP30-AS1 * (-0.3862) + expression value of AC011445.1 * 0.4593 + expression value of AP005205.2 * (-0.4020) + expression value of DNM3OS * (0.3120) + expression value of AC027348.1 * (-0.8224).

Table 1
Univariate Cox analysis of immune-related lncRNAs.

id	HR	HR.95L	HR.95H	pvalue
AC040169.1	0.62648	0.480216	0.817292	0.000566
AC011445.1	1.472847	1.155252	1.877753	0.001781
AC027348.1	0.605601	0.43929	0.834877	0.002201
AP005205.2	0.63808	0.46868	0.868708	0.004317
AC010531.6	0.656519	0.489907	0.879795	0.004841
AC083880.1	0.630212	0.456412	0.870194	0.005039
UBXN10-AS1	0.71376	0.562469	0.905743	0.005528
HCG14	0.655622	0.483959	0.888175	0.006419
AC091153.3	0.663274	0.485976	0.905257	0.009676
AC020916.2	1.530921	1.104282	2.122391	0.010615
USP30-AS1	0.737146	0.576721	0.942196	0.014873
AC135050.6	0.685481	0.503844	0.932598	0.016207
CDC37L1-DT	0.702602	0.523147	0.943614	0.018993
AC073046.1	1.358341	1.050206	1.756884	0.01964
LINC02004	0.688872	0.500912	0.94736	0.021871
KMT2E-AS1	0.735848	0.564351	0.959461	0.023473
KRT7-AS	1.238102	1.025378	1.494957	0.026384
ATP2A1-AS1	0.742762	0.569485	0.968762	0.028227
AL137802.2	0.680427	0.481615	0.961309	0.02898
AL035701.1	0.687185	0.48917	0.965355	0.030519
TGFB2-AS1	0.793979	0.643097	0.980262	0.031924
DLG3-AS1	0.722241	0.533965	0.976904	0.034721
AL451165.2	0.774204	0.608127	0.985636	0.037764
DNM3OS	1.341397	1.01427	1.77403	0.039465
AP001453.2	0.785556	0.623085	0.990391	0.041188
AC010326.3	0.71923	0.523362	0.988402	0.042165
PCAT6	0.823436	0.681595	0.994794	0.044001

id	HR	HR.95L	HR.95H	pvalue
AL391069.2	0.787554	0.624064	0.993875	0.04425
U62317.1	0.710871	0.50925	0.992319	0.044934
AC133552.5	0.792207	0.630022	0.996142	0.046257
AC103769.1	0.757677	0.574379	0.99947	0.049563

Table 2
Multiple Cox analysis of the EOC-specific immune-related lncRNAs.

id	coef	HR	HR.95L	HR.95H	pvalue
KRT7-AS	0.20788	1.231066	1.006212	1.506167	0.043368
USP30-AS1	-0.38621	0.679631	0.51683	0.893714	0.005705
AC011445.1	0.459328	1.58301	1.231711	2.034504	0.000333
AP005205.2	-0.40195	0.669012	0.485082	0.922683	0.014264
DNM3OS	0.311961	1.366101	1.034135	1.804631	0.028073
AC027348.1	-0.82237	0.439388	0.303846	0.635395	0.000012

Verification of 6-lncRNA signature for survival prediction

We constructed a good 6-lncRNA signature for survival prediction and verified it. The training set risk scores were calculated according to the risk score calculation formula and the lncRNA expression profiles of the TCGA EOC cohort. Then, using the median risk score as a cut-off point, patients were divided into high-risk and low-risk groups. In the training set, the mortality of patients increased with the increase in risk score (Fig. 4A). The survival status of the low-risk group was better than that of the high-risk group, and the OS of the low-risk group was significantly higher than that of the high-risk group, and the heat map showed the expression of 6 lncRNAs in the training set (Fig. 4B). The AUC value of 3-year and 5-year OS were 0.715 and 0.798, respectively (Fig. 4C). In order to evaluate the accuracy of the prediction model, the prediction model was validated in the validation set and the combination set. Patients in the validation and combination sets were also divided into low-risk and high-risk groups based on the median risk score of the training set. The results showed that the mortality of patients in both validation and combination sets increased with the increase of risk score (Figs. 4D and 4G). The survival status of the low-risk group was better than that of the high-risk group, and the OS of the low-risk group was higher than that of the high-risk group, and the expression of 6 lncRNAs in validation and combination sets was also shown by heat map (Figs. 4E and 4H). In the validation set, the AUC value of 3-year and 5-year OS were 0.675 and 0.614, respectively (Fig. 4F). In the combination set, the AUC value of 3-year and 5-year OS were 0.682 and 0.693, respectively (Fig. 4I).

Predictive model as an independent prognostic factor evaluation

We included clinicopathological features such as age, histological grade, and FIGO stage, as well as risk score in the analysis. Univariate and multivariate Cox regression analyses were used to determine whether 6-lncRNA signature was an independent prognostic factor. Univariate Cox analysis results showed that age and risk score were independent prognostic factors for EOC patients in the training and combination set, while only risk score was independent prognostic factor in the validation set. The results of multivariate Cox analysis also proved that the predictive model was reliable independent prognostic indicator in the training, validation and combination set (Table 3).

Table 3
Univariate and multivariate Cox analysis of the clinical features for EOC patients in each set.

Variables	Univariate analysis		P	Multivariate analysis		P
	HR	95%CI		HR	95%CI	
TCGA training set						
Age (≤ 55 vs > 55)	1.52	1.03–2.25	0.04	1.39	0.93–2.06	1.39
Grade (G1 vs G2 vs G3 vs G4)	1.24	0.78–1.97	0.40	1.04	0.65–1.68	1.04
Stage (Stage I vs Stage II vs Stage III vs Stage IV)	1.46	0.97–2.21	0.07	1.44	0.93–2.23	1.44
Riskscore (high/low)	1.70	1.49–1.93	0.00	1.68	1.47–1.92	1.68
TCGA verification set						
Age (≤ 55 vs > 55)	1.14	0.69–1.88	0.61	1.16	0.70–1.93	0.56
Grade (G1 vs G2 vs G3 vs G4)	2.09	0.84–5.20	0.11	2.08	0.83–5.22	0.12
Stage (Stage I vs Stage II vs Stage III vs Stage IV)	1.21	0.76–1.92	0.43	1.13	0.68–1.87	0.63
Riskscore (high/low)	1.47	1.14–1.89	0.00	1.47	1.15–1.88	0.00
TCGA combination set						
Age (≤ 55 vs > 55)	1.36	1.00-1.85	0.05	1.32	0.97–1.97	0.08
Grade (G1 vs G2 vs G3 vs G4)	1.39	0.92–2.09	0.12	1.25	0.82–1.88	0.30
Stage (Stage I vs Stage II vs Stage III vs Stage IV)	1.33	0.98–1.80	0.07	1.29	0.94–1.77	0.11
Riskscore (high/low)	1.63	1.45–1.83	0.00	1.61	1.43–1.81	0.00

Construction and Verification of nomogram

We developed a nomogram for predicting survival risk in EOC patients based on the entire TCGA EOC cohort. Prognostic indicators such as age, grade, stage, and risk score were included in the Nomogram (Fig. 5A). The calibration chart showed a high consistency between the prediction and actual observation of the 3-year and 5-year survival rates for EOC patients (Fig. 5B-C). In addition, the 3-year and 5-year OS of

the nomogram were 0.688 and 0.711, respectively (Fig. 5D-E), which was significantly more valuable than a single clinicopathological index.

