An Immune-Related Long Noncoding RNA Signature as a Prognostic Biomarker for Human Endometrial Cancer

Ziwei Wang
Department of Obstetrics and Gynecology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology,Wuhan 430022, China

Yan Liu
Department of Obstetrics and Gynecology, Union Hospital, Tongji Medical College,Huazhong University of Science and Technology, Wuhan 430022, China

Jun Zhang
Department of Obstetrics and Gynecology, Union Hospital,Tongji Medical College,Huazhong University of Science and Technology, Wuhan 430022,China

Rong Zhao
Department of Obstetrics and Gynecology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

Xing Zhou
Department of Obstetrics and Gynecology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

Hongbo Wang (hb_wang1969@sina.com)
Wuhan Union Hospital https://orcid.org/0000-0001-7090-1750

Research article

Keywords: endometrial cancer, immune-related lncRNA, prognosis, biomarker

Posted Date: December 1st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-115869/v1

Version of Record: A version of this preprint was published at Journal of Oncology on December 10th, 2021. See the published version at https://doi.org/10.1155/2021/9972454.

Abstract

Endometrial cancer is one of the most common malignant tumors threatening women's health. Recently, immunity and long noncoding RNA (lncRNA) have become hot topics in oncology. Here, to search for prognostic biomarkers, immune-related lncRNAs were identified by collecting endometrial cancer samples' information from The Cancer Genome Atlas (TCGA) database, and immune-related genes from Gene set enrichment analysis (GSEA) gene sets. These included ELN-AS1, AC103563.7, PCAT19, AF131215.5, LINC01871, AC084127.1, NRAV, SCARNA9. AL049539.1, POC1B-AS1, AC108134.4, and AC019080.5. Models based on these 12 immune-related lncRNAs were constructed. Survival analysis showed that the survival rate in the high-risk group was significantly lower than in the low-risk group. Independent prognostic analysis results showed that the patient's age, pathological grade, FIGO stage, and risk status were risk factors. Clinical correlation analysis showed that the 12 immune-related lncRNAs correlated with patients' age, pathological grade, and FIGO stage. After principal component analysis and functional annotation, to further determine the effects of these lncRNAs on prognosis, samples were divided into training and validation groups. Multivariate Cox regression analysis of the training group showed that the patient's age, FIGO stage, and risk status were prognostic risk factors. A nomogram constructed with the risk factors was used to estimate the patient's survival rate. C-indexes were calculated and multi-index ROC curves were plotted to evaluate the accuracy and stability of the nomogram. Finally, Kaplan-Meier survival analysis showed that age, pathological grade, FIGO stage, and risk status were all related to patients' survival. In summary, we identified 12 immune-related lncRNAs that affect the prognosis of endometrial cancer, and that may act as therapeutic targets and molecular biomarkers for the disease.

1 Introduction

Endometrial cancer is one of the most common malignant tumors in women. However, the mechanisms underlying endometrial cancer occurrence and development are unknown. Recent studies indicate that hypertension, diabetes, obesity, and estrogen replacement therapy are risk factors for endometrial cancer[1–4]. The main clinical manifestation of endometrial cancer is postmenopausal vaginal bleeding[5]. Currently, the main treatment methods are surgery, chemotherapy, radiotherapy, and neoadjuvant therapy [6–9]. In general, the prognosis of endometrial cancer is good, although some patients with higher FIGO stages have a poor prognosis [10–12].

lncRNA is a type of noncoding RNA that exceeds 200 nucleotides in length [13]. Recent studies have found that lncRNA dynamically regulate gene expression at multiple levels [14–18]. lncRNAs play many roles; for instance, they are involved in the growth and development of the body [19], as well as in the occurrence and development of many diseases, such as cardiovascular [20, 21] and neurodegenerative [22] diseases, and malignant tumors [23].

Recently, tumor immunotherapy has become a hot topic. Immune cells and molecules are complex components in the tumor microenvironment and can promote or inhibit malignant tumor progression.

The immune response in the tumor microenvironment has an important effect on tumor proliferation, invasion, and metastasis [24–29]. Many studies have focused on the immune response in the tumor microenvironment to regulate tumor progression. For example, in breast and colon cancer and melanoma models, by reprograming the TAM population to a pro-inflammatory phenotype and increasing tumor immunogenicity, anti-MARCO monoclonal antibodies can inhibit tumor progression [30].

In this study, we explored the transcriptome data of endometrial cancer samples in The Cancer Genome Atlas (TCGA) database. Immune-related lncRNAs co-expressed with immune-related genes were identified, and prognostic nomogram was constructed to predict the overall survival rate. These lncRNAs may become potential therapeutic targets and molecular biomarkers of endometrial cancer.

2 Materials And Methods

2.1 Data Sources

Fragments per kilobase million (FPKM) RNA seq data of 575 samples were downloaded from the TCGA database (https://portal.gdc.cancer.gov/), including 23 normal samples and 552 endometrial cancer samples. Clinical information on endometrial cancer samples was downloaded from the UCSC Xena (https://xenabrowser.net). We downloaded the gene sets IMMUNE_RESPONSE (systematic name: M13664), IMMUNE_SYSTEM_PROCESS (systematic name: M19817), GO_NEGATIVE_REGULATION_OF_ADAPTIVE_IMMUNE_RESPONSE (systematic name: M10422), GO_NEGATIVE_REGULATION_OF_IMMUNE_RESPONSE (systematic name: M15641), GO_T_CELL_ACTIVATION_INVOLVED_IN_IMMUNE_RESPONSE (Systematic name: M10714) from the Gene set enrichment analysis (GSEA) database (http://software.broadinstitute.org/gsea/index.jsp). Immune-related genes were obtained from the IMMUNE_RESPONSE (systematic name: M13664) and IMMUNE_SYSTEM_PROCESS (systematic name: M19817) gene sets. Functional annotation was performed using gene sets IMMUNE_RESPONSE (systematic name: M13664), GO_NEGATIVE_REGULATION_OF_ADAPTIVE_IMMUNE_RESPONSE (systematic name: M10422), GO_NEGATIVE_REGULATION_OF_IMMUNE_RESPONSE (systematic name: M15641), GO_T_CELL_ACTIVATION_INVOLVED_IN_IMMUNE_RESPONSE (Systematic name: M10714).

2.2 Acquisition of Immune-Related lncRNAs

Through sorting and analyzing the transcriptome data of the endometrial cancer samples from the TCGA database, the expression matrices of mRNA and lncRNA of endometrial cancer samples were obtained. Immune-related genes were obtained from the IMMUNE_RESPONSE (systematic name: M13664) and IMMUNE_SYSTEM_PROCESS (systematic name: M19817) gene sets. Immune-related genes expression matrix was obtained through collecting and analyzing the mRNA expression matrix and immune-related genes together using the limma package of R software. Lastly, the limma package of R software was used to screen the immune-related lncRNAs that had a co-expression relationship with immune-related genes, and an immune-related lncRNAs expression matrix was obtained (correlation filter = 0.5, p-value filter = 0.001).

2.3 Cox Regression, Survival, Independent Prognostic, and Clinical Correlation Analyses

Through collecting and analyzing the clinical data, univariate (p-value filter = 0.01) and multivariate Cox regression analysis were performed using the survival package of R software. The co-expression network was plotted using Cytoscape software (Cytoscape_v3.7.2). The survival and survminer packages of R software were used to draw the survival curve of high-risk and low-risk groups. Risk curves were drawn using the pheatmap package of R software. Univariate and multivariate independent prognostic analyses were performed using the survival package of the R software, and a multi-index ROC curve was plotted using the survivalROC package of the R software to evaluate the accuracy of the constructed model. The ggpubr package of R software was used to perform clinical correlation analysis.

2.4 Immune and Stromal Scores

We used the limma and estimate packages of R software to calculate the immune scores and stromal scores of all samples. R software was then used to analyze the immune scores and stromal scores of different risk states.

2.5 Principal Component and Gene Set Enrichment Analyses

The principal component analysis was performed using the limma and scatterplot3d packages of R software. Functional annotation of the 12 immune-related lncRNAs was performed using GSEA software (GSEA_4.0.2).

2.6 Prognostic Nomogram and Survival Analysis for Independent Prognostic Risk Factors

Using the foreign, survival, and caret packages of R software, 70% of the tumor samples were placed into the training group, and 30% into the validation group. The rms, foreign, and survival packages of R software were used to perform multivariate Cox regression analysis of the training group and calculate the C-index of the training and validation groups. Nomogram was constructed using the rms, foreign, and survival packages of R software. The survival and timeROC packages of R software were used to construct multi-index ROC curves for the training and validation groups. The survival package of R software was used to perform Kaplan-Meier survival analysis for the training group.

2.7 Data Statistics

All statistical analysis was performed using R software (R-3.6.1), strawberry-Perl-5.30.0.1. P-values < 0.05 were considered statistically significant.

3 Results

3.1 Immune-Related lncRNAs Associated with Prognosis

To identify immune-related lncRNAs associated with prognosis, a total of 332 immune-related genes were selected from the GSEA data sets. Through collecting and analyzing data, we identified 137 immune-related genes and 363 immune-related lncRNAs with a co-expression relationship with these immune-related genes. Next, univariate Cox regression analysis of all immune-related lncRNAs was performed, and a forest plot was obtained (Fig. 1A). As shown in Fig. 1, green columns indicated protective lncRNAs (HR < 1), and red indicated risk-associated lncRNAs (HR > 1). Next, 34 lncRNAs were subjected to multivariate Cox regression analysis, and 12 immune-related lncRNAs associated with prognosis were obtained (Table 1): ELN-AS1, AC103563.7, PCAT19, AF131215.5, LINC01871, AC084117.1, NRAV, SCARNA9, AL049539.1, POC1B-AS1, AC108134.4, AC019080.5. Based on the median of risk scores, all endometrial cancer samples were divided into high-risk and low-risk groups: risk score = (expression level of ELN-AS1 * 0.229) + (expression level of AC103563.7 * 0.313) + (expression level of PCAT19 * -0.277) + (expression level of AF131215.5 * 0.252) + (expression level of LINC01871 * -0.357) + (expression level of AC084117.1 * 0.449) + (expression level of NRAV * -0.433) + (expression level of SCARNA9*-0.339) + (expression level of AL049539.1 * 0.476) + (expression level of POC1B-AS1 * -0.758) + (expression level of AC108134.4 * -0.262) + (expression level of AC019080.5 * 0.899). Finally, an immune-related lncRNAs-immune-related genes network was constructed using Cytoscape software (Fig. 1B).

Figure 1. Immune-related lncRNAs associated with prognosis. (A) Forest plot of univariate Cox regression analysis. The p-value, hazard ratio, and 95% confidence intervals of the immune-related genes are shown. Red and green indicate a risk-associated (HR > 1) and a protective (HR < 1) lncRNA, respectively. (B) Immune-related lncRNAs-immune-related genes network. Green and red represent immune-related genes and 12 immune-related lncRNAs, respectively.

Table 1
Multivariate Cox regression analysis of the 12 immune-related lncRNAs

LncRNA	coefficient	HR	HR.95L	HR.95H	p-value
ELN-AS1	0.229	1.257	0.963	1.640	0.092
AC103563.7	0.313	1.367	1.018	1.836	0.038
PCAT19	-0.277	0.758	0.564	1.018	0.065
AF131215.5	0.252	1.287	0.924	1.794	0.136
LINC01871	-0.357	0.700	0.560	0.873	0.002
AC084117.1	0.449	1.567	1.065	2.307	0.023
NRAV	-0.433	0.648	0.407	1.034	0.069
SCARNA9	-0.339	0.712	0.575	0.883	0.002
AL049539.1	0.476	1.610	1.041	2.488	0.032
POC1B-AS1	-0.758	0.469	0.203	1.083	0.076
AC108134.4	-0.262	0.770	0.552	1.073	0.122
AC019080.5	0.899	2.457	1.352	4.464	0.003

Regression coefficients, p-value, hazard ratio, and 95% confidence interval of the immune-related lncRNAs are shown.

3.2 Survival Analysis and Risk Curves

To compare the survival rate of different risk statuses based on grouping results, survival analysis was performed, and survival curves were obtained (Fig. 2A). As shown in Fig. 2, the survival rate of the high-risk group was lower than that of the low-risk group. Risk curves for the high-risk and low-risk groups were then obtained (Figs. 2B, 2C). The results show that the risk score of the high-risk group was higher than that of the low-risk group, and the survival time of the high-risk group was shorter than that of the low-risk group. Next, a heat map was obtained to compare the expression levels of the 12 immune-related lncRNAs in different risk status (Fig. 2D). As shown in Fig. 2, the expression levels of AC103563.7, AF131215.5, AC084117.1, and AL049539.1 in the high-risk group were higher than in the low-risk group, and the expression levels of NRAV, ELN − AS1, PCAT19, AC108134.4, LINC01871, and SCARNA9 in the low-risk group were higher than in the high-risk group. Survival curves were then plotted for the 12 immune-related lncRNAs to analyze the effects of these lncRNAs on survival (Fig. 2E-2H). As shown in Fig. 2, the overall survival rate associated with LINC01871, AC108134.4, and POC1B − AS1 in the low-expression group was lower than in the high expression group. The overall survival rate associated with AC019080.5 in the low-expression group was higher than in the high expression group ($P < 0.05$).

Figure 2. Survival analysis and risk curves. (A) Survival curve of endometrial cancer. Red and blue indicate high-risk and low-risk groups, respectively. (B) Risk curve of endometrial cancer. Red and green

indicate high-risk and low-risk groups, respectively. (C) Scatter plot of different survival status of endometrial cancer patients. Red and green dots denote patients that are dead or alive, respectively. (D) Hierarchical clustering of 12 immune-related lncRNAs. Differences in the expression level of 12 immune-related lncRNAs in different risk status. (E-H) Survival curves of immune-related lncRNAs. Blue and red represent low and high expression groups, respectively.

3.3 Independent Prognostic and Clinical Correlation Analyses

To analyze the impact of patients' age, pathologic grade, and FIGO stage on prognosis, through collecting and analyzing the clinical data, we performed univariate and multivariate independent prognostic analyses (Fig. 3A, 3B). The results show that patients' age, pathological grade, and FIGO stage were all associated with prognosis and were all risk factors for endometrial cancer. The older the patient, the higher the pathological grade and FIGO stage, the poorer the prognosis. We then constructed a multi-index ROC curve to evaluate the accuracy of all models (Fig. 3C). As shown in Fig. 3, risk score (area under the curve (AUC) = 0.709), age (AUC = 0.614), grade (AUC = 0.652), stage (AUC = 0.709) analysis a good accuracy of the constructed model. Clinical correlation analysis was then performed for the 12 immune-related lncRNAs and patients' age, pathological grade, and FIGO stage (Fig. 3D-3F). As shown in Fig. 3, AC103563.7, AL049539.1, ELN-AS1, NRAV, and POC1B-AS1 were associated with age, pathological grade, and FIGO stage, AC108134.4, and PCAT19 with pathological grade and FIGO stage, AF131215.5, and SCARNA9 with pathological grade.

Figure 3. Independent prognostic analysis and clinical correlation analysis. (A) Forest plot of univariate independent prognostic analysis. The p-value, hazard ratios, and 95% confidence intervals are shown. Red and green indicate a risk-related (HR > 1) and a protective (HR < 1) factor, respectively. (B) Forest plot of multivariate independent prognostic analysis. The p-value, hazard ratios, and 95% confidence intervals are shown. Red and green indicate a risk-related (HR > 1) and a protective (HR < 1) factor, respectively. (C) Plot of multi-index ROC curve showing risk score (AUC = 0.709), age (AUC = 0.614), grade (AUC = 0.652), and stage (AUC = 0.709). (D) lncRNA expression levels in patients aged below (red) or above (blue) 65. (E) lncRNA expression levels in different pathological grades. (F) lncRNA expression levels in different FIGO stages. *p < 0.05, **p < 0.01, ***p < 0.001. ns, p > 0.05.

3.4 Immune Scores and Stromal Scores

To compare tumor microenvironment differences at different risk states using the ESTIMATE algorithm, ESTIMATE, immune, and stromal scores of all samples were calculated. By collecting the data, we obtained ESTIMATE, immune, and stromal scores of different risk status, the results show that ESTIMATE scores ranged from − 3166.978 to 3990.147, immune scores ranged from − 1359.509 to 3614.677, and stromal scores ranged from − 2224.623 to 860.431. Next, we draw box plots of ESTIMATE scores, immune scores, and stromal scores at different risk status (Fig. 4A-4C), The average ESTIMATE scores, immune scores, and stromal scores of the high-risk group were lower than those of the low-risk group.

Figure 4. Immune scores and stromal scores. (A) Distribution of ESTIMATE scores of different risk status. Box-plot shows that there is a significant difference between high-risk and low-risk groups ($p < 0.05$). (B) Distribution of immune scores of different risk status. Box-plot shows that there is a significant difference between high-risk and low-risk groups ($p < 0.05$). (C) Distribution of stromal scores of different risk status. Box-plot shows that there is a significant difference between high-risk and low-risk groups ($p < 0.05$).

3.5 Principal Component and Gene Set Enrichment Analyses

Principal component analyses of the expression of all immune-related lncRNAs and 12 immune-related lncRNAs associated with prognosis were performed to determine whether there were differences in distribution between the high-risk and low-risk groups (Fig. 5A, 5B). As shown in Fig. 5, the high-risk and low-risk groups showed better separations, and the immune status of the high-risk and low-risk groups differed. Next, GSEA was performed on the 12 immune-related lncRNAs (Fig. 5C-5F). As shown in Fig. 5, compared with the high-risk group, the low-risk group was enriched in the gene sets IMMUNE_RESPONSE (systematic name: M13664), GO_NEGATIVE_REGULATION_OF_ADAPTIVE_IMMUNE_RESPONSE (systematic name: M10422), GO_NEGATIVE_REGULATION_OF_IMMUNE_RESPONSE (systematic name: M15641), GO_T_CELL_ACTIVATION_INVOLVED_IN_IMMUNE_RESPONSE (Systematic name: M10714). Therefore, these immune-related lncRNAs may be associated with the regulation of immune response.

Figure 5. Principal component and gene set enrichment analyses. (A) Principal components analysis between low-risk and high-risk groups based on all immune-related lncRNAs. Red and green indicate high-risk and low-risk groups, respectively (B) Principal components analysis between low-risk and high-risk groups based on 12 immune-related lncRNAs associated with prognosis. Red and green indicate high-risk and low-risk groups, respectively (C-F) GSEA indicate significant enrichment in immune-related phenotypes in the low-risk patients.

3.6 Prognostic Nomogram for Overall Survival rate

To further evaluate the effect of the model constructed based on the 12 immune-related lncRNAs on prognosis, based on the above results, we used the patient's risk status based on these lncRNAs as an independent risk factor affecting prognosis and divided all samples into a training (70%) and a validation (30%) group. A multivariate Cox regression analysis using training samples was performed to analyze the patient's age, pathological grade, FIGO stage, and risk status effects on prognosis (Table 2). The results showed that the age, FIGO stage, and risk status were all associated with prognosis. A prognostic nomogram that integrated all significant independent factors for overall survival was constructed to predict the survival rate (Fig. 6A). Next, a multi-index ROC curve using the training samples was made to evaluate the accuracy of the model (Fig. 6B). The results showed the 3-year survival (AUC = 0.808) and 5-year survival (AUC = 0.831) rates. The C-index of training group was 0.794 (standard error ± 0.029). Therefore, the prognostic nomogram showed good accuracy. Next, we used the verification samples to draw a multi-index ROC curve to evaluate the stability of the prognostic nomogram (Fig. 6C). As shown in Fig. 5, the 3-year survival (AUC = 0.834) and 5-year survival (AUC = 0.843) were larger than the AUC values

of the training samples, and the C-index was 0.818 (standard error ± 0.036), which was larger than the C-index of the training samples. In summary, the prognostic nomogram showed good stability.

Table 2
Multivariate Cox regression analysis of clinical characteristics

Variable	coefficient	HR	lower.95	upper.95	p-value
Age ≥ 65	0.727	2.069	1.158	3.697	0.014
Grade 2	1.131	3.097	0.676	14.201	0.146
Grade 3	0.976	2.654	0.604	11.665	0.196
Stage II	0.726	2.067	0.881	4.847	0.095
Stage III	1.173	3.231	1.687	6.188	0.000
Stage IV	1.517	4.559	2.006	10.358	0.000
High-risk group	1.064	2.897	1.402	5.986	0.004
Age ≥ 65 was compared to age < 65, Grade 2 and Grade 3 were compared to Grade 1, Stage II, Stage III and Stage IV were compared to Stage Ⅰ, and High-risk group was compared to Low-risk group. Regression coefficients, p-value, hazard ratio, and 95% confidence interval of the clinical characteristics are shown.					

Figure 6. Prognostic nomogram for overall survival rate. (A) Survival nomogram. (to use the nomogram, an individual patient's value is located on each variable axis, and a line is drawn upward to determine the number of points received for each variable value. The sum of these numbers is located on the Total Points axis, and a line is drawn downward to the survival axes to determine the likelihood of 3- or 5-year survival). (B) Multi-index ROC curve of training samples. Red and blue indicate 3-year and 5-year survival, respectively. (C) Multi-index ROC curve of validation samples. Red and blue indicate 3-year and 5-year survival, respectively.

3.7 Survival Analysis for Independent Prognostic Risk Factors

To analyze the effects of age, pathological grade, FIGO stage, and risk status on overall survival. Each risk factor was subjected to Kaplan-Meier survival analysis, and survival curves were drawn to analyze the impact of each factor on prognosis. (Fig. 7A-7D). As shown in Fig. 7, age, pathological grade, FIGO stage, and risk status based on the 12 immune-related lncRNAs all affected the patients' prognosis ($p < 0.05$).

Figure 7. Survival analysis of independent prognostic risk factors. (A) Survival curve of patients younger or older than 65. (B) Survival curve of groups in Grade 1, Grade 2, and Grade 3. (C) Survival curve of

groups in stage Ⅰ, stage Ⅱ, stage Ⅲ, and stage Ⅳ. (D) Survival curve of patients in the high-risk and low-risk groups divided based on the 12 immune-related lncRNAs.

4 Discussion

Recent studies have found that the immune system has a dual effect on tumor progression. The immune components in the tumor microenvironment can both promote or inhibit the progression of malignant tumors [31]. Infiltrating immune cells and cytokines secreted by immune cells in the tumor microenvironment and chemokines are involved in tumor progression [32–34]. In recent years, with the widespread application of the immune checkpoint inhibitors PD-1/PD-L1 and CTLA-4, more and more studies have been devoted to suppressing the progress of tumors by regulating the immune components in the tumor microenvironment [35–37]. For example, through regulating mitotic checkpoints and chromosome stability, TIF1γ inhibits tumor progression [38]. lncRNAs have complex functions and regulate the occurrence and development of malignant tumors through various mechanisms. For example, by inhibiting CUL4A-mediated LATS1 ubiquitination and increasing YAPS127 phosphorylation, lncRNA uc.134 can inhibit the progression of liver cancer [39]. In gastric cancer, low-expression lncRNA LINC00261 can inhibit tumor metastasis via regulating EMT [40]. In liver cancer, through degrading HNRNPA2B1 via ubiquitination, which reduces the stability of p52 and p65 mRNAs and inhibiting the NF-κB signaling pathway in HCC cells, lncRNA miR503HG inhibits tumor metastasis [41].

lncRNAs also regulate tumor progression by regulating immune components in the tumor microenvironment. For example, in lung and breast cancer, lncRNA NKILA can up-regulate tumor-specific cytotoxic T lymphocytes and type 1 helper T (TH1) cells' sensitivity to activation-induced cell death by inhibiting NF-κB activity, thereby facilitating immune escape [42]. In colorectal cancer, through regulating SATB2, lncRNA SATB2-AS1 can regulate the expression of TH-1 type chemokines transcription and immune cell density in the tumor microenvironment, thus suppressing tumor metastasis [43].

However, it is relatively unknown whether lncRNAs modulate endometrial cancer progression via immune components-regulation in the tumor microenvironment. In the present work, the expression matrix of immune-related lncRNAs was analyzed through collecting the immune gene sets from the GSEA database and the endometrial cancer sample information from the TCGA database. Univariate and multifactorial Cox regression analyses were performed, and 12 immune-related lncRNAs were identified as having an important influence on endometrial cancer: ELN-AS1, AC103563.7, PCAT19, AF131215.5, LINC01871, AC084127.1, NRAV, SCARNA9, AL049539.1, POC1B-AS1, AC108134.4, AC019080.5.

Among the 12 immune-related lncRNAs, PCAT19 is downregulated in non-small cell lung carcinoma and regulates the development of prostate cancer [44–46]. Through modulating the miR-182/PDK4 axis, PCAT19 promotes the proliferation of laryngocarcinoma cells [47]. SCARNA9 is downregulated in cervical cancer [48]. Nevertheless, the specific functions of ELN-AS1, AC103563.7, AF131215.5, LINC01871, AC084127.1, NRAV, AL049539.1, POC1B-AS1, AC108134.4, and AC019080.5 in the tumor microenvironment remain unknown. We speculate that these lncRNAs may regulate the occurrence and

development of endometrial cancer by regulating immune components in the tumor microenvironment. However, to understand the respective functions, further experimental validation is required. It is foreseen that these 12 immune-related lncRNAs might represent new molecular biomarkers and therapeutic targets for endometrial cancer.

5 Conclusion

By sorting and analyzing the transcriptome information of endometrial cancer samples from the TCGA database, we identified 12 immune-related lncRNAs. These molecules might play important regulatory roles in the occurrence and development of endometrial cancer and may represent potential therapeutic targets. However, their specific roles and mechanisms need further experimental validation.

6 Abbreviations

TCGA, The Cancer Genome Atlas; GSEA, gene set enrichment analysis; HR, hazard ratio; RNA seq, RNA sequencing. NES, normalized enrichment score.

Declarations

7 Disclosure

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

8 Author Contributions

Ziwei Wang, Yan Liu, and Hongbo Wang conceived and designed the study and contributed to writing of the manuscript. Ziwei Wang, Yan Liu, and Rong Zhao performed the analysis procedures. Jun Zhang and Xing Zhou analyzed the results. Ziwei Wang, Rong Zhao, and Xing Zhou contributed with analysis tools. All authors reviewed the manuscript.

9 Funding

This study was financially supported by a National Natural Science Foundation of China Grant (No. 81974409).

10 Acknowledgments

The authors are very grateful for the valuable data provided by GSEA and the TCGA databases. The authors also want to thank Yingchao Zhao for their assistance in data collecting.

11 Data Availability Statement

We obtained RNA seq data from TCGA (http://cancergenome.nih.gov/) and clinical information from the UCSC Xena (https://xenabrowser.net). All gene sets were downloaded from GSEA (http://software.broadinstitute.org/gsea/index.jsp).

12 Ethical guidelines

Not applicable.

References

1. Colombo N, et al. Endometrial cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology: official journal of the European Society for Medical Oncology. 2013;24 **Suppl 6**:vi33–8.
2. Ghanbari Andarieh M, et al. Risk Factors for Endometrial Cancer: Results from a Hospital-Based Case-Control Study. Asian Pacific journal of cancer prevention: APJCP. 2016;17(10):4791–6.
3. Gong T-T, Wang Y-L, Ma X-X. Age at menarche and endometrial cancer risk: a dose-response meta-analysis of prospective studies. Scientific reports. 2015;5:14051.
4. Setiawan VW, et al. Type I and II endometrial cancers: have they different risk factors? Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2013;31(20):2607–18.
5. Sjögren LL, Mørch LS, Løkkegaard E. Hormone replacement therapy and the risk of endometrial cancer: A systematic review. Maturitas. 2016;91:25–35.
6. Katz MS, Smith L, Simcock R. Treatment toxicity in endometrial cancer: can we identify and manage it better? Lancet Oncol. 2018;19(5):602.
7. Meyer LA, et al. Postoperative Radiation Therapy for Endometrial Cancer: American Society of Clinical Oncology Clinical Practice Guideline Endorsement of the American Society for Radiation Oncology Evidence-Based Guideline. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2015;33(26):2908–13.
8. Bestvina CM, Fleming GF. Chemotherapy for Endometrial Cancer in Adjuvant and Advanced Disease Settings. Oncologist. 2016;21(10):1250–9.
9. de Haydu C, et al. An update on the current pharmacotherapy for endometrial cancer. Expert opinion on pharmacotherapy. 2016;17(4):489–99.
10. Weiderpass E, et al. Trends in corpus uteri cancer mortality in member states of the European Union. European journal of cancer (Oxford England: 1990). 2014;50(9):1675–84.
11. Siegel RL, Miller KD, Jemal A, *Cancer statistics*, 2015. CA: a cancer journal for clinicians, 2015. **65**(1).
12. Ferlay J, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer. 2015;136(5):E359–86.
13. Chen L-L. Linking Long Noncoding RNA Localization and Function. Trends Biochem Sci. 2016;41(9):761–72.

14. Chen H, et al., *Non-coding Transcripts from Enhancers: New Insights into Enhancer Activity and Gene Expression Regulation.* Genomics, proteomics & bioinformatics, 2017. **15**(3): p. 201–207.
15. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol. 2013;20(3):300–7.
16. Beermann J, et al. Non-coding RNAs in Development and Disease: Background, Mechanisms, and Therapeutic Approaches. Physiological reviews. 2016;96(4):1297–325.
17. Dykes IM, Emanueli C. Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA. Genomics, proteomics bioinformatics, 2017. 15(3): 177–86.
18. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet. 2016;17(1):47–62.
19. Chen Z. Progress and prospects of long noncoding RNAs in lipid homeostasis. Molecular metabolism. 2016;5(3):164–70.
20. Wang K, et al. APF lncRNA regulates autophagy and myocardial infarction by targeting miR-188-3*p.*. Nature communications. 2015;6:6779.
21. Piccoli M-T, et al. Inhibition of the Cardiac Fibroblast-Enriched lncRNA Prevents Cardiac Fibrosis and Diastolic Dysfunction. Circulation research. 2017;121(5):575–83.
22. Perry RB-T, et al., *Regulation of Neuroregeneration by Long Noncoding RNAs.* Molecular cell, 2018. 72(3).
23. Martens-Uzunova ES, et al. Long noncoding RNA in prostate, bladder, and kidney cancer. European urology. 2014;65(6):1140–51.
24. Ahmed N, et al. Tumour microenvironment and metabolic plasticity in cancer and cancer stem cells: Perspectives on metabolic and immune regulatory signatures in chemoresistant ovarian cancer stem cells. Sem Cancer Biol. 2018;53:265–81.
25. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nature medicine. 2013;19(11):1423–37.
26. Brown JM, Recht L, Strober S. The Promise of Targeting Macrophages in Cancer Therapy. Clinical cancer research: an official journal of the American Association for Cancer Research. 2017;23(13):3241–50.
27. Ngambenjawong C, Gustafson HH, Pun SH. Progress in tumor-associated macrophage (TAM)-targeted therapeutics. Adv Drug Deliv Rev. 2017;114:206–21.
28. Quail DF, Joyce JA. The Microenvironmental Landscape of Brain Tumors. Cancer cell. 2017;31(3):326–41.
29. Sun Y. Tumor microenvironment and cancer therapy resistance. Cancer letters. 2016;380(1):205–15.
30. Georgoudaki A-M, et al. Reprogramming Tumor-Associated Macrophages by Antibody Targeting Inhibits Cancer Progression and Metastasis. Cell reports. 2016;15(9):2000–11.
31. Schreiber RD, Old LJ, Smyth MJ, *Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion.* Science (New York, N.Y.), 2011. **331**(6024): p. 1565–1570.

32. Schietinger A, et al. Tumor-Specific T Cell Dysfunction Is a Dynamic Antigen-Driven Differentiation Program Initiated Early during Tumorigenesis. Immunity. 2016;45(2):389–401.
33. Guillerey C, Huntington ND, Smyth MJ. Targeting natural killer cells in cancer immunotherapy. Nature immunology. 2016;17(9):1025–36.
34. Elkabets M, et al. IL-1β regulates a novel myeloid-derived suppressor cell subset that impairs NK cell development and function. Eur J Immunol. 2010;40(12):3347–57.
35. Ebert PJR, et al. MAP Kinase Inhibition Promotes T Cell and Anti-tumor Activity in Combination with PD-L1 Checkpoint Blockade. Immunity. 2016;44(3):609–21.
36. Francis DM, Thomas SN. Progress and opportunities for enhancing the delivery and efficacy of checkpoint inhibitors for cancer immunotherapy. Adv Drug Deliv Rev. 2017;114:33–42.
37. Syn NL, et al. De-novo and acquired resistance to immune checkpoint targeting. Lancet Oncol. 2017;18(12):e731–41.
38. Pommier RM, et al. TIF1γ Suppresses Tumor Progression by Regulating Mitotic Checkpoints and Chromosomal Stability. Cancer research. 2015;75(20):4335–50.
39. Ni W, et al. A novel lncRNA uc.134 represses hepatocellular carcinoma progression by inhibiting CUL4A-mediated ubiquitination of LATS1. J Hematol Oncol. 2017;10(1):91.
40. Fan Y, et al. Decreased expression of the long noncoding RNA LINC00261 indicate poor prognosis in gastric cancer and suppress gastric cancer metastasis by affecting the epithelial-mesenchymal transition. J Hematol Oncol. 2016;9(1):57.
41. Wang H, et al. Long noncoding RNA miR503HG, a prognostic indicator, inhibits tumor metastasis by regulating the HNRNPA2B1/NF-κB pathway in hepatocellular carcinoma. Theranostics. 2018;8(10):2814–29.
42. Huang D, et al. NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. Nature immunology. 2018;19(10):1112–25.
43. Xu M, et al. LncRNA SATB2-AS1 inhibits tumor metastasis and affects the tumor immune cell microenvironment in colorectal cancer by regulating SATB2. Mol Cancer. 2019;18(1):135.
44. Hua JT, et al., *Risk SNP-Mediated Promoter-Enhancer Switching Drives Prostate Cancer through lncRNA PCAT19*. Cell, 2018. 174(3).
45. Gao P, et al., *Biology and Clinical Implications of the 19q13 Aggressive Prostate Cancer Susceptibility Locus*. Cell, 2018. 174(3).
46. Acha-Sagredo A, et al. Long non-coding RNA dysregulation is a frequent event in non-small cell lung carcinoma pathogenesis. British journal of cancer. 2020;122(7):1050–8.
47. Xu S, Guo J, Zhang W, *lncRNA PCAT19 promotes the proliferation of laryngocarcinoma cells via modulation of the miR-182/PDK4 axis.* Journal of cellular biochemistry, 2019. **120**(8): p. 12810–12821.
48. Roychowdhury A, et al. Deregulation of H19 is associated with cervical carcinoma. Genomics. 2020;112(1):961–70.

Figures

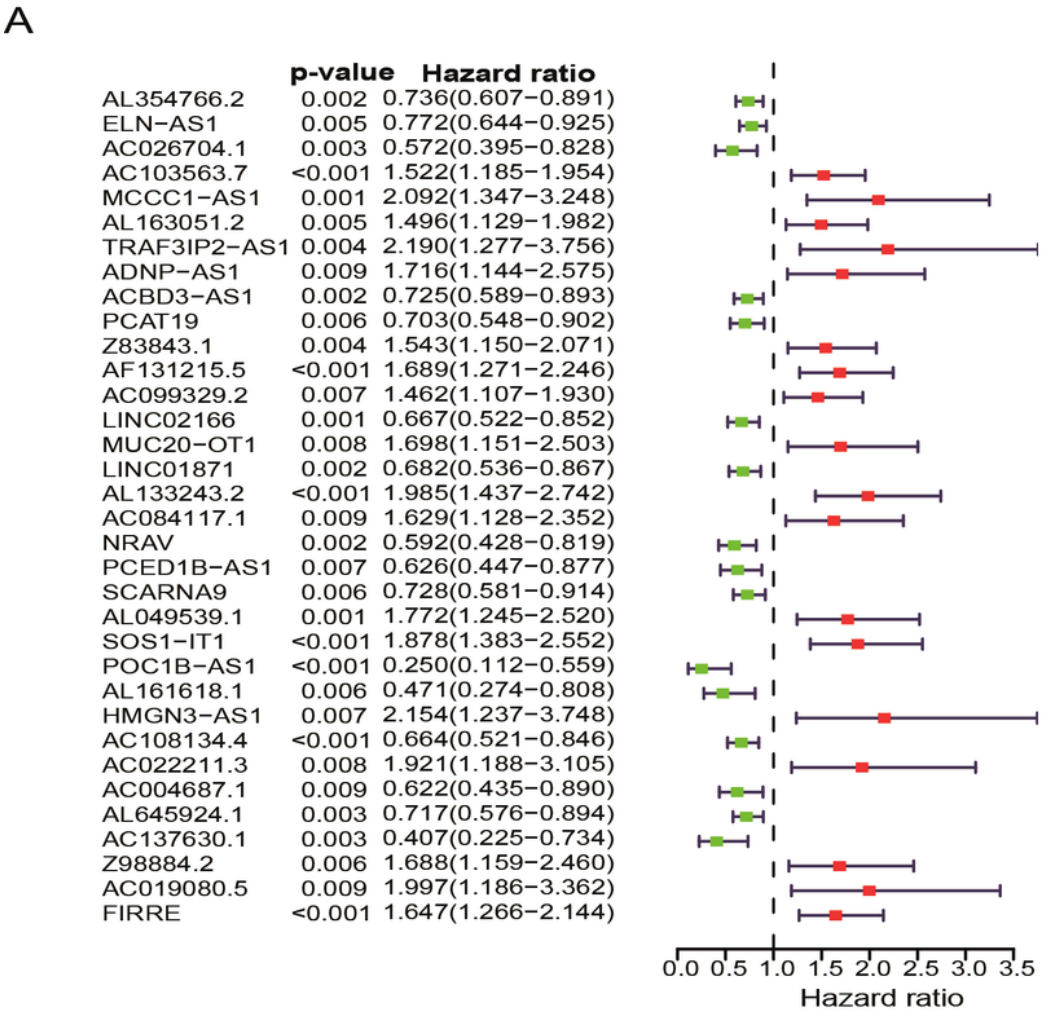

Figure 1

Immune-related lncRNAs associated with prognosis. (A) Forest plot of univariate Cox regression analysis. The p-value, hazard ratio, and 95% confidence intervals of the immune-related genes are shown. Red and green indicate a risk-associated (HR > 1) and a protective (HR < 1) lncRNA, respectively. (B) Immune-

related lncRNAs-immune-related genes network. Green and red represent immune-related genes and 12 immune-related lncRNAs, respectively.

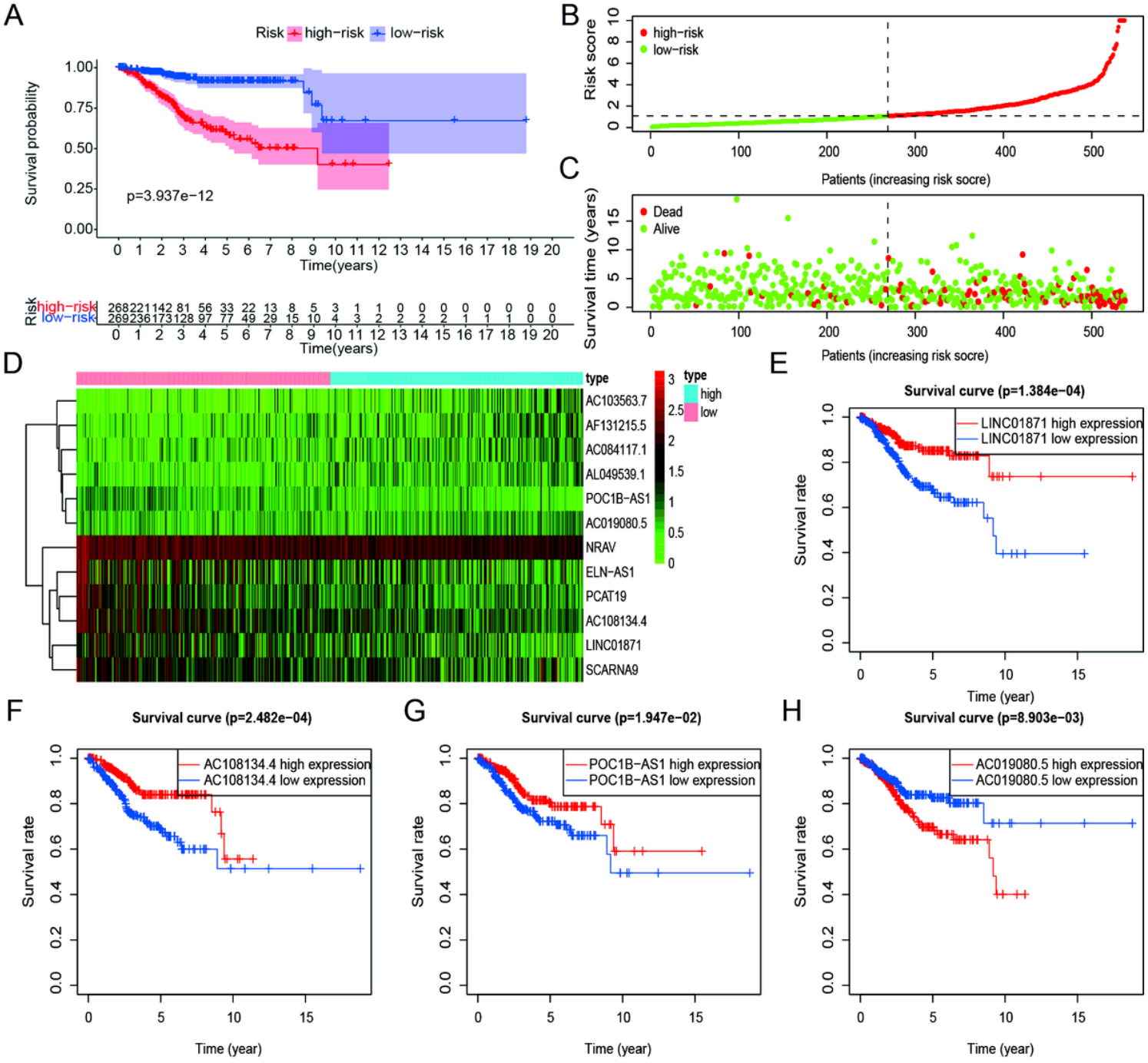

Figure 2

Survival analysis and risk curves. (A) Survival curve of endometrial cancer. Red and blue indicate high-risk and low-risk groups, respectively. (B) Risk curve of endometrial cancer. Red and green indicate high-risk and low-risk groups, respectively. (C) Scatter plot of different survival status of endometrial cancer patients. Red and green dots denote patients that are dead or alive, respectively. (D) Hierarchical clustering of 12 immune-related lncRNAs. Differences in the expression level of 12 immune-related

lncRNAs in different risk status. (E-H) Survival curves of immune-related lncRNAs. Blue and red represent low and high expression groups, respectively.

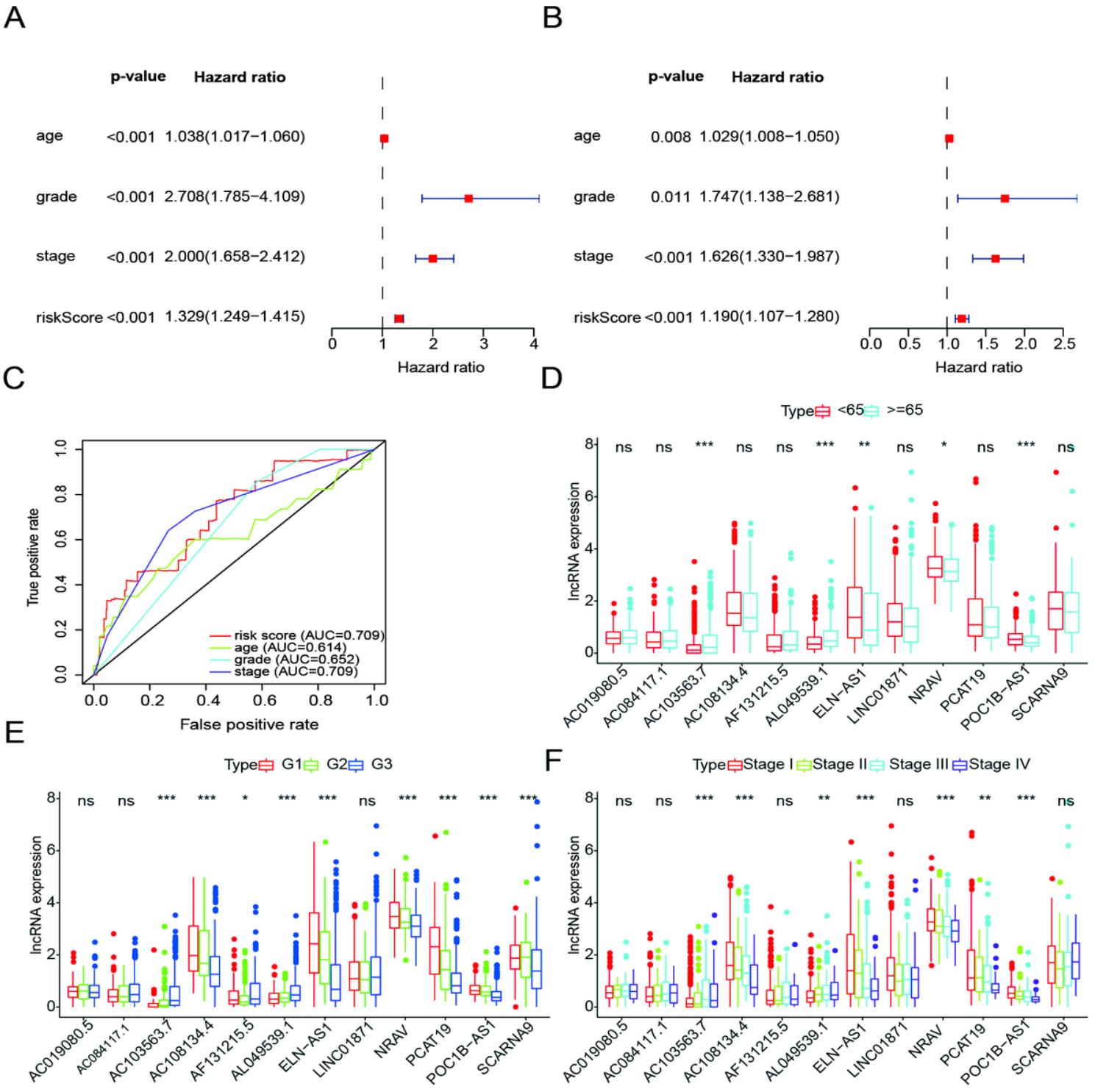

Figure 3

Independent prognostic analysis and clinical correlation analysis. (A) Forest plot of univariate independent prognostic analysis. The p-value, hazard ratios, and 95% confidence intervals are shown. Red and green indicate a risk-related (HR > 1) and a protective (HR < 1) factor, respectively. (B) Forest plot of multivariate independent prognostic analysis. The p-value, hazard ratios, and 95% confidence intervals

are shown. Red and green indicate a risk-related (HR > 1) and a protective (HR < 1) factor, respectively. (C) Plot of multi-index ROC curve showing risk score (AUC = 0.709), age (AUC = 0.614), grade (AUC = 0.652), and stage (AUC = 0.709). (D) lncRNA expression levels in patients aged below (red) or above (blue) 65. (E) lncRNA expression levels in different pathological grades. (F) lncRNA expression levels in different FIGO stages. *p < 0.05, **p < 0.01, ***p < 0.001. ns, p > 0.05.

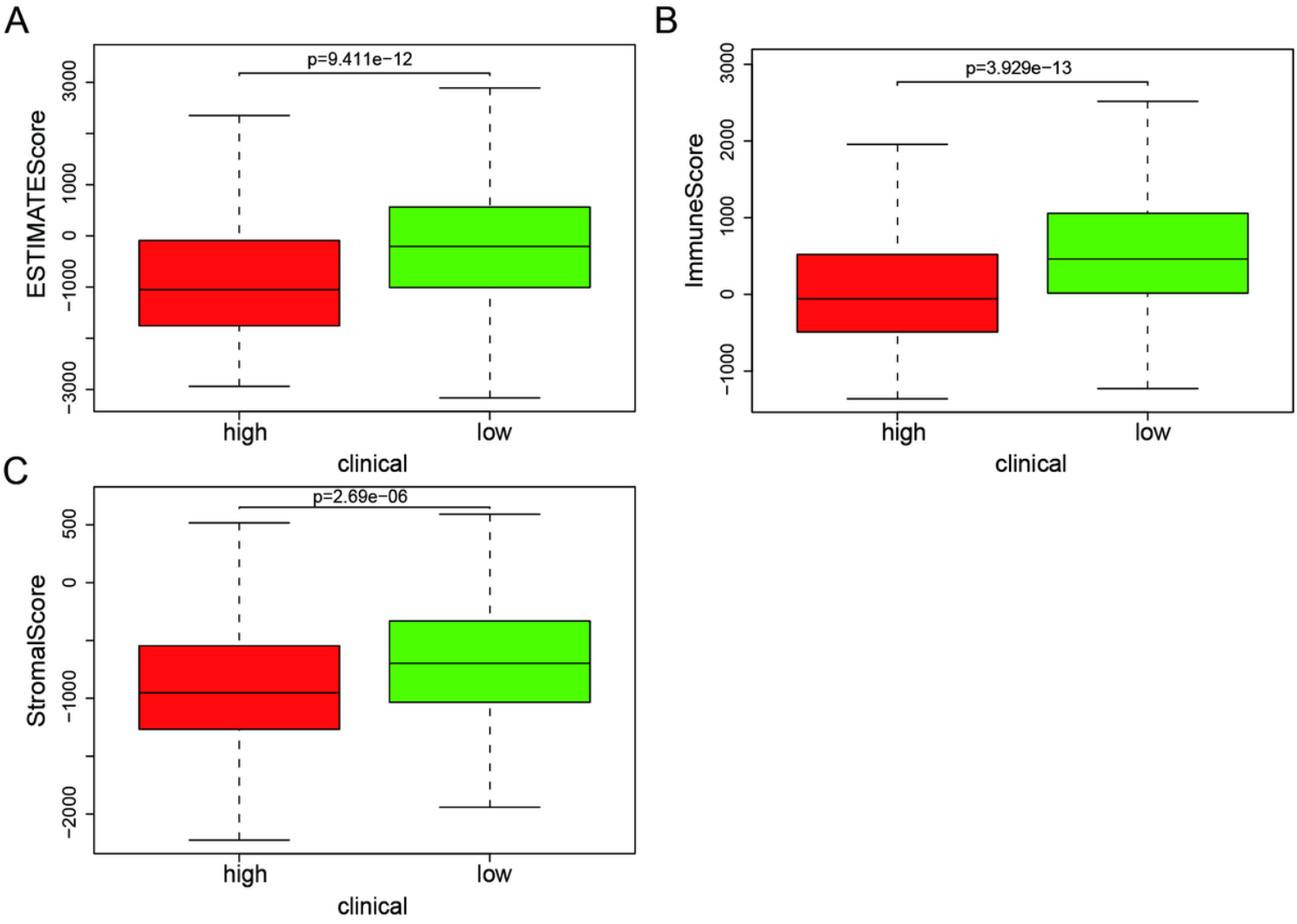

Figure 4

Immune scores and stromal scores. (A) Distribution of ESTIMATE scores of different risk status. Box-plot shows that there is a significant difference between high-risk and low-risk groups (p < 0.05). (B) Distribution of immune scores of different risk status. Box-plot shows that there is a significant difference between high-risk and low-risk groups (p < 0.05). (C) Distribution of stromal scores of different risk status. Box-plot shows that there is a significant difference between high-risk and low-risk groups (p < 0.05).

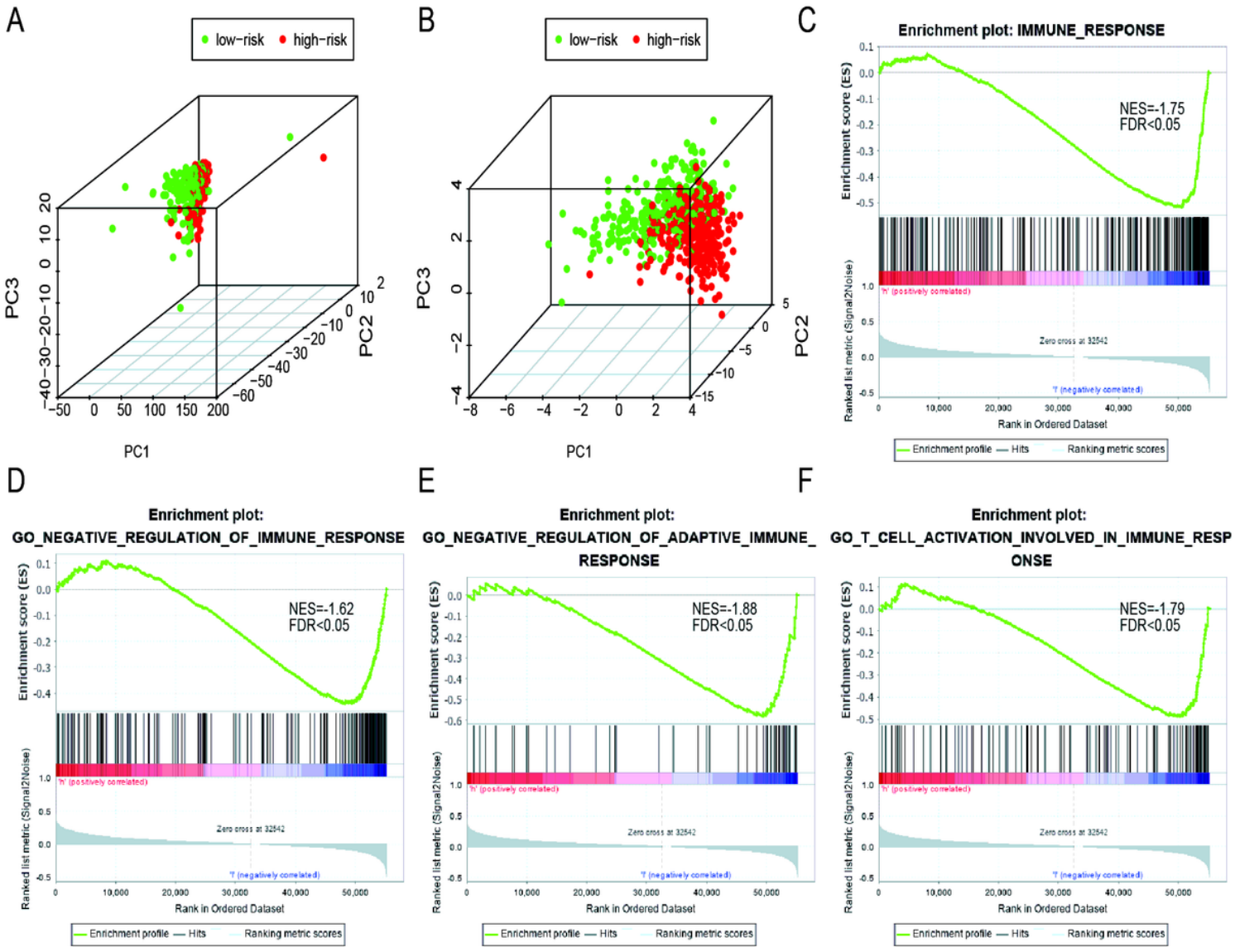

Figure 5

Principal component and gene set enrichment analyses. (A) Principal components analysis between low-risk and high-risk groups based on all immune-related lncRNAs. Red and green indicate high-risk and low-risk groups, respectively (B) Principal components analysis between low-risk and high-risk groups based on 12 immune-related lncRNAs associated with prognosis. Red and green indicate high-risk and low-risk groups, respectively (C-F) GSEA indicate significant enrichment in immune-related phenotypes in the low-risk patients.

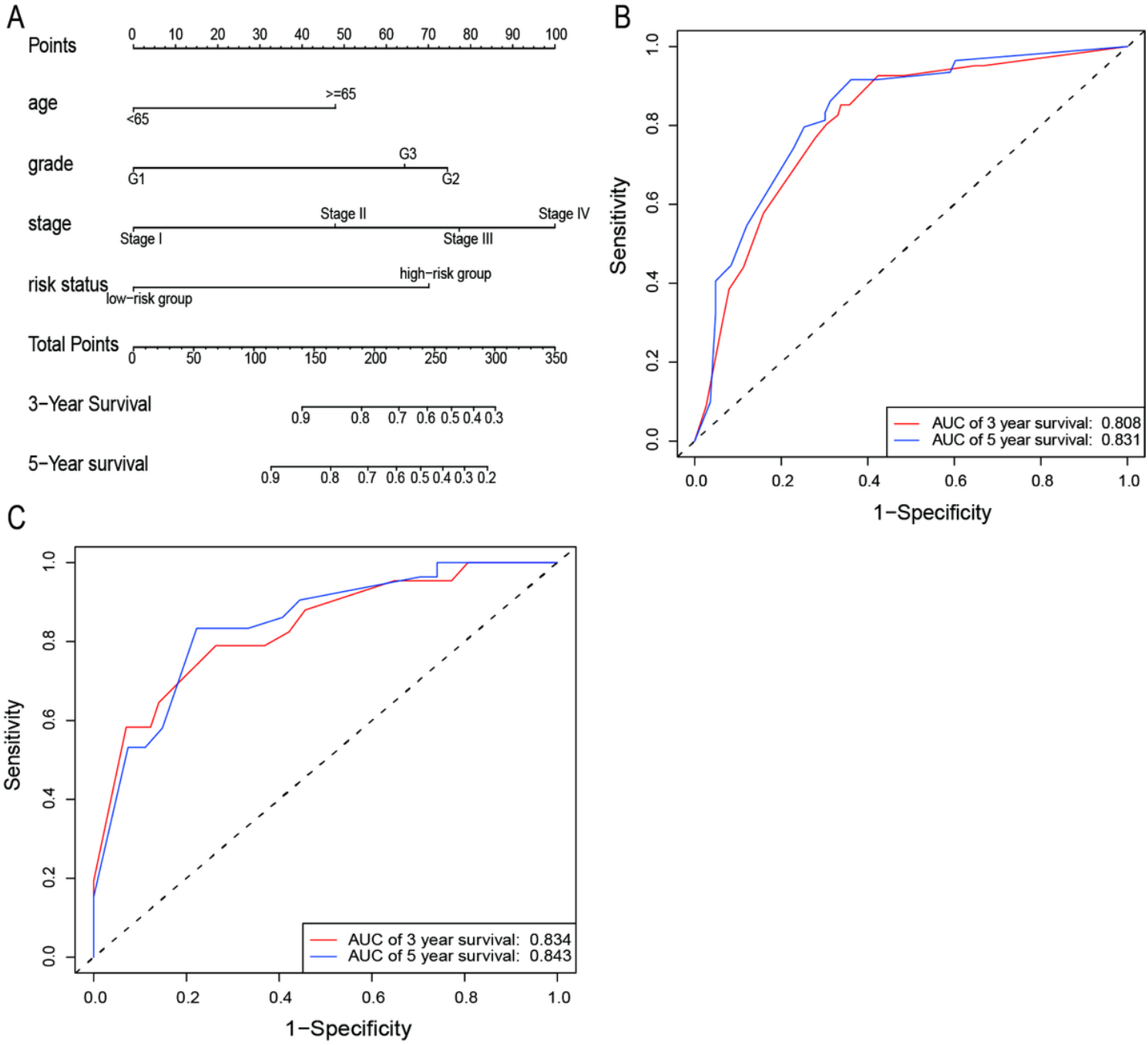

Figure 6

Prognostic nomogram for overall survival rate. (A) Survival nomogram. (to use the nomogram, an individual patient's value is located on each variable axis, and a line is drawn upward to determine the number of points received for each variable value. The sum of these numbers is located on the Total Points axis, and a line is drawn downward to the survival axes to determine the likelihood of 3- or 5-year survival). (B) Multi-index ROC curve of training samples. Red and blue indicate 3-year and 5-year survival, respectively. (C) Multi-index ROC curve of validation samples. Red and blue indicate 3-year and 5-year survival, respectively.

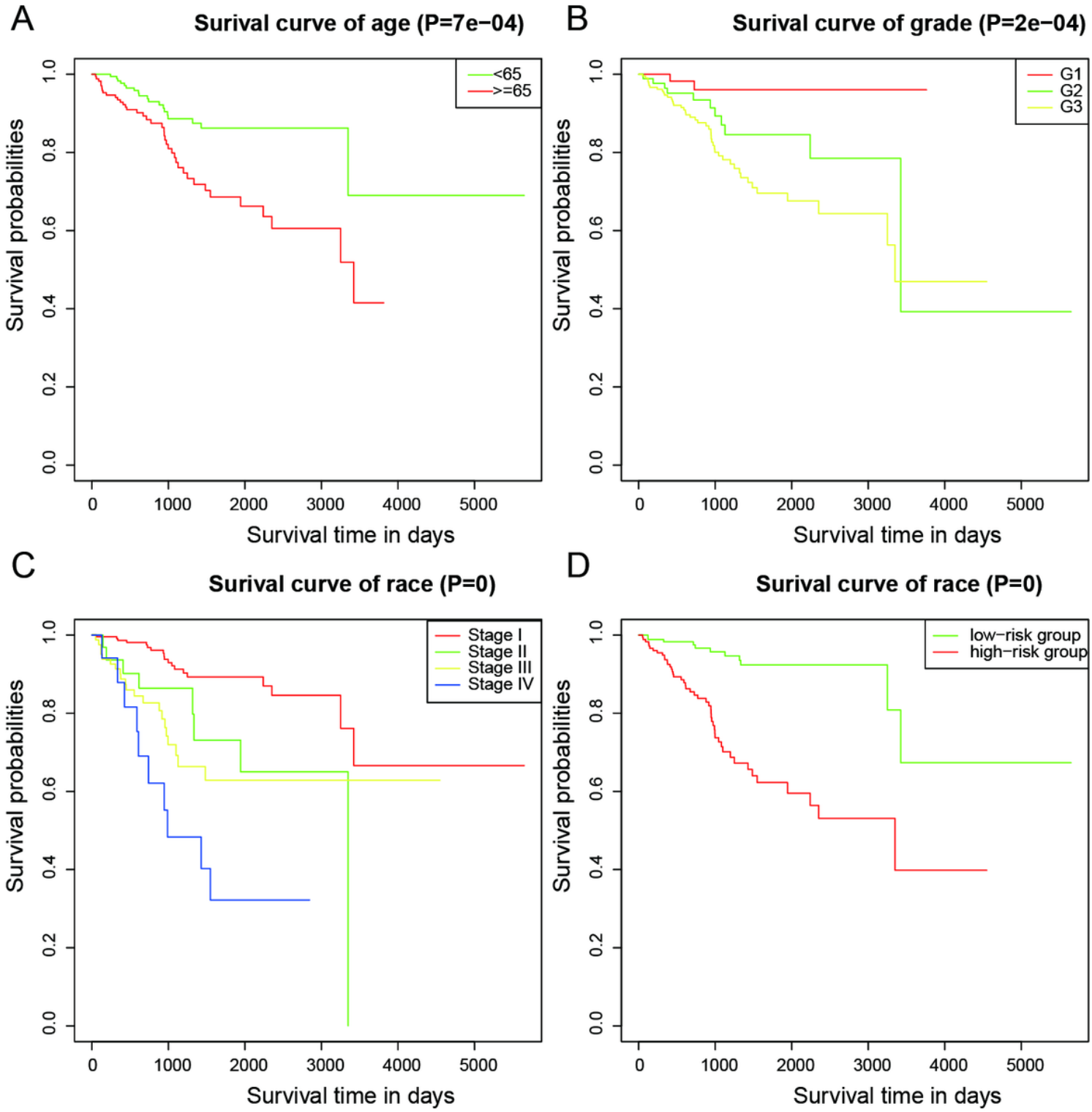

Figure 7

Survival analysis of independent prognostic risk factors. (A) Survival curve of patients younger or older than 65. (B) Survival curve of groups in Grade 1, Grade 2, and Grade 3. (C) Survival curve of groups in stage Ⅰ, stage Ⅱ, stage Ⅲ, and stage Ⅳ. (D) Survival curve of patients in the high-risk and low-risk groups divided based on the 12 immune-related lncRNAs.

Supplementary Files

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

- SupplementaryData.docx