Immune-related Long Noncoding RNA Signature for Patients with Cervical Cancer

Jianfeng Zheng
Nanjing Medical University
Jinyi Tong (tongjinyi252@zju.edu.cn)
Hangzhou First People's Hospital https://orcid.org/0000-0003-2576-8764
Benben Cao
Zhejiang Chinese Medical University
Xia Zhang
Nanjing Medical University
Zheng Niu
Hangzhou First People's Hospital

Research article

Keywords: immune-related lncRNAs, cervical cancer, signature, nomogram, infiltrating immune cells, GSEA, ceRNA network

Posted Date: August 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-49217/v1

Abstract

Background: Cervical cancer (CC) is a common gynecological malignancy for which prognostic and therapeutic biomarkers are urgently needed. The signature based on immune-related lncRNAs(IRLs) of CC has never been reported. This study aimed to establish an IRL signature for patients with CC.

Methods: The RNA-seq dataset was obtained from the TCGA, GEO, and GTEx database. The immune scores（IS）based on single-sample gene set enrichment analysis (ssGSEA) were calculated to identify the IRLs, which were then analyzed using univariate Cox regression analysis to identify significant prognostic IRLs.

A risk score model was established to divide patients into low-risk and high-risk groups based on the median risk score of these IRLs. This was then validated by splitting TCGA dataset(n=304) into a training-set(n=152) and a valid-set(n=152). The fraction of 22 immune cell subpopulations was evaluated in each sample to identify the differences between low-risk and high-risk groups. Additionally, a ceRNA network associated with the IRLs was constructed.

Results: A cohort of 326 CC and 21 normal tissue samples with corresponding clinical information was included in this study. Twenty-eight IRLs were collected according to the Pearson's correlation analysis between immune score and lncRNA expression ($P < 0.01$). Four IRLs (BZRAP1-AS1, EMX2OS, ZNF667-AS1, and CTC-429P9.1) with the most significant prognostic values ($P < 0.05$) were identified which demonstrated an ability to stratify patients into low-risk and high-risk groups by developing a risk score model. It was observed that patients in the low-risk group showed longer overall survival (OS) than those in the high-risk group in the training-set, valid-set, and total-set. The area under the curve (AUC) of the receiver operating characteristic curve (ROC curve) for the four IRLs signature in predicting the one-, two-, and three-year survival rates were larger than 0.65. In addition, the low-risk and high-risk groups displayed different immune statuses in GSEA. These IRLs were also significantly correlated with immune cell infiltration.

Conclusions: Our results showed that the IRL signature had a prognostic value for CC. Meanwhile, the specific mechanisms of the four-IRLs in the development of CC were ascertained preliminarily.

Introduction

Cervical cancer(CC)is a malignant gynecologic tumor threatening the health of women. The morbidity and mortality for CC ranks fourth worldwide among women[1]. Infection with high-risk human papillomavirus (HPV), especially HPV16 and HPV18, is the main etiologic risk factor for CC, and plays an important role in diagnostic tests [2]. Surgery is the main treatment method for CC in early stages while advanced-stage CC can be treated with radiotherapy, chemotherapy, or concurrent chemoradiation, thereby improving the survival rate of CC patients [3]. However, a considerable number of CC patients have poor prognosis due to metastasis or recurrence within two years after treatment[4]. Hence, effective

prevention to reduce morbidity and individual treatments to improve the prognosis of CC are important for obstetricians and gynecologists.

The immune system can recognize tumor antigens expressed on the surface of tumor cells. It generates an immune response via the activation of effector cells and triggers the release of a series of effector molecules to attack and eliminate tumor cells and to inhibit tumor growth[5]. The immune imbalance in tumor microenvironment plays an important role in the occurrence and development of cancer. With the development of cellular molecular biology and immunology, immunotherapy has become a new treatment approach for cervical cancer[6]. Tumor immunotherapy acts mainly by increasing the immunogenicity of tumor cells and the effect of cell damage sensitivity, stimulate and enhance the anti-tumor immune response, with the aid of biological agents, doping effects of immune cells and molecules into the body, together, the body's immune system can not only kill cancer cells remaining small, still can prevent tumor metastasis and recurrence[7].

Long noncoding RNAs (lncRNAs), defined arbitrarily as transcripts lacking protein-coding potential, are RNAs with more than 200 nucleotides[8]. LncRNAs are not only associated with the invasion, migration, and proliferation of CC, but are also involved in autophagy and epithelial-mesenchymal transition (EMT) [9]. Emerging studies have demonstrated that lncRNAs also modulate the immune response to tumors. For instance, the down-regulated lncRNA AGER has been demonstrated to be closely related to T-cell status in lung cancer [10]. lncRNA GM16343 is regulated by Interleukin 36β to strengthen the anti-tumor immune response of $CD8^+$ T-cells[11]. lncRNA LINK-A has been demonstrated to downregulate antigen presentation and intrinsic suppression in triple-negative breast cancer[12]. LncRNAs can also affect the development of cancer by regulating NK cells. LncRNA GAS5 can inhibit tumor growth, and the overexpression of GAS5 can regulate Mir-544/RUNX3 to enhance the killing effect of NK cells in liver cancer[13]. LncRNAs can also modulate tumor immunity by regulating Treg cells. The expression of Mir-448 can be increased by interfering with lncRNA SNHG1, which down-regulates the expression of indoleamine 2,3-dioxygenase (IDO), inhibits the differentiation of Treg cells, and reduces the immune escape of breast cancer cells[14]. EGFR is an important member of the tyrosine kinase receptor family. By binding to EGFR specifically, LNC-EGFR promotes Treg cell differentiation and promotes the immune escape of liver cancer cells[15]. LncRNAs can affect the tumor microenvironment and thus play an important role in immunotherapy. Nonetheless, the effect research on immune-related lncRNAs in CC is rarely reported.

The purpose of our research was to identify an IRL signature, which might serve as prognostic and therapeutic biomarkers in CC. We developed a prognostic signature and mechanisms of IRLs using single-sample gene set enrichment analysis (ssGSEA), survival analysis, a Cox regression risk model, gene set enrichment analysis (GSEA), ceRNA network, and other analysis methods.

Methods

Datasets and preprocessing

The RNA sequencing profiles associated with CC were obtained from The Cancer Genome Atlas (TCGA) [16] (https://toil.xenahubs.net), which consisted of 306 CC and 13 normal tissue samples. The low-expression genes were filtered, and the genes whose expression level was greater than 0 in more than a third of the samples were retained. Additionally, the RNAs were identified as mRNAs or lncRNAs based on their annotation information in the GENCODE database[17](https://www.gencodegenes.org/). GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array platform was used to obtain the microarray dataset GSE6791 from gene expression omnibus (GEO) repository (GeneExpressionOmnibus,GEO□http://www.ncbi.nlm.nih.gov/geo/) [18], of which twenty CC and eight normal tissue samples were included. All 326 CC samples and 21 normal tissue samples were contained corresponding clinical information. For the RNA sequencing profiles obtained from TCGA TARGET GTEx, empirical Bayes and linear regression along with Benjamini and Hochberg multiple comparison methods from limma package[19](Version 3.10.3, http://www.bioconductor.org/packages/2.9/bioc/html/limma.html□ were performed to gain adjusted-P-value and |logFC| (adj.P.Value□0.05, |logFC|□0.585). For GSE6791, GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) was used with the seqmap[20] tool to map probes to mRNA and lncRNA sequences. The immune-related genes (IRGs) were downloaded from the InnateDB database (http://www.innatedb.com)[21]. Focusing on screening genes that were up- and down-regulated consistently, we used venn analysis to select the intersection genes of the aforementioned datasets. Single-sample gene set enrichment analysis (ssGSEA)[22] was used to identify the immune scores(IS□of each sample. Pearson's correlation coefficient between lncRNAs and IS was calculated for each corresponding samples to identify immune-related lncRNAs□IRLs□(P□0.01).

Signature development of IRLs

Univariate Cox regression analysis with hazard ratio (HR) was gained from overall survival (OS) and OS time from the candidate IRLs. HR□1 indicated that expression was higher as well as the risk and the survival rate was lower. LncRNAs that are up-regulated in the tumor should, in principle, have an HR greater than 1. The Kaplan–Meier analysis was generated by survminer(version 0.4.3)in R package based on the expression value, survival time and survival status to determine the optimal cut point. The log-rank test was performed based on survival□version 2.42-6□in R package to sort the IRLs with a significant prognostic value ($P < 0.05$), then the survival curves were drawn. Multivariate Cox regression analysis was performed based on the expression value, OS, and OS time of IRLs in each sample. Subsequently the individual prognostic risk model for corresponding samples was established. $Expr_{gene}$ indicated the expression of corresponding IRL for each sample. All the samples were divided into low-risk and high-risk according to the median of risk scores the following study.All the samples in TCGA were regarded as a total-set. Training- and valid-sets were constructed by dividing the total-set equally into two parts to validate the risk score formula. The Kaplan-Meier survival curves of all three sets were drawn to determine the prognostic difference between the risk groups. The one-year, two-year and three-year survival receiver operating characteristic (ROC) curves predicted by the risk model were drawn.

Clinical and pathological characteristics of the risk score model

Clinical and pathological characteristics, including age, pathologic M, pathologic N, pathologic T, clinical stage, neoplasm histologic grade, neoplasm cancer status, primary therapy outcome success, radiation therapy, tobacco smoking history, along with the risk score, the immune score and the expression of IRLs in each sample were included in the analysis for illustration of a heat map. We observed the differences in clinical pathological characteristics in risk groups or the risk score of clinical pathological characteristics using the Student's t-test. A scatter plot was drawn using GraphPad Prism 5[23] based on significant clinicopathological characteristics.

Nomogram model construction and visualization

Univariate Cox regression analyses were performed respectively based on the risk groups and clinicopathological factors(age, pathologic-M, pathologic-N, pathologic-T, clinical-stage, neoplasm-histologic-grade, radiation-therapy and tobacco smoking- history). Multivariate Cox regression analysis was conducted on the factors with a p□0.05. Using the IRL signature along with the factors with a p□0.05, a nomogram was constructed to make the multivariate Cox regression analysis more clear.

Gene set enrichment analysis

To elucidate the biological differences between risk groups, a gene set enrichment analysis (GSEA)[24] of “c5.bp.v7.1.symbols.gmt” background was carried out using GSEA(version 4.0.3) software. A nominal p value < 0 .05(NOM p-val < 0.05) was considered significant. We focued on selecting immune-related terms for display.

Evaluation of infiltrating immune cells

To further observe the differences in the abundance of infiltrating immune cells of the risk groups, we used the CIBERSORT algorithm, a deconvolution method [25], coupled with LM22 that distinguished 22 immune cell subpopulations from Cibersort (a web server) and a heat map for all the samples was drawn using ggplot2(Version: 3.2.1). The Student's t-test was applied to find significant immune cell subpopulations in risk groups to chart Violin Plot by boxplot□version 0.3.2□in R package.

Construction of an IRL-associated ceRNA network

The ceRNA network was constructed to explore the association among IRLs, miRNAs, and mRNAs based on the ceRNA hypothesis[26]. First, the Pearson correlation coefficient analysis between the IRLs and

immune-related DEGs was performed to obtain lncRNA-mRNA pairs(r > 0 and adj.p.value < 0.01). The target miRNAs of IRLs were predicted using DIANA-LncBase v2[27] (http://carolina.imis.athenainnovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-experimental). The mRNAs targeted by miRNAs were predicted by integrating miRWalk2.0[28] (version:3.4.0,http://chianti.ucsd.edu/cytoscape-3.4.0/), miRanda, RNA22, and Targetscan databases. Only miRNA-mRNA pairs recognized by all four databases were considered as candidate targets. Subsequently, the ceRNA network based on the same miRNA of lncRNA-miRNA and miRNA-mRNA pairs was established and visualized using Cytoscape (version:3.4.0□http://chianti.ucsd.edu/cytoscape-3.4.0/) [29].Sorting the mRNA with a significant prognostic value ((P < 0.05) in the ceRNA network, the Kaplan–Meier analysis was generated using survminer (version 0.4.3) to determine the optimal cut point, and the log-rank test was performed based on survival (version 2.42-6) in R package.

Results

Differential analysis of genes

In the aggregate, 1995 differentially expressed lncRNAs, with corresponding clinical information were extracted from TCGA and GSE6791 databases mentioned above. Of these, 567 lncRNAs were highly expressed and 1428 lncRNAs were expressed at low levels in the CC samples (Table 1). A total of 64 lncRNAs were obtained from the intersection of the two databases (Fig. 1A), among which 10 lncRNAs were consistently up-regulated and 54 lncRNAs were consistently down-regulated in the two databases. A total of 1040 IRGs were identified from the InnateDB. The ssGSEA analysis was based on 201 IRGs obtained from the intersection of TCGA and InnateDB databases (Fig. 1B) to compute IS. Finally, a cohort of 28-lRLs was obtained based on the Pearson correlation coefficient analysis between the 64 lncRNAs and IS of 201 IRGs(P□0.01).

Table 1
Differential mRNAs and lncRNAs

		mRNA	lncRNA
TCGA	up	3305	472
	down	4163	733
	total	7468	1205
GSE6791	up	4680	95
	down	3133	695
	total	7813	790

Univariate Cox regression and Kaplan–Meier survival analysis

Univariate Cox regression and Kaplan–Meier survival analysis were performed for the cohort of 28 IRLs. Four IRLs (BZRAP1-AS1, EMX2OS, ZNF667-AS1, and CTC-429P9.1) with low expression were found to be in accordance with our expectation and were then included in the signature development. The results are shown in Table 2. The Kaplan–Meier survival analysis revealed that four IRLs were related to OS in CC significantly (p 0.05) (Fig. 2). The four IRLs were defined as protective factors due to their HR value 1, which showed that the high expression of the four IRLs was associated with lower OS.

Table 2
K-M analysis, Univariate Cox regression analysis and Pearson correlation analysis of 4-IRLs

	K-M analysis	Univariate Cox regression analysis				Pearson correlation analysis		differential expression analysis
lncRNA	p.value	HR	lower.95	upper.95	p.value	r	p.adj.value	up_down
BZRAP1-AS1	0.0002	0.82	0.72	0.93	0.00	0.32	2.84E-08	down_lnc
EMX2OS	0.0010	0.91	0.82	1.00	0.05	-0.25	3.01E-05	down_lnc
ZNF667-AS1	0.0019	0.87	0.77	0.99	0.03	-0.17	0.007	down_lnc
CTC-429P9.1	0.0159	0.93	0.82	1.06	0.27	-0.24	5.15E-05	down_lnc

Construction of the prognostic risk model and validation using four IRLs

A total of 304 TCGA samples were regarded as the total-set so that there were both 152 samples in the training-set and the valid-set as described in the methods. The regression coefficient β was first generated from the training-set ($\beta_{ZNF667-AS1}$= -0.152565, β_{EMX2OS}=-0.019887, $\beta_{BZRAP1-AS1}$=-0.17831, $\beta_{CTC-429P9.1}$ = 0.0096755, Table 3) to establish the risk score(RS) formula. Hence, the prognostic risk model for corresponding samples was established using the following formula:

Table 3
β of each 4-IRL

Ensemble_ID	lncRNA	β
ENSG00000166770.10	ZNF667-AS1	-0.152565
ENSG00000229847.8	EMX2OS	-0.019887
ENSG00000265148.5	BZRAP1-AS1	-0.17831
ENSG00000269427.1	CTC-429P9.1	0.0096755

risk score (RS) = $expr_{BZRAP1-AS1}$*(-0.152565) + $expr_{EMX2OS}$*(-0.019887) + $expr_{BZRAP1-AS1}$*(-0.17831) + $expr_{CTC-429P9.1}$*0.0096755

$Expr_{gene}$ indicated the expression value of corresponding IRL for each sample. An RS higher than the median was identified as a high-risk group while an RS lower than the medium was identified as a low-risk group. Using this approach, a risk model signature based on four IRLs was constructed, which was further validated in the total-set and valid-set using the same β to confirm the prediction potential of the 4-IRL signature. The Kaplan-Meier survival curves based on logRank test revealed that OS in the low-risk group was markedly longer than that in the other group in all three sets ($P_{Training-set}$ =0.0068, $P_{Valid-set}$ =0.02, $P_{Total-set}$ =0.0015, Fig. 3A, 3B, 3C). The one-year, two-year and three-year survival ROC curves predicted by the risk model indicted that the AUCs were larger than 0.65 (0.695, 0.66, 0.676, Fig. 3D),thus predicting that the risk score model could efficiently forecast over 65% of OS for CC patients. Therefore, the risk model signature based on four IRLs was accurate in predicting OS of CC patients.

Clinicopathological characteristics of the risk score model

A heat map was constructed by combining the expression values of four IRLs and their clinicopathological characteristics (Fig. 4A). The higher the RS, the lower the IS. RS and IS values showed a significant negative correlation based on the scatterplot of the correlation coefficient (r=-0.14, p = 0.01631, Fig. 4B). The scatterplot for the distribution of IS in risk groups showed that the IS of the high-risk group was significantly lower than that of the low-risk group (Fig. 4C). Furthermore, the high expression of the four IRLs could be seen with low RS, which suggested that the up-regulation of the four IRLs were associated with better prognosis. In contrast, low expression of the four IRLs had the opposite consequence. An RS contrast of neoplasm histologic grade showed that the RS in stages IIB-III-IV was significantly greater than that in stages I-II-IIA (Fig. 4C). This part of the result suggests that a high-risk score has an adverse effect on prognosis, which may be caused by a decrease in immune score.

Nomogram model construction and visualization

The univariate and multivariate Cox regression analyses of clinicopathological characteristics and the four-IRL signature for total TCGA dataset demonstrated that age, AJCC stage along with the four-IRL signature were all independent prognostic factors (P□0.05,Table 4). Nonetheless, neoplasm histologic grade, radiation therapy, and a history of tobacco-smoking were not correlated with prognosis independently. To better predict prognosis at one-, three-, and five-year OS of CC patients, we constructed a new nomogram using variables such as age, AJCC stage, and the four-IRL signature (Fig. 5).

Table 4
Univariate and multivariate cox regression analyses of clinical parameters

	Univariate		**Multivariate**	
Variable	HR	p.value	HR	p.value
Age,years				
>=60 vs < 60	1.86253	0.01696	3.689112	0.048533
Grade				
III-IV vs I-II	0.889401	0.661005		
T stage				
T3-T4 vs T1-T2	3.64268	0.000074	3.869304	0.060251
N stage				
N1 vs N0	2.694785	0.0046	2.319201	0.132728
M stage				
M1 vs M0	3.64563	0.020534	2.24E-07	0.998132
AJCC stage				
IIB-III-IV vs I-II-IIA	1.861189	0.009193	0.200615	0.057357
Radiotherapy				
Yes vs no	1.091559	0.828872		
tobacco_smoking_history				
> 1 vs = 1	1.469726	0.12418		
Four-lncRNA signature				
High risk vs Low risk	2.137914	0.001933	3.047015	0.032286

Gene set enrichment analysis

GSEA analysis of the two risk groups was carried out to predict enrichment status disparities of molecular mechanism functions. The enrichment analysis showed that 118 biological functions were markedly enriched in the low-risk group, whereas only one biological function(GO_ATP_SYNTHESIS_COUPLED_ELECTRON_TRANSPORT)was enriched in the high-risk group. In the enrichment status enriched in the low-risk group, we sought out couple immune-related responses as shown in Fig. 6.

Evaluation of infiltrating immune cells

Previous studies have shown that infiltrating immune cells are closely related to the prognosis and treatment of malignant tumors[29]. From the GSEA analysis in our study, we discovered that the four-lRL signature was associated with many immune characteristics. Hence, the abundance of twenty-two infiltrating immune cells (Fig. 7A) were estimated which showed that the abundance of nine infiltrating immune cells (B cells naïve, B cells memory, T cells CD4 memory resting, NK cells resting, macrophages M1, dendritic cells activated, mast cells resting, mast cells activated and neutrophils) was significantly different (P□0.05) between the risk groups (Fig. 7B). The abundance of four infiltrating immune cells (B cells naïve, T cells CD4 memory resting, macrophages M1 and mast cells resting) in the low-risk group were significantly higher than that in the high-risk group as shown by the Student's t-test(P□0.05).

Construction of an lRL-associated ceRNA network

Four-lRLs,forty-five hub mRNAs□and thirty-eight miRNAs were involved in ceRNA network (Fig. 8). A total of 46 lncRNA-miRNA pairs, 232 miRNA-mRNA pairs, and 55 lncRNA-mRNA pairs were identified. The downregulation of 12 mRNAs in the ceRNA network was significantly related to OS in CC. Among them, the downregulation of seven mRNAs (CXCL12, FREM1, IGF1, IRF4, NFATC2, NTN1, and STAT6) had an adverse effect on prognosis, which was in line with what was expectated (Fig. 9).

Discussion

In this study, a cohort of 326 CC and 21 normal tissue samples from two datasets (TCGA, GSE6791) were included to identify the differential lncRNAs for patients with CC. A total of 1040 IRGs were collected from the InnateDB. Four lRLs (BZRAP1-AS1, EMX2OS, ZNF667-AS1, and CTC-429P9.1) were identified after ssGSEA analysis, the Pearson correlation coefficient analysis, univariate Cox regression analysis, and Kaplan–Meier survival analysis between the lncRNAs and IS. The prognostic risk model based on the

four IRLs could divide CC patients into two risk groups according to the median RS, which was validated by dividing the total samples equally. The univariate and multivariate cox regression analyses of clinicopathological characteristics showed that age, AJCC stage, and the four-IRL signature were all independent prognostic factors. A nomogram was constructed based on age, AJCC stage, and the four-IRL signature to predict OS for CC patients more clearly. We also found that RS and IS showed significant negative correlation, which indicated that high RS had an adverse effect on prognosis due to a decrease in IS.

In our study, the four IRLs (BZRAP1-AS1, EMX2OS, ZNF667-AS1, and CTC-429P9.1) were identified to be protective against CC; only ZNF667-AS1 had been previously reported in CC. In the existing literature, ZNF667-AS1 with low expression has been identified to be negatively correlated with the OS, tumor size, and FIGO stage in CC[30]. Additionally, ZNF667-AS1 can competitively bind to miR-93-3p, which targets PEG3, to regulate the progression of CC[31]. Recent research has shown that inhibiting PEG3 would promote the immune escape of cancer cells[32]. BZRAP1-AS1 was found to be a novel biomarker associated with prostate cancer(PC), being down-regulated in PC samples [33]. It was shown, however, to be highly expressed in hepatocellular carcinoma and inhibited the transcription of THBS1 by recruiting DNMT3b to its promoter region[34]. Previous research has revealed that the downregulation of EMX2OS in classical papillary thyroid cancer might independently predict shorter recurrence-free survival [35], while the overexpression of EMX2OS in ovarian cancer and EMX2OS/miR-654/AKT3 axis may target PD-L1(programmed cell death protein 1) to suppress the initiation and progression of cancer [36]. Accumulating evidences have suggested that Thrombospondin-1 (THBS1) may affect tumor immunity[37]. PD1-PDL1(PD1 ligand) has already been shown to be an important immune checkpoint pathway, which can be used by cancer cells to evade immune attacks[38]. Thus, BZRAP1-AS1 and EMX2OS may play a dual role in cancer and directly or indirectly regulate tumor immunology. By contrast, no reports concerning CTC-429P9.1 in cancer have been published, and therefore, the role of CTC-429P9.1 remains unclear.

Our GSEA analysis showed that certain immune-related enrichment statuses were dramatically enriched in low-risk group. Non-canonical WNT signaling has been demonstrated to be closely associated with cancer stem cell survival, bulk-tumor expansion and invasion/metastasis, which have the potential for tumor immunological[39]. Ubiquitination has been shown to be crucial for tumor immunity[40].

The ceRNA network was established to explore specific mechanisms in the development of CC based on the competing endogenous RNAs theory. Combining the mRNAs of the most outstanding prognosis with the connectivity of miRNAs, the following ceRNA relationships have been established, which should be used as a follow-up verification: EMX2OS/hsa-miR-3153/CXCL12, EMX2OS/hsa-miR-3928-3p/FREM1, BZRAP1-AS1/hsa-miR-30b-3p/IGF1, EMX2OS/hsa-miR-92a-2-5p/IRF4, BZRAP1-AS1/hsa-miR-30b-3p/NFATC2, BZRAP1-AS1/hsa-miR-30b-3p/NTN1, BZRAP1-AS1/hsa-miR-541-3p/STAT6. CXCL12 binds primarily to CXCR4 which can cause a mass of signal routing including the immunity[41]. FREM1 has been authenticated as an immune-related gene which may be a potential target for immunotherapy in breast cancer and clear cell renal cell carcinoma[42, 43]. The insulin-like growth factor 1(IGF1) pathway

has been proven to contribute to the suppression of immune tumor microenvironment(TME) in gynecologic cancers[44]. IRF4^{+} Tregs have been confirmed to be correlated with poor prognosis in patients with multiple cancers[45]. NFATC2 regulates IL-21 expression in human CD4 + CD45RO + T lymphocytes[46].STAT6 is a factor converge on intracellular determinants of cell functions and drives the recruitment and polarization of tumor-associated macrophages (TAMs)[47]. Thus, our IRLs need to be further investigated.

Compared to a certain number of studies carried out to explore IRL signatures in several human malignancies [48–54], this was the first study based on a comprehensive analysis that focused on IRL signature in CC. However, our study has some limitations. Additional data sets are needed to validate our study. The specific molecular mechanisms in which these genes may be involved should be verified through in vitro and in vivo experiments.

Conclusions

We established a four-IRL-based signature with a prognostic value for CC, which could stratify CC patients into low- and high-risk groups. Meanwhile, the specific mechanisms of the four-IRLs in the development of CC were preliminarily ascertained. This may provide the basis for tumor prevention and immunotherapy in the future.

Abbreviations

CC: cervical cancer; IRL: immune-related lncRNA; IS: immune score; ssGSEA: single-sample gene set enrichment analysis; OS: overall survival; AUC: area under the curve; ROC: receiver operating characteristic; HPV: human papillomavirus; EMT: epithelial-mesenchymal transition; GSEA: gene set enrichment analysis; TCGA: The Cancer Genome Atlas; GEO: gene expression omnibus; HR: hazard ratio

Declarations

Availability of data and materials

The RNA sequencing profiles are able to be gained from The Cancer Genome Atlas (TCGA) (https://toil.xenahubs.net) and Gene Expression Omnibus (GEO)(http://www.ncbi.nlm.nih.gov/geo/). The immune-related genes can be downloaded from the InnateDB database (http://www.innatedb.com).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflicts of interest.

Authors' contributions

JZ, BC, NZ, and XZ contributed to data analysis, interpretation, and drafting. JZ and JT contributed to study design, study supervision, and final approval of the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The author, Jianfeng Zheng, gratefully acknowledges Miss Lynn for providing assistance with English writing.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
2. Hu Z, Ma D. The precision prevention and therapy of HPV-related cervical cancer: new concepts and clinical implications. Cancer Med. 2018;7(10):5217–36.
3. Vu M, Yu J, Awolude OA, Chuang L. Cervical cancer worldwide. Curr Probl Cancer. 2018;42(5):457–65.
4. Liontos M, Kyriazoglou A, Dimitriadis I, Dimopoulos MA, Bamias A. Systemic therapy in cervical cancer: 30 years in review. Crit Rev Oncol Hematol. 2019;137:9–17.
5. Barabas AZ, Cole CD, Lafreniere R. Antibody-initiated beneficial and harmful immune responses. Immunol Res. 2018;66(6):783–9.
6. Ventriglia J, Paciolla I, Pisano C, Cecere SC, Di Napoli M, Tambaro R, Califano D, Losito S, Scognamiglio G, Setola SV, et al. Immunotherapy in ovarian, endometrial and cervical cancer: State of the art and future perspectives. Cancer Treat Rev. 2017;59:109–16.
7. O'Donnell JS, Teng MWL, Smyth MJ. Cancer immunoediting and resistance to T cell-based immunotherapy. Nat Rev Clin Oncol. 2019;16(3):151–67.
8. Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell. 2018;172(3):393–407.

9. Tornesello ML, Faraonio R, Buonaguro L, Annunziata C, Starita N, Cerasuolo A, Pezzuto F, Tornesello AL, Buonaguro FM. The Role of microRNAs, Long Non-coding RNAs, and Circular RNAs in Cervical Cancer. Front Oncol. 2020;10:150.
10. Pan Z, Liu L, Nie W, Miggin S, Qiu F, Cao Y, Chen J, Yang B, Zhou Y, Lu J, et al. Long non-coding RNA AGER-1 functionally upregulates the innate immunity gene AGER and approximates its anti-tumor effect in lung cancer. Mol Carcinog. 2018;57(3):305–18.
11. Mao D, Hu C, Zhang J, Feng C, Zhang Z, Wang J, Man Z, Zhu Z, Wang Y, Zhao H, et al. Long Noncoding RNA GM16343 Promotes IL-36β to Regulate Tumor Microenvironment by CD8(+)T cells. Technol Cancer Res Treat. 2019;18:1533033819883633.
12. Hu Q, Ye Y, Chan LC, Li Y, Liang K, Lin A, Egranov SD, Zhang Y, Xia W, Gong J, et al. Oncogenic lncRNA downregulates cancer cell antigen presentation and intrinsic tumor suppression. Nat Immunol. 2019;20(7):835–51.
13. Fang P, Xiang L, Chen W, Li S, Huang S, Li J, Zhuge L, Jin L, Feng W, Chen Y, et al. LncRNA GAS5 enhanced the killing effect of NK cell on liver cancer through regulating miR-544/RUNX3. Innate Immun. 2019;25(2):99–109.
14. Pei X, Wang X, Li H. LncRNA SNHG1 regulates the differentiation of Treg cells and affects the immune escape of breast cancer via regulating miR-448/IDO. Int J Biol Macromol. 2018;118(Pt A):24–30.
15. Jiang R, Tang J, Chen Y, Deng L, Ji J, Xie Y, Wang K, Jia W, Chu WM, Sun B. The long noncoding RNA lnc-EGFR stimulates T-regulatory cells differentiation thus promoting hepatocellular carcinoma immune evasion. Nat Commun. 2017;8:15129.
16. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675–8.
17. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012;22(9):1760–74.
18. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R. NCBI GEO: mining tens of millions of expression profiles–database and tools update. Nucleic Acids Res. 2007;35(Database issue):D760–5.
19. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
20. Jiang H, Wong WH. SeqMap: mapping massive amount of oligonucleotides to the genome. Bioinformatics. 2008;24(20):2395–6.
21. Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock RE, Brinkman FS, Lynn DJ. InnateDB: systems biology of innate immunity and beyond–recent updates and continuing curation. Nucleic Acids Res. 2013;41(Database issue):D1228–33.

22. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
23. Berkman SJ, Roscoe EM, Bourret JC. Comparing self-directed methods for training staff to create graphs using Graphpad Prism. J Appl Behav Anal. 2019;52(1):188–204.
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.
25. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol. 2018;1711:243–59.
26. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. Nature. 2014;505(7483):344–52.
27. Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, Zagganas K, Tsanakas P, Floros E, Dalamagas T, et al. DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic Acids Res. 2016;44(D1):D231–8.
28. Dweep H, Gretz N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods. 2015;12(8):697.
29. Jochems C, Schlom J. Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity. Exp Biol Med (Maywood). 2011;236(5):567–79.
30. Zhao LP, Li RH, Han DM, Zhang XQ, Nian GX, Wu MX, Feng Y, Zhang L, Sun ZG. Independent prognostic Factor of low-expressed LncRNA ZNF667-AS1 for cervical cancer and inhibitory function on the proliferation of cervical cancer. Eur Rev Med Pharmacol Sci. 2017;21(23):5353–60.
31. Li YJ, Yang Z, Wang YY, Wang Y. Long noncoding RNA ZNF667-AS1 reduces tumor invasion and metastasis in cervical cancer by counteracting microRNA-93-3p-dependent PEG3 downregulation. Mol Oncol. 2019;13(11):2375–92.
32. Yang F, Wang T, Du P, Fan H, Dong X, Guo H. M2 bone marrow-derived macrophage-derived exosomes shuffle microRNA-21 to accelerate immune escape of glioma by modulating PEG3. Cancer Cell Int. 2020;20:93.
33. Tan J, Jin X, Wang K. Integrated Bioinformatics Analysis of Potential Biomarkers for Prostate Cancer. Pathol Oncol Res. 2019;25(2):455–60.
34. Wang W, Chen G, Wang B, Yuan Z, Liu G, Niu B, Chen Y, Zhou S, He J, Xue H. Long non-coding RNA BZRAP1-AS1 silencing suppresses tumor angiogenesis in hepatocellular carcinoma by mediating THBS1 methylation. J Transl Med. 2019;17(1):421.
35. Gu Y, Feng C, Liu T, Zhang B, Yang L. The downregulation of lncRNA EMX2OS might independently predict shorter recurrence-free survival of classical papillary thyroid cancer. PLoS One. 2018;13(12):e0209338.
36. Duan M, Fang M, Wang C, Wang H, Li M. LncRNA EMX2OS Induces Proliferation, Invasion and Sphere Formation of Ovarian Cancer Cells via Regulating the miR-654-3p/AKT3/PD-L1 Axis. Cancer

Manag Res. 2020;12:2141–54.
37. Huang T, Sun L, Yuan X, Qiu H. Thrombospondin-1 is a multifaceted player in tumor progression. Oncotarget. 2017;8(48):84546–58.
38. Li X, Shao C, Shi Y, Han W. Lessons learned from the blockade of immune checkpoints in cancer immunotherapy. J Hematol Oncol. 2018;11(1):31.
39. Katoh M. Canonical and non-canonical WNT signaling in cancer stem cells and their niches: Cellular heterogeneity, omics reprogramming, targeted therapy and tumor plasticity (Review). Int J Oncol. 2017;51(5):1357–69.
40. Mansour MA. Ubiquitination: Friend and foe in cancer. Int J Biochem Cell Biol. 2018;101:80–93.
41. Zhou W, Guo S, Liu M, Burow ME, Wang G. Targeting CXCL12/CXCR4 Axis in Tumor Immunotherapy. Curr Med Chem. 2019;26(17):3026–41.
42. Zhang Z, Li J, He T, Ding J. Bioinformatics Identified 17 Immune Genes as Prognostic Biomarkers for Breast Cancer: Application Study Based on Artificial Intelligence Algorithms. Front Oncol. 2020;10:330.
43. Luo J, Xie Y, Zheng Y, Wang C, Qi F, Hu J, Xu Y. Comprehensive insights on pivotal prognostic signature involved in clear cell renal cell carcinoma microenvironment using the ESTIMATE algorithm. Cancer Med. 2020;9(12):4310–23.
44. Yahya MA, Sharon SM, Hantisteanu S, Hallak M, Bruchim I. The Role of the Insulin-Like Growth Factor 1 Pathway in Immune Tumor Microenvironment and Its Clinical Ramifications in Gynecologic Malignancies. Front Endocrinol (Lausanne). 2018;9:297.
45. Alvisi G, Brummelman J, Puccio S, Mazza EM, Tomada EP, Losurdo A, Zanon V, Peano C, Colombo FS, Scarpa A, et al. IRF4 instructs effector Treg differentiation and immune suppression in human cancer. J Clin Invest. 2020;130(6):3137–50.
46. El-Said H, Fayyad-Kazan M, Aoun R, Borghol N, Skafi N, Rouas R, Vanhamme L, Mourtada M, Ezzeddine M, Burny A, et al. MiR302c, Sp1, and NFATc2 regulate interleukin-21 expression in human CD4 + CD45RO + T lymphocytes. J Cell Physiol. 2019;234(5):5998–6011.
47. Szebeni GJ, Vizler C, Kitajka K, Puskas LG. Inflammation and Cancer: Extra- and Intracellular Determinants of Tumor-Associated Macrophages as Tumor Promoters. Mediators Inflamm. 2017;2017:9294018.
48. Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, Jin X, Yin J, Chen L, Zhang Y, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. Nat Commun. 2020;11(1):1000.
49. Shen Y, Peng X, Shen C. Identification and validation of immune-related lncRNA prognostic signature for breast cancer. Genomics. 2020;112(3):2640–6.
50. Wang W, Zhao Z, Yang F, Wang H, Wu F, Liang T, Yan X, Li J, Lan Q, Wang J, et al. An immune-related lncRNA signature for patients with anaplastic gliomas. J Neurooncol. 2018;136(2):263–71.

51. Wei C, Liang Q, Li X, Li H, Liu Y, Huang X, Chen X, Guo Y, Li J. Bioinformatics profiling utilized a nine immune-related long noncoding RNA signature as a prognostic target for pancreatic cancer. J Cell Biochem. 2019;120(9):14916–27.
52. Zhang Y, Zhang L, Xu Y, Wu X, Zhou Y, Mo J. **Immune-related long noncoding RNA signature for predicting survival and immune checkpoint blockade in hepatocellular carcinoma**. *Journal of Cellular Physiology* 2020.
53. Zhong Z, Hong M, Chen X, Xi Y, Xu Y, Kong D, Deng J, Li Y, Hu R, Sun C, et al. Transcriptome analysis reveals the link between lncRNA-mRNA co-expression network and tumor immune microenvironment and overall survival in head and neck squamous cell carcinoma. BMC Med Genomics. 2020;13(1):57.
54. Zhou M, Zhang Z, Zhao H, Bao S, Cheng L, Sun J. An Immune-Related Six-lncRNA Signature to Improve Prognosis Prediction of Glioblastoma Multiforme. Mol Neurobiol. 2018;55(5):3684–97.

Figures

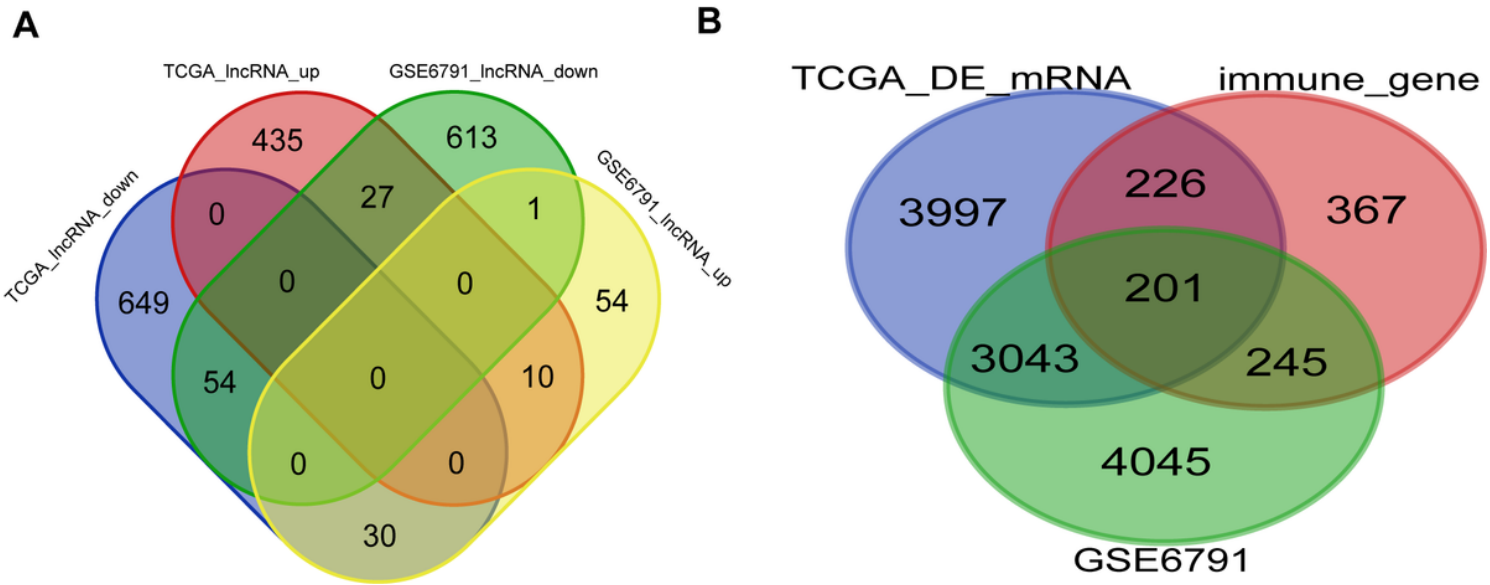

Figure 1

Venn analysis (A) Venn diagram of Differential lncRNAs (B) Venn diagram of Differential mRNA and immune-related genes

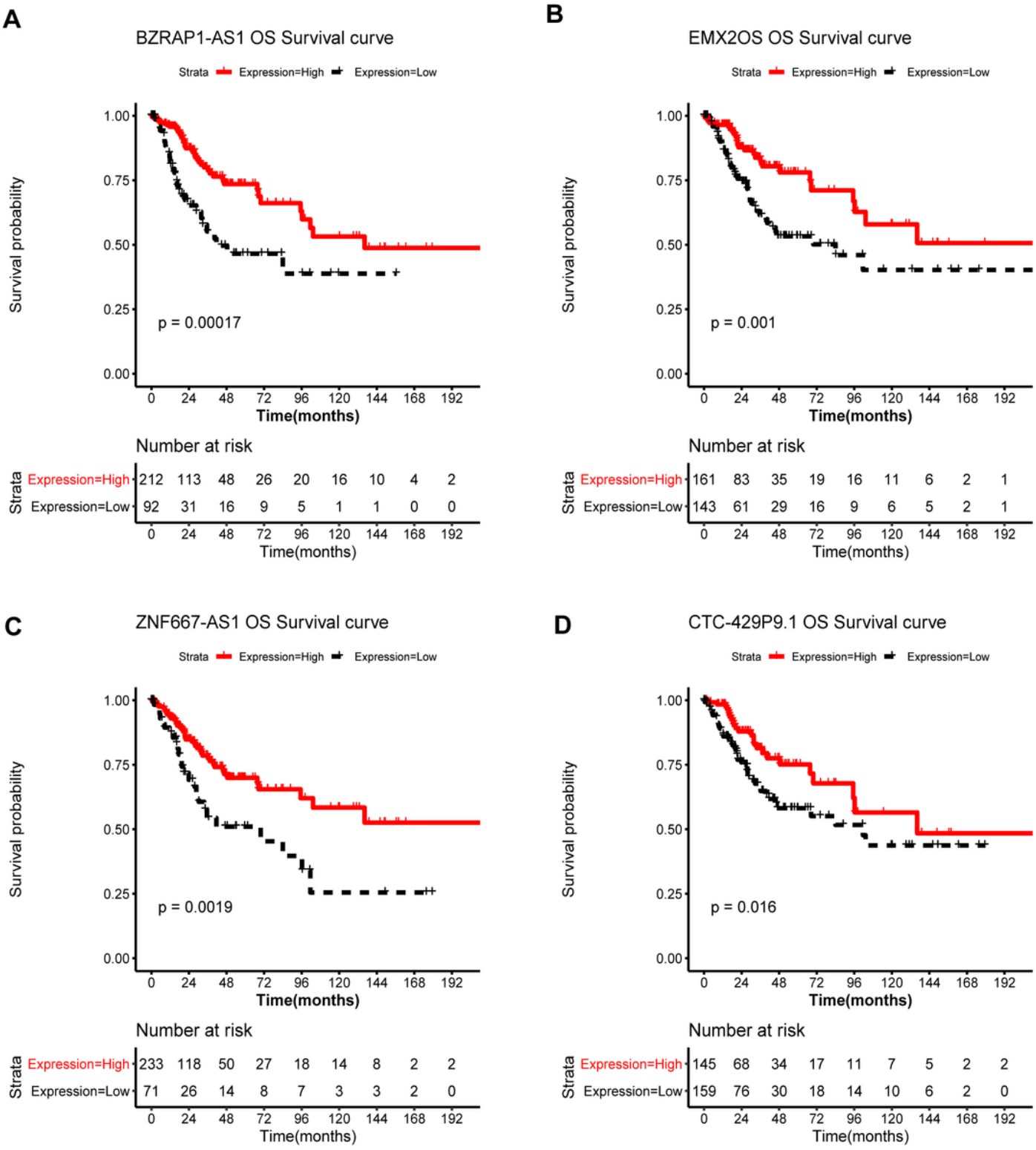

Figure 2

Kaplan-Meier curve of 4-IRLs

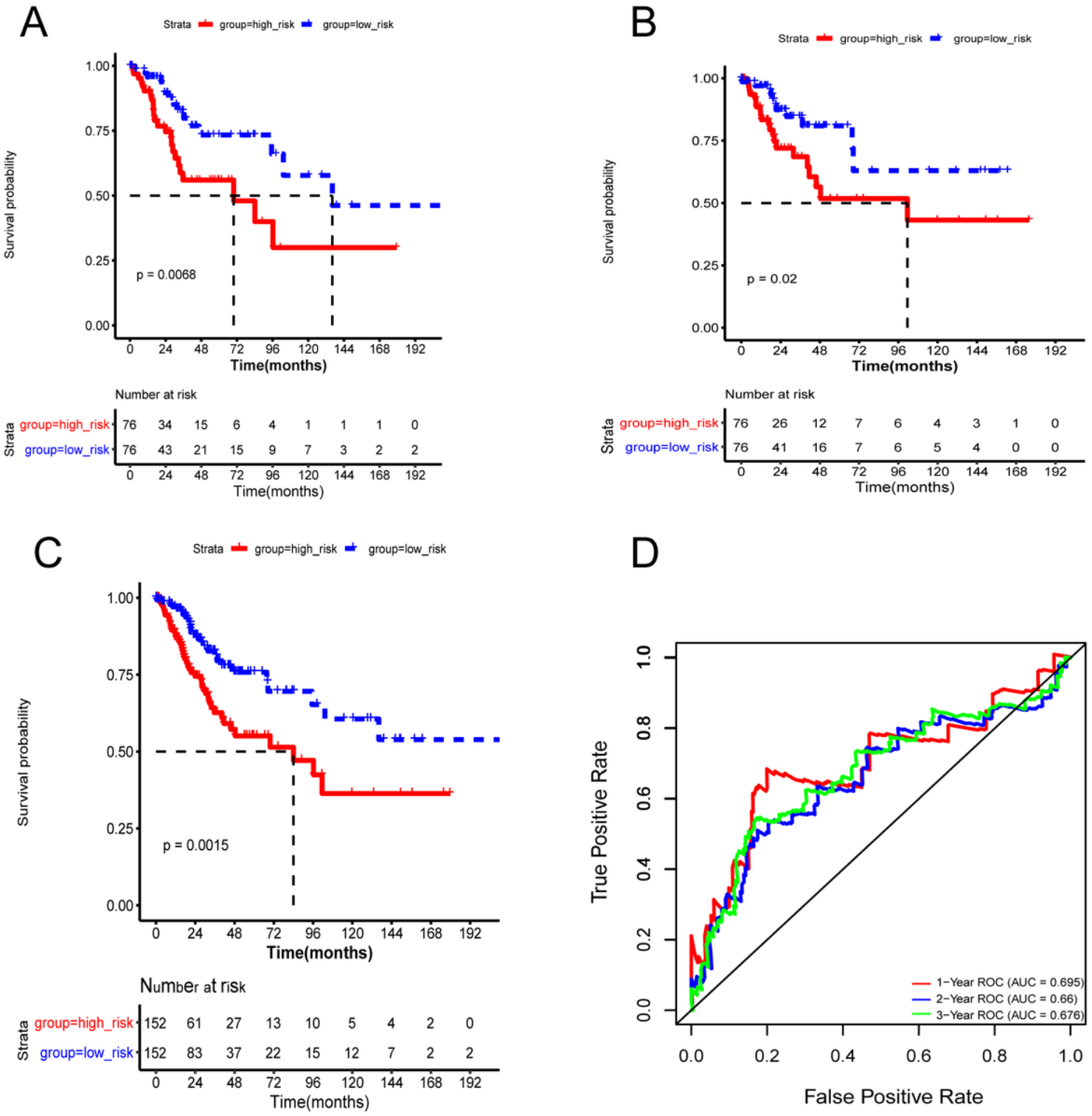

Figure 3

The 4-IRLs Prognostic Risk Model and Validating (A) Kaplan-Meier survival curves of OS among CC patients from different groups stratified by the signature in the Training-set, the Valid-set, and the Total-set. (B) Time-dependent receiver operating characteristic (ROC) curve for predicting overall survival (OS) of the Risk Model

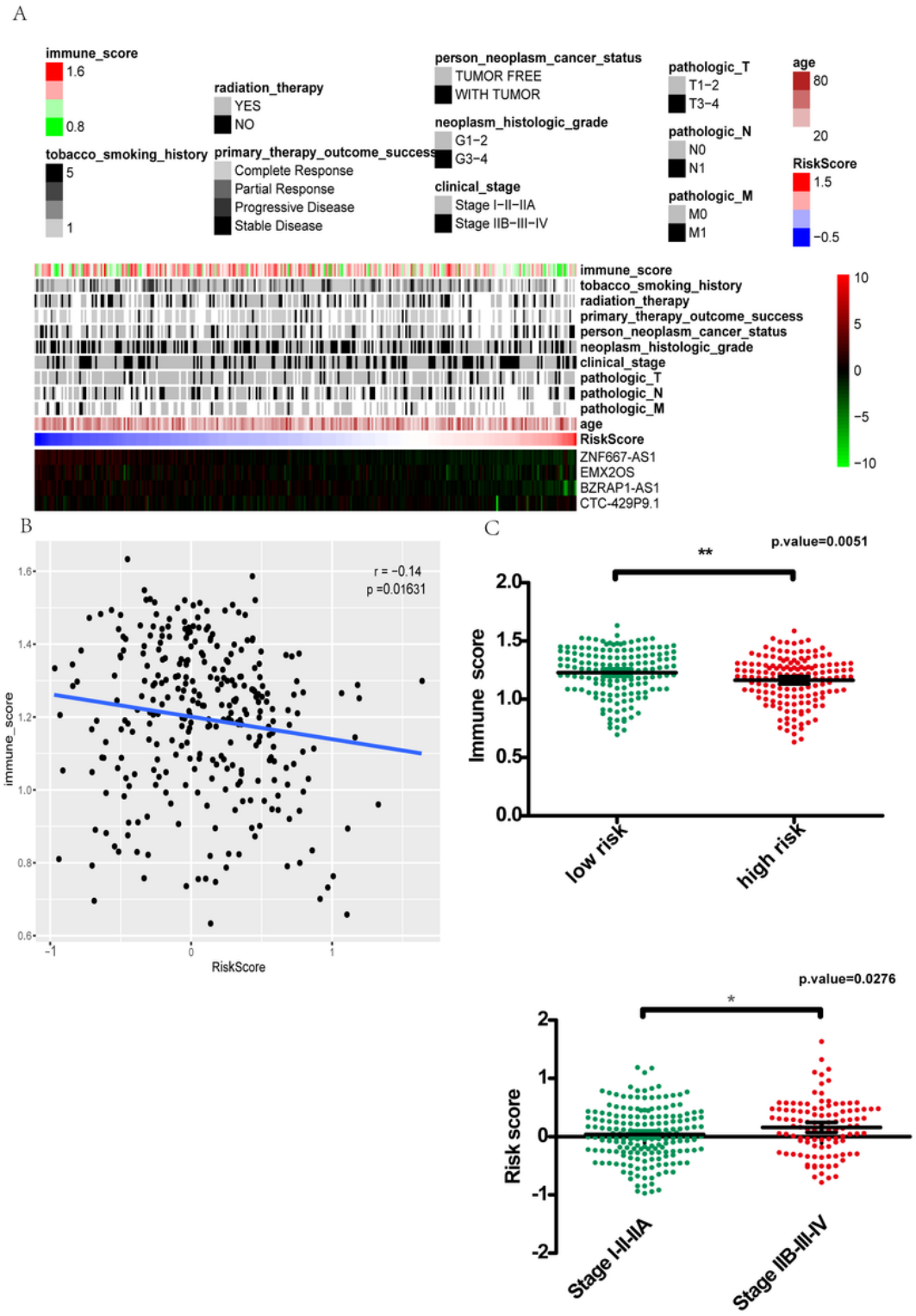

Figure 4

Clinical pathological characteristics of risk score model (A) distribution of immune score, clinicopathologic features, 4 IRLs expression, and risk score (B) correlations between Risk score and immune score (C) immune score in different risk groups and Risk Score in tumor different stages

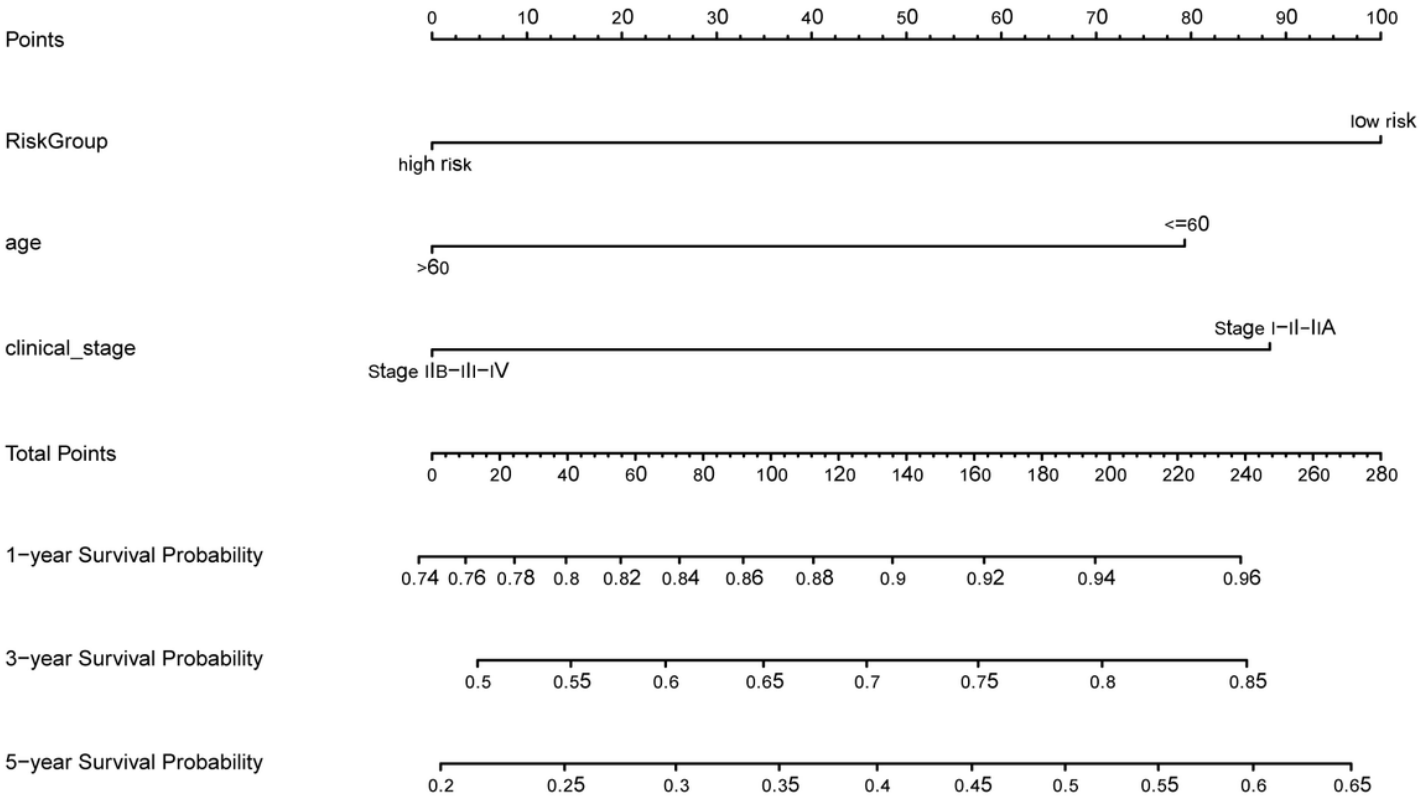

Figure 5

A nomogram based on the signature and clinical information.

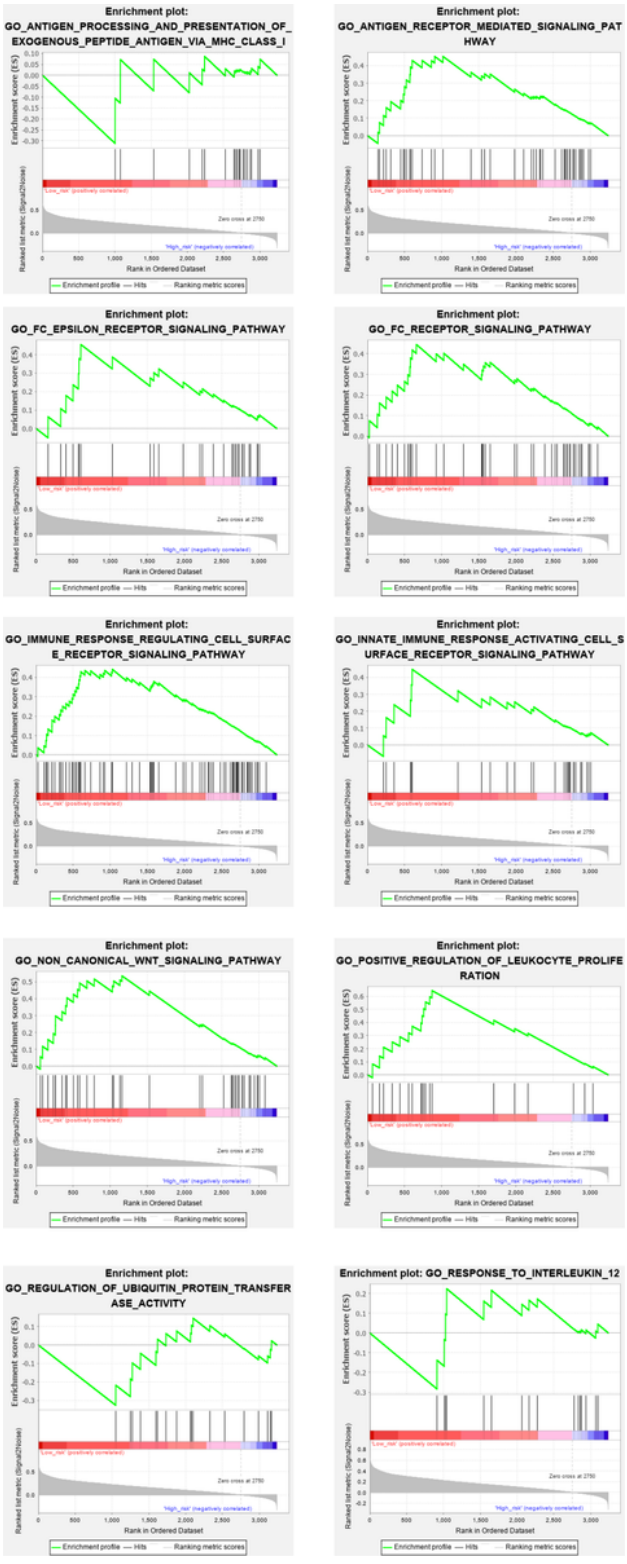

Figure 6

immunologic characteristics regulated via the GSEA

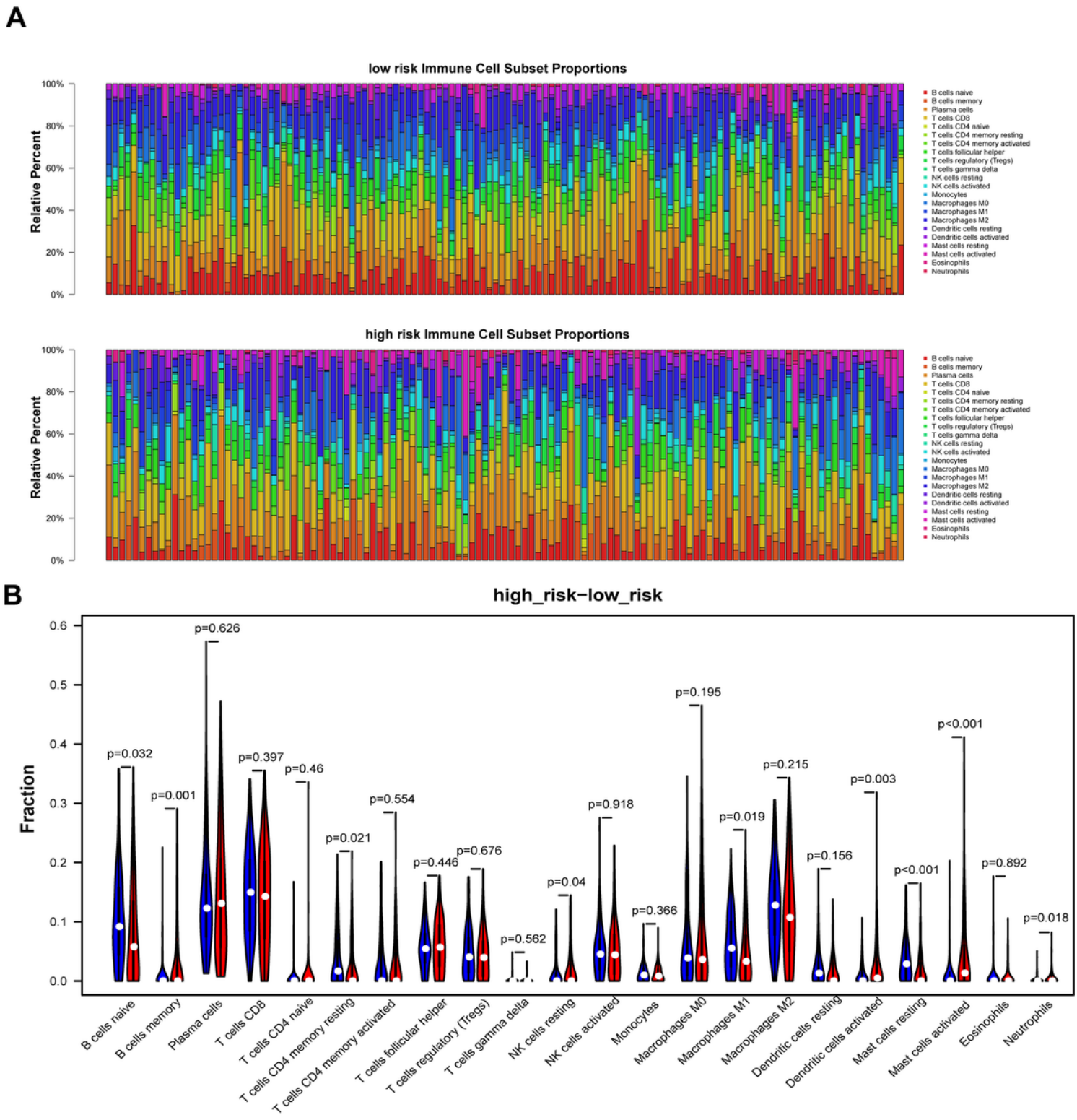

Figure 7

Evaluation of tumor-infiltrating immune cells (A) The landscape of immune infiltration in risk groups (B) The difference of 22 tumor-infiltrating immune cells among risk groups