Evaluation of immune infiltrate function in low-risk and high-risk populations

PCA was used to visualize patient distribution based on genome-wide expression set, IRGs expression set, IR-lncRs expression set, and 6-lncRNA signature. According to the genome-wide and the IRGs expression set, the PCA showed no significant separation between the groups (Fig. 6A-B). Based on the IR-lncRs expression set, patients in the low-risk and high-risk groups tended to be divided into two groups (Fig. 6C). However, with the 6-lncRNA signature, patients in the low-risk group and the high-risk group clearly showed different distribution directions (Fig. 6D). We also analyzed the relationship between 6-lncRNA signature and immune cell infiltration, and found a positive correlation between macrophages and risk score (Fig. 6E). GSEA was used to further explore the biological function of 6-lncRNA signature. GSEA analysis showed that 6-lncRNA signature in the high-risk group was highly expressed mainly in cell adhesion, mitogen-activated protein kinase (MAPK) signaling pathway, various cancer-related pathways, and ErbB receptor-related signaling pathway (Fig. 7). In the low-risk group, more were associated with spliceosomes and proteasomes, which are involved in the cell cycle, regulate apoptosis and enhance the immune response.

Expression of 6-lncRNA and immune checkpoint related genes

We compared the expression levels of six lncRNAs (KRT7-AS, USP30-AS1, AC011445.1, AP005205.2, DNM3OS and AC027348.1) in normal ovarian tissues and ovarian cancer tissues (Fig. 8A). In order to further understand the interaction between the six lncRNAs, we analyzed their expression correlations (Fig. 8B). At the same time, eight immune checkpoint related genes, including PDCD1 (PD-1), CD274(PD-L1), PDCD1LG1, PDCD1LG2, CTLA-4, HAVCR2, LAG-3 and CD96, were selected to analyze their expression differences between high-risk and low-risk groups. We found that there were four up-regulated immune checkpoint genes in the high-risk group with 6-lncRNA signature, namely CD274, PDCD1, LAG-3 and PDCD1LG1 (Fig. 8C-J).

Discussion

At present, EOC has insidious onset, early metastasis and high recurrence rate. Platinum-based chemotherapy occupies an important position in EOC drug treatment, and it is very common for EOC patients with initial platinum-sensitive to develop cancer recurrence and platinum resistance [17]. The use of targeted therapies has led to significant changes in the treatment model of EOC, extending the survival of EOC patients and bringing new hopes. Targeted therapy for EOC refers to the use of small molecule compounds or monoclonal antibodies to specifically interfere with the molecular targets of tumor cells to achieve the purpose of anti-tumor, mainly including angiogenesis inhibitors, poly ADP-ribose polymerase(PARP) inhibitors and immune checkpoint inhibitors (ICBs) [18].The efficacy of targeted

therapy is closely related to the mutation status of relevant genes, and the occurrence of gene mutations is probabilistic, so the efficacy of targeted therapy is limited to some extent [19]. Studies have shown that lncRNA plays an important role in the occurrence, development, metastasis, invasion and other biological behaviors of EOC, and its expression is dysregulated in ovarian cancer tissues, showing potential as emerging tumor markers [20, 21]. Emerging evidences suggest that lncRNAs play a regulatory role in controlling cancer immunity [10]. For example, NKILA lncRNA promotes tumor immune escape by sensitizing T cells to activation-induced cell death [22].

At present, IR-lncR prognostic signatures have been reported in lung cancer, cervical cancer, breast cancer and other cancers [23–25]. However, previous studies on the prognostic biomarkers of EOC mainly focused on mRNAs and microRNAs, and relatively few studies focused on the prognostic value of IR-lncRs in EOC [26, 27]. Therefore, in order to better evaluate the immune status and prognosis of EOC patients, we focused on IR-lncRs this time and developed a risk scoring model in EOC patients based on IR-lncRs for the first time.

In this study, we integrated the gene profiles of TCGA and GTEx, combined with IRGs from ImmPort, and obtained 421 differentially expressed IR-lncRs through co-expression analysis. Univariate and multivariate Cox regression analysis were performed on the differentially expressed IR-lncRs of 236 EOC patients in the training set. Finally, a 6-lncRNA signature (KRT7-AS, USP30-AS1, AC011445.1, AP005205.2, DNM3OS, and AC027348.1) was determined, which could classify EOC patients into high-risk and low-risk groups, with significant difference in OS between the two groups ($P < 0.001$). At the same time, we verified the signature in the validation set and the combination set, and the results showed that 6-lncRNA signature had good predictive ability. Independent prognostic analysis confirmed that the 6-lncRNA signature was superior to other clinicopathologic features in predicting survival.

Among the 6-lncRNA signature, studies have found that KRT7-AS can promote tumor cell proliferation and migration, and KRT7 may promote epithelial-to-mesenchymal transition (EMT) of OC through TGF-β/Smad2/3 signaling pathway [28–30]. In our study, USP30-AS1 was regarded as a tumor suppressor factor. However, there are not many relevant studies on USP30-AS1 in EOC. In the future, it may be necessary to combine in vivo and in vitro experiments to reveal the relationship between USP30-AS1 and EOC. Mitra et al. knocked out the DNM3OS gene in OC cells, performed RNA sequencing and pathway analysis of differentially expressed genes, and found that several EMT-linked pathways were affected, the expression of EMT-related proteins in OC cells was reduced, and migration and invasion were inhibited [2].

Immune cell infiltration in TME plays a key role in tumorigenesis and progression and affects the clinical prognosis of cancer patients [31]. Macrophages, as an important component of TME, have demonstrated the ability to inhibit T cells recruitment and function as well as other aspects of tumor immunity, and are associated with adverse disease outcomes [32]. In this study, there was a positive correlation between macrophages and risk score, which we speculated might be related to the poor prognosis of patients in the high-risk group at the level of tumor immunity. In addition, ICBs, as an emerging anticancer target, have been approved for a variety of malignancies, which may change the treatment model of EOC in the

future. In the 6-lncRNA signature constructed by us, CD274, PDCD1, LAG-3 and PDCD1LG1 genes in the high-risk group were highly expressed, which may have implications for the selection of immunotherapy targets and populations in EOC. However, there are some limitations in our study. First, our signature was only validated internally, not further validated with other external data. Secondly, it is necessary to further study the functions and mechanisms of these six IR-lncRNAs in combination with basic experiments. In addition, a larger sample size is needed to verify the accuracy of 6-lncRNA signature in the future.

Conclusions

In summary, an immune-related lncRNA prognostic evaluation model for EOC was established, which consisted of six lncRNAs (KRT7-AS, USP30-AS1, AC011445.1, AP005205.2, DNM3OS and AC027348.1). The results showed that the model is reliable in predicting the prognosis of clinical patients. We expect that this model will provide ideas for the development of new biomarkers and guide the individualized treatment of patients with EOC.

Abbreviations

AUC
Area under the curve; DEGs:Differentially expressed genes; DE-IRGs:Differentially expressed IRGs; EOC:Epithelial ovarian cancer; GESA:Gene Set Enrichment Analysis; GTEx:Genotype-Tissue Expression; ICBs:Immune checkpoint inhibitors; ImmPort:Immunology Database and Analysis Portal; IRGs:Immune-related genes; IR-lncR:Immune-related lncRNA; LASSO:Least Absolute Shrinkage and Selection Operator; LncRNA:Long non-coding RNA; OC:Ovarian cancer; OS:Overall survival; PARP:Poly ADP-ribose polymerase; PCA:Principal component analysis; ROC:Receiver operating characteristic; TCGA:The Cancer Genome Atlas; Timer:Tumor Immune Estimation Resource; TME:Tumor microenvironment.

Declarations

Acknowledgment

The authors would like to thank the TCGA, GTEx, ImmPort and Timer database for open access to the database.

Authors' Contributions

YP participated in the design of the work, the statistical analysis of the data and drafted this article. HW, QH and JJW collected data and designed charts. MJZ designed the study and analyzed for the work. All authors read and approved the final article.

Funding Information

Not applicable.

Availability of Data and Materials

The datasets generated and analyzed during this study are available in the TCGA, GTEx, ImmPort and Timer database.

Ethics approval and consent to participate

Not applicable.

Consent for Publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Torre LA, Trabert B, DeSantis CE, Miller KD, Samimi G, Runowicz CD, et al. Ovarian cancer statistics, 2018. CA Cancer J Clin. 2018;68(4):284-96. https://doi.org/10.3322/caac.21456
2. Mitra R, Chen X, Greenawalt EJ, Maulik U, Jiang W, Zhao Z, et al. Decoding critical long non-coding RNA in ovarian cancer epithelial-to-mesenchymal transition. Nat Commun. 2017;8(1):1604.
3. De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, et al. Cancer survival in Europe 1999–2007 by country and age: results of EUROCARE-5—a population-based study. The Lancet Oncology. 2014;15(1):23-34. https://doi.org/10.1016/S1470-2045(13)70546-1
4. Lheureux S, Braunstein M, Oza AM. Epithelial ovarian cancer: Evolution of management in the era of precision medicine. CA Cancer J Clin. 2019;69(4):280-304.
5. Ledermann JA, Raja FA, Fotopoulou C, Gonzalez-Martin A, Colombo N, Sessa C, et al. Newly diagnosed and relapsed epithelial ovarian carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2013;24 Suppl 6:vi24-32.
6. Corrado G, Salutari V, Palluzzi E, Distefano MG, Scambia G, Ferrandina G. Optimizing treatment in recurrent epithelial ovarian cancer. Expert Rev Anticancer Ther. 2017;17(12):1147-58.
7. Armstrong DK, Alvarez RD, Bakkum-Gamez JN, Barroilhet L, Behbakht K, Berchuck A, et al. Ovarian Cancer, Version 2.2020, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2021;19(2):191-226.
8. Dinger ME, Pang KC, Mercer TR, Mattick JS. Differentiating protein-coding and noncoding RNA: challenges and ambiguities. PLoS Comput Biol. 2008;4(11):e1000176.
9. Beermann J, Piccoli MT, Viereck J, Thum T. Non-coding RNAs in Development and Disease: Background, Mechanisms, and Therapeutic Approaches. Physiol Rev. 2016;96(4):1297-325.

10. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol. 2021;22(2):96-118.
11. Zhao J, Du P, Cui P, Qin Y, Hu C, Wu J, et al. LncRNA PVT1 promotes angiogenesis via activating the STAT3/VEGFA axis in gastric cancer. Oncogene. 2018;37(30):4094-109.
12. Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res. 2011;71(20):6320-6.
13. Xu QR, Tang J, Liao HY, Yu BT, He XY, Zheng YZ, et al. Long non-coding RNA MEG3 mediates the miR-149-3p/FOXP3 axis by reducing p53 ubiquitination to exert a suppressive effect on regulatory T cell differentiation and immune escape in esophageal cancer. J Transl Med. 2021;19(1):264.
14. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.
15. Chen D, Lu T, Tan J, Li H, Wang Q, Wei L. Long Non-coding RNAs as Communicators and Mediators Between the Tumor Microenvironment and Cancer Cells. Front Oncol. 2019;9:739.
16. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017;541(7637):321-30.
17. Kuroki L, Guntupalli SR. Treatment of epithelial ovarian cancer. BMJ. 2020;371:m3773.
18. Jiang Y, Wang C, Zhou S. Targeting tumor microenvironment in ovarian cancer: Premise and promise. Biochim Biophys Acta Rev Cancer. 2020;1873(2):188361.
19. Matei D, Nephew KP. Epigenetic Attire in Ovarian Cancer: The Emperor's New Clothes. Cancer Res. 2020;80(18):3775-85.
20. Liang H, Zhao X, Wang C, Sun J, Chen Y, Wang G, et al. Systematic analyses reveal long non-coding RNA (PTAF)-mediated promotion of EMT and invasion-metastasis in serous ovarian cancer. Mol Cancer. 2018;17(1):96.
21. Wu DD, Chen X, Sun KX, Wang LL, Chen S, Zhao Y. Role of the lncRNA ABHD11-AS1 in the tumorigenesis and progression of epithelial ovarian cancer through targeted regulation of RhoC. Mol Cancer. 2017;16(1):138.
22. Huang D, Chen J, Yang L, Ouyang Q, Li J, Lao L, et al. NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. Nat Immunol. 2018;19(10):1112-25.
23. Li JP, Li R, Liu X, Huo C, Liu TT, Yao J, et al. A Seven Immune-Related lncRNAs Model to Increase the Predicted Value of Lung Adenocarcinoma. Front Oncol. 2020;10:560779.
24. Chen P, Gao Y, Ouyang S, Wei L, Zhou M, You H, et al. A Prognostic Model Based on Immune-Related Long Non-Coding RNAs for Patients With Cervical Cancer. Front Pharmacol. 2020;11:585255.
25. Ma W, Zhao F, Yu X, Guan S, Suo H, Tao Z, et al. Immune-related lncRNAs as predictors of survival in breast cancer: a prognostic signature. J Transl Med. 2020;18(1):442.
26. Millstein J, Budden T, Goode EL, Anglesio MS, Talhouk A, Intermaggio MP, et al. Prognostic gene expression signature for high-grade serous ovarian cancer. Ann Oncol. 2020;31(9):1240-50.

27. Yokoi A, Matsuzaki J, Yamamoto Y, Yoneoka Y, Takahashi K, Shimizu H, et al. Integrated extracellular microRNA profiling for ovarian cancer screening. Nat Commun. 2018;9(1):4319.
28. Huang B, Song JH, Cheng Y, Abraham JM, Ibrahim S, Sun Z, et al. Long non-coding antisense RNA KRT7-AS is activated in gastric cancers and supports cancer cell progression by increasing KRT7 expression. Oncogene. 2016;35(37):4927-36.
29. Chen F, Chen Z, Guan T, Zhou Y, Ge L, Zhang H, et al. N(6) -Methyladenosine Regulates mRNA Stability and Translation Efficiency of KRT7 to Promote Breast Cancer Lung Metastasis. Cancer Res. 2021;81(11):2847-60.
30. An Q, Liu T, Wang MY, Yang YJ, Zhang ZD, Liu ZJ, et al. KRT7 promotes epithelialmesenchymal transition in ovarian cancer via the TGFbeta/Smad2/3 signaling pathway. Oncol Rep. 2021;45(2):481-92.
31. Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. Nat Immunol. 2013;14(10):1014-22.
32. DeNardo DG, Ruffell B. Macrophages as regulators of tumour immunity and immunotherapy. Nat Rev Immunol. 2019;19(6):369-82.

Figures

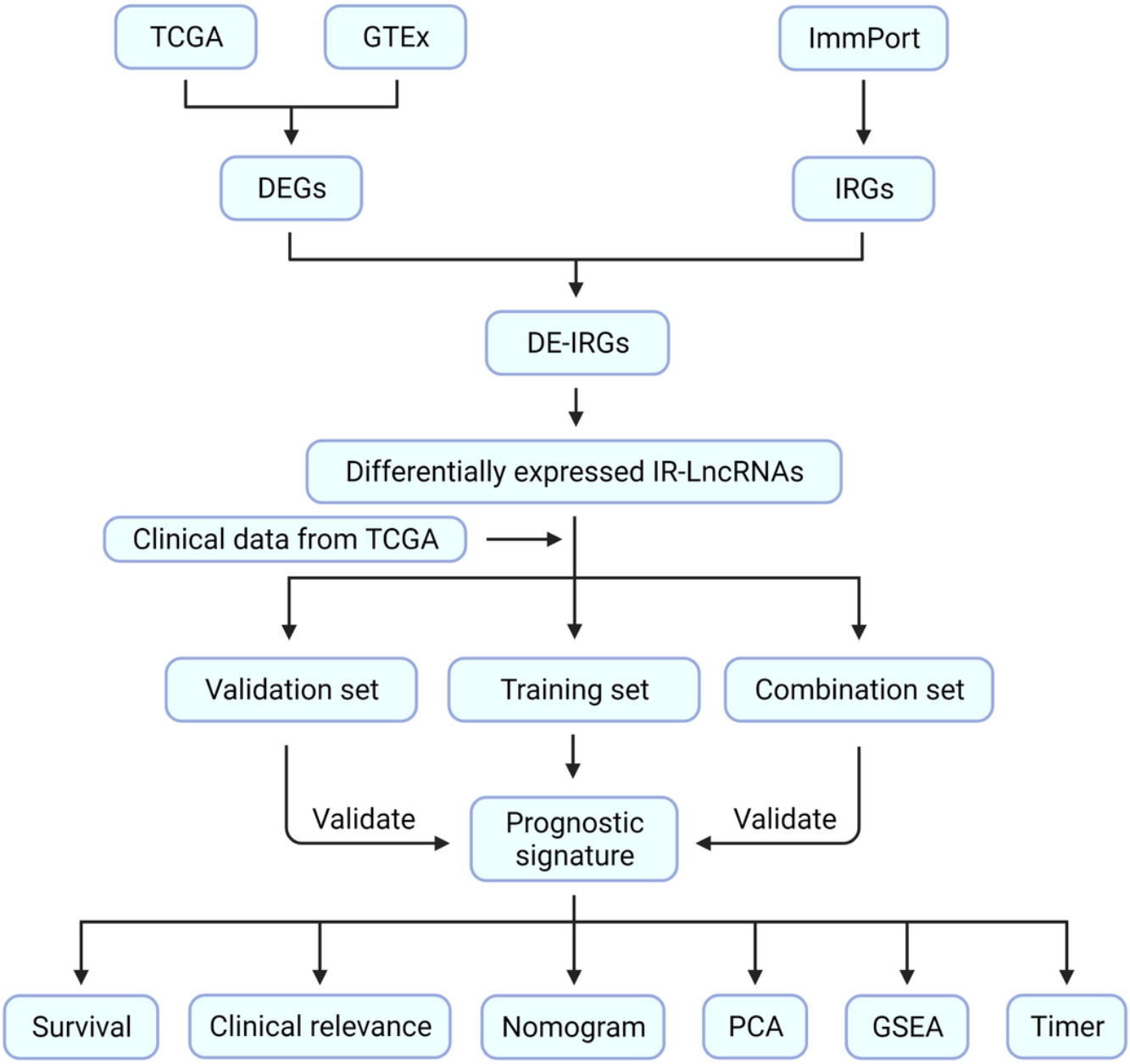

Figure 1

The flow diagram of the whole study.

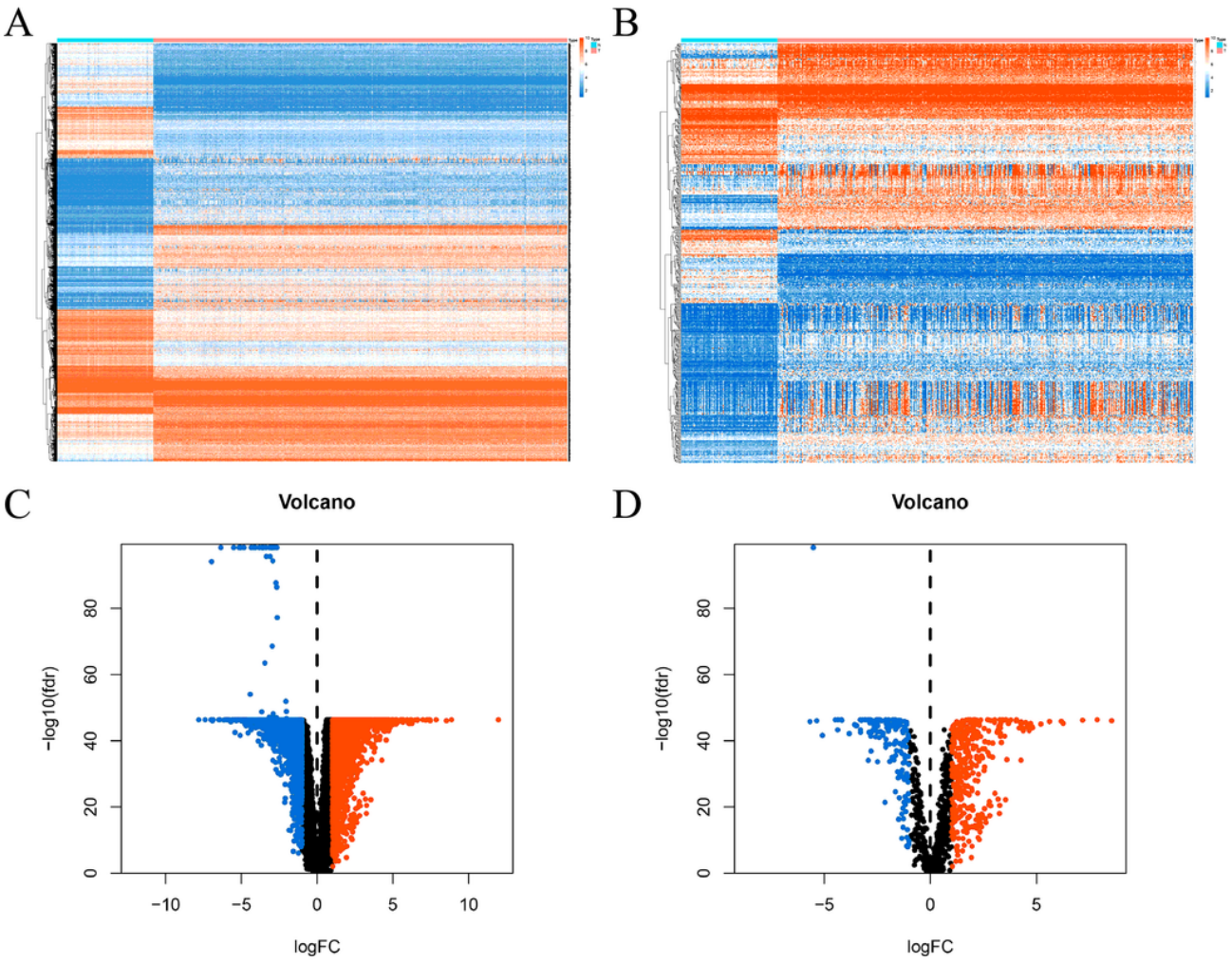

Figure 2

DE-IRGs were identified from 379 cases of ovarian cancer and 88 cases of normal ovarian tissue. (A, C) Heat map and volcano of DEGs in EOC. (B, D) Heat map and volcano of DE-IRGs in EOC.

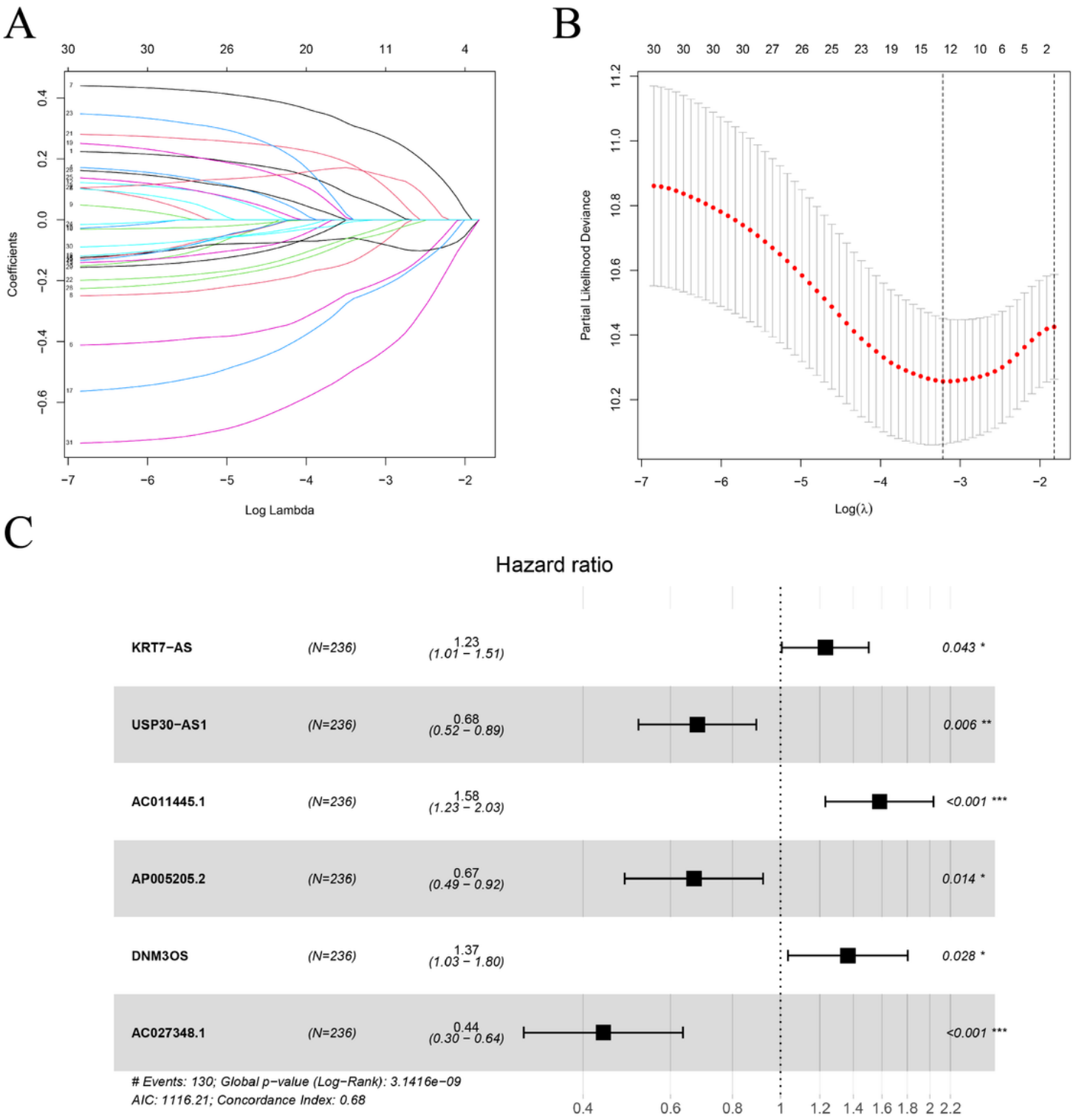

Figure 3

Construction of immune-related lncRNAs prognostic model based on the training set. (A) LASSO coefficient profiles of the 32 candidates in the training set. (B) A plot of thousand-fold cross-validation error rates. Selection of the optimal parameter (lambda) in the LASSO model. (C) Forest plot of six candidate immune-related LncRNAs associated with the survival of EOC were screened by multivariate Cox regression analysis.

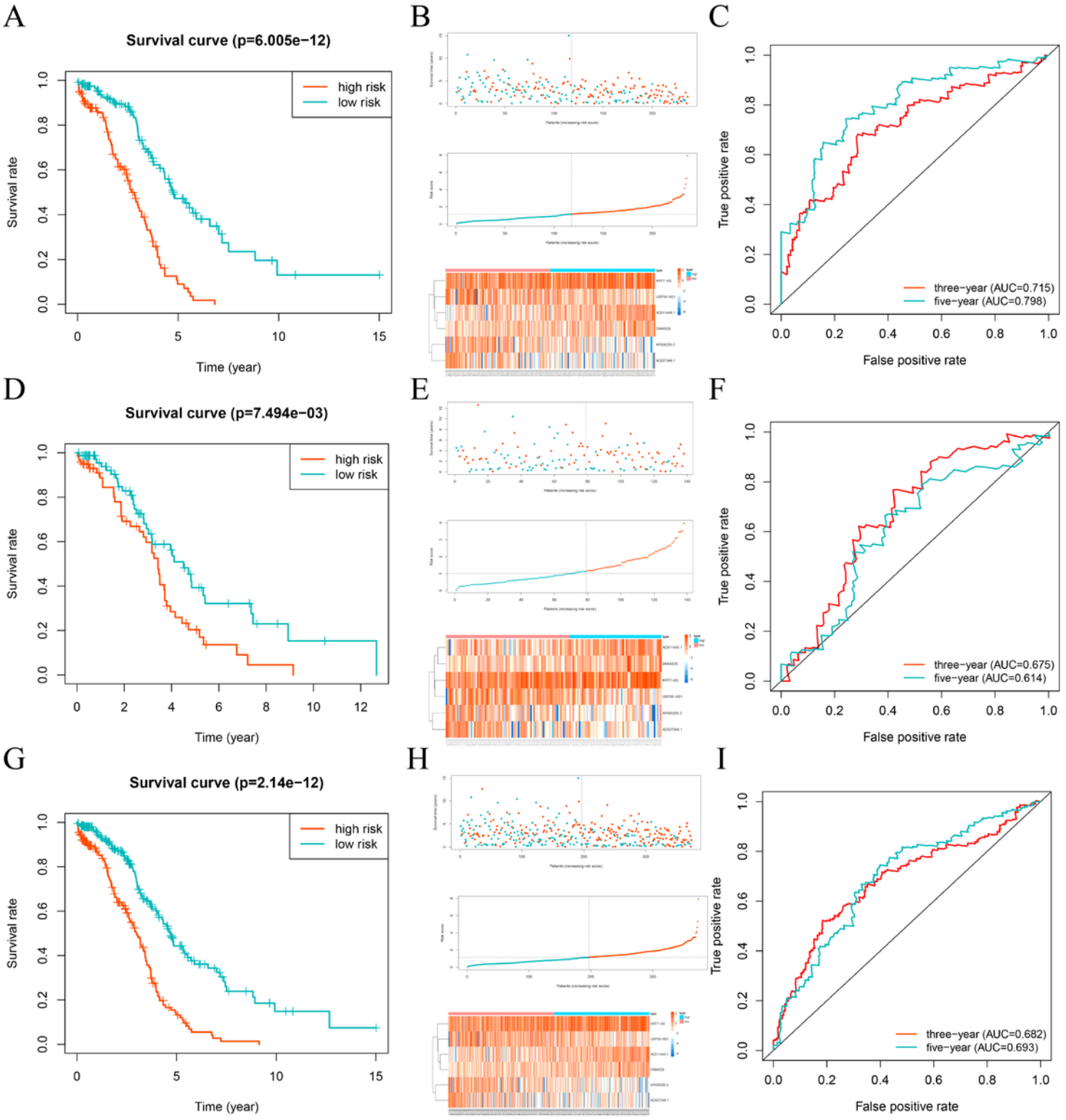

Figure 4

Verification of survival prediction ability and analysis of risk score of 6-lncRNA signature in EOC. (A,D,G) Time-dependent ROC analysis, survival state chart, risk score, risk curve, heat map of lncRNA expression, and Kaplan–Meier curve of the 6-lncRNA signature in the training set. (B,E,H) Time-dependent ROC analysis, survival state chart, risk score, risk curve, heat map of lncRNA expression, and Kaplan–Meier curve of the 6-lncRNA signature in the validation set. (C,F,I) Time-dependent ROC analysis, survival state

chart, risk score, risk curve, heat map of lncRNA expression, and Kaplan–Meier curve of the 6-lncRNA signature in the combination set.

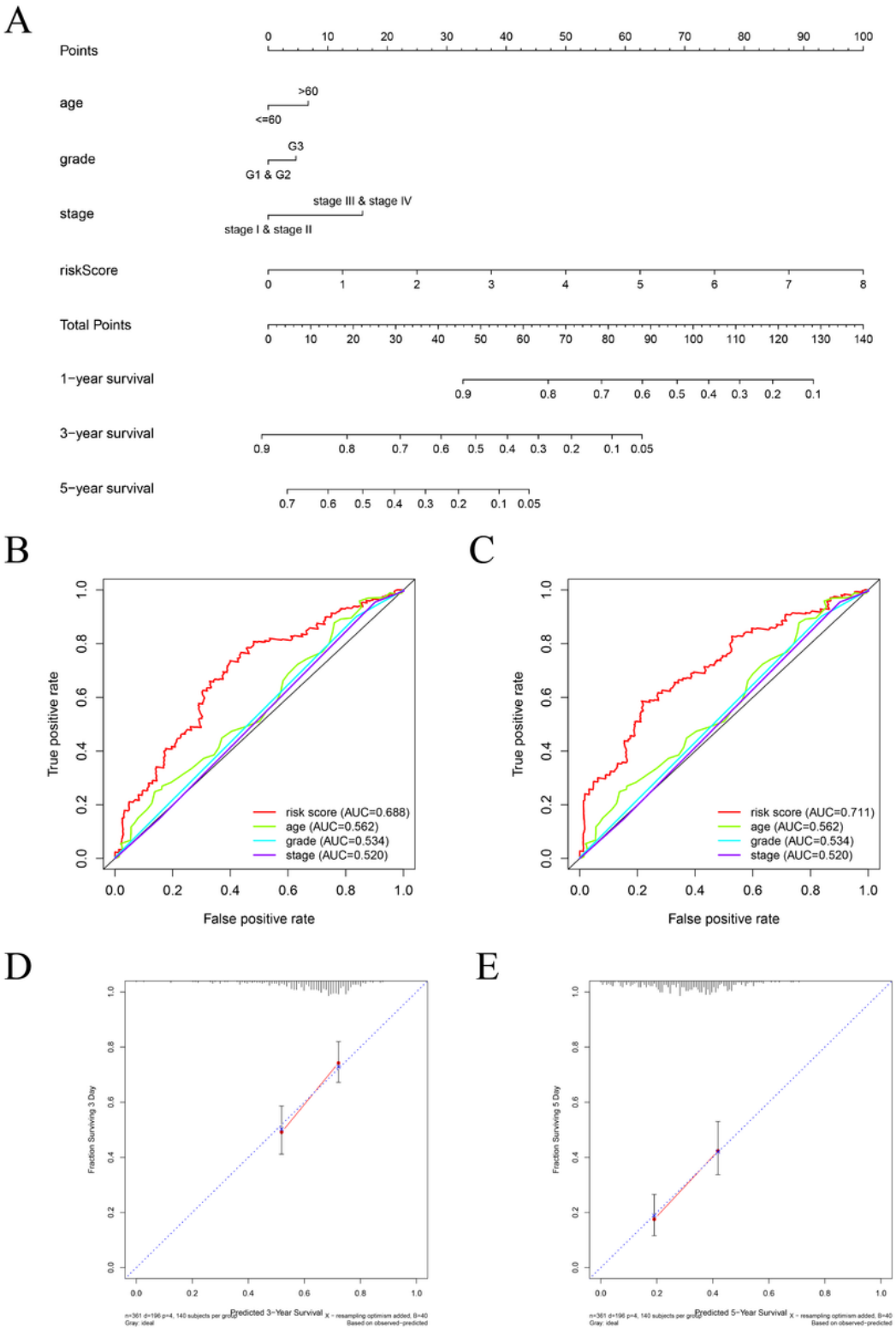

Figure 5

Nomogram for predicting overall survival probability of EOC patients. (A) The nomogram was built based on age, histological grade, FIGO stage, and risk score. (B,C) The time-dependent ROC of nomogram for 3-,

and 5-year overall survival in EOC patients. (D,E) The calibration curves for internal validation of the nomogram at 3-, and 5-year.

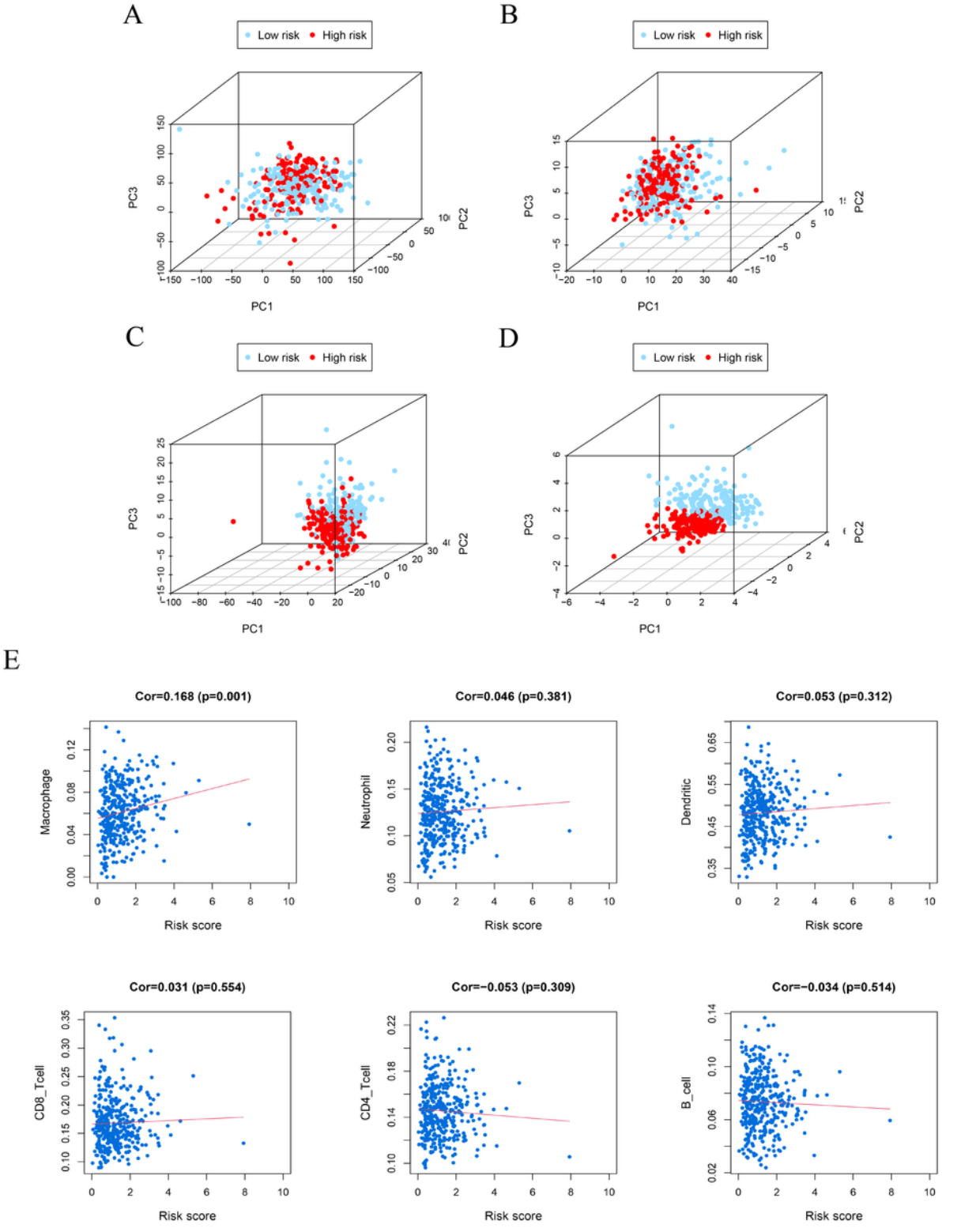

Figure 6

The immune status and immune cell infiltration in high-risk and low-risk groups were evaluated by principal component analysis (PCA). (A) PCA map based on genome-wide expression set. (B) PCA map based on immune-related gene expression set. (C) PCA map based on immune-related lncRNA set. (D)

PCA map based on 6-lncRNA signature. (E) The relationship between risk score and infiltration abundance of six kinds of immune cells was analyzed based on 6-lncRNA signature, which were macrophage, neutrophil, dendritic cell, CD4T cell, CD8T cell and B cell.

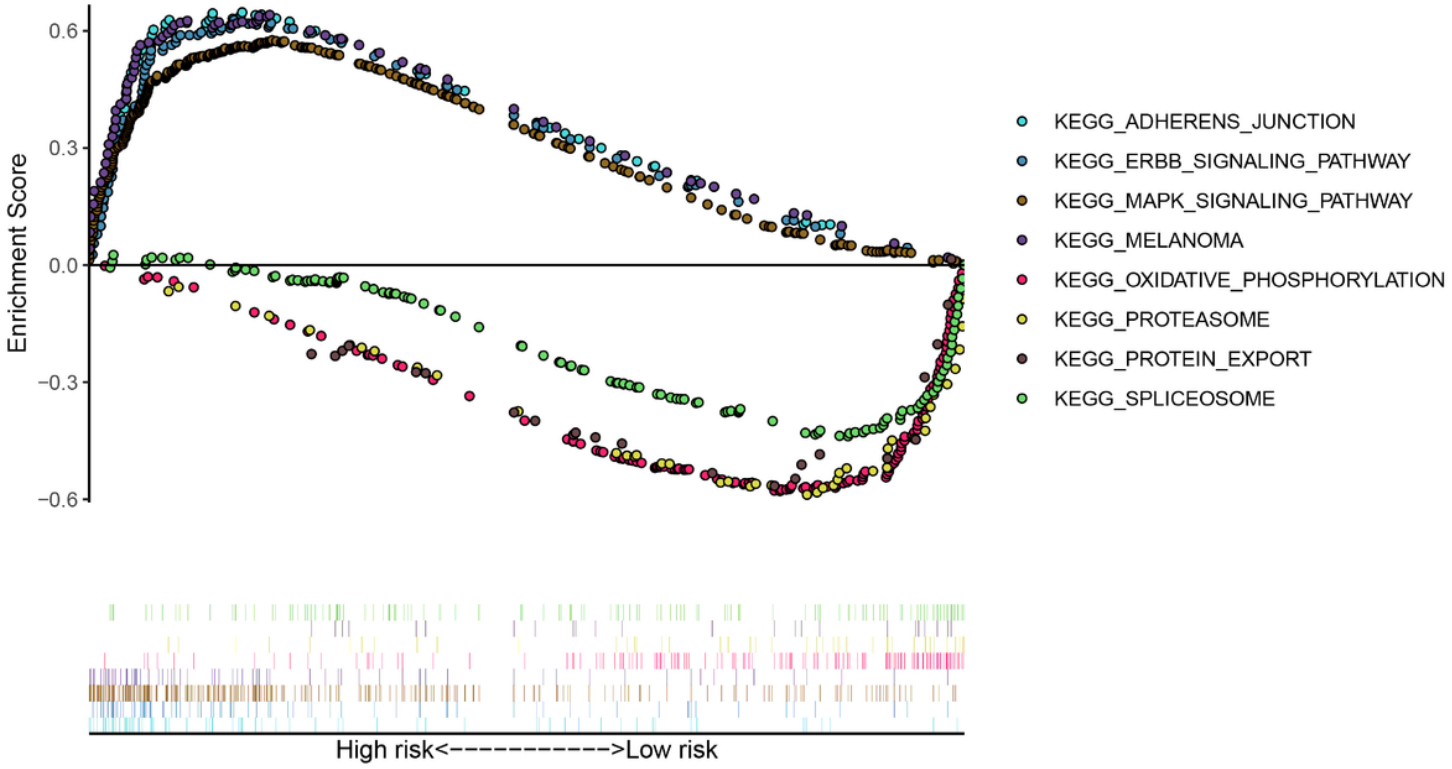

Figure 7

Gene set enrichment analysis (GSEA) between high-risk and low-risk groups based on the prediction model.

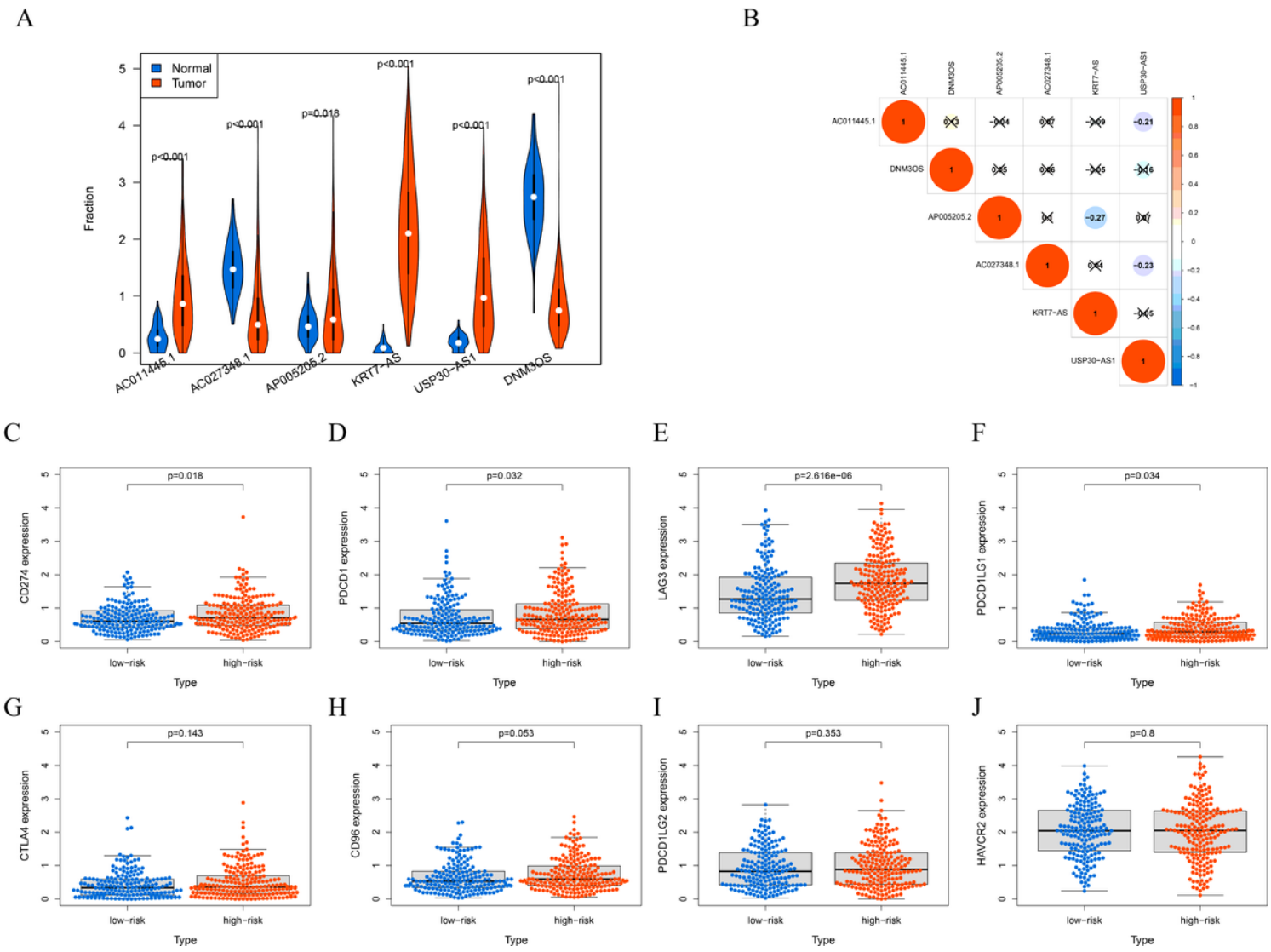

Figure 8

Expression of 6-lncRNA and immune checkpoint related genes. (A)The expression levels of 6 kinds of lncRNAs in normal ovarian tissues and ovarian cancer tissues. (B) The Pearson correlation among 6 kinds of lncRNAs. Box plots visualized the correlation between risk score and immune checkpoint related genes, CD274 (C), PDCD1 (D), LAG3 (E), PDCD1LG1 (F), CTLA-4 (G), CD96 (H), PDCD1LG2 (H) and HAVCR2 (G).

Supplementary Files

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

- MingjunZhangSupplementaryTableS1.xlsx
- MingjunZhangSupplementaryTableS2.xlsx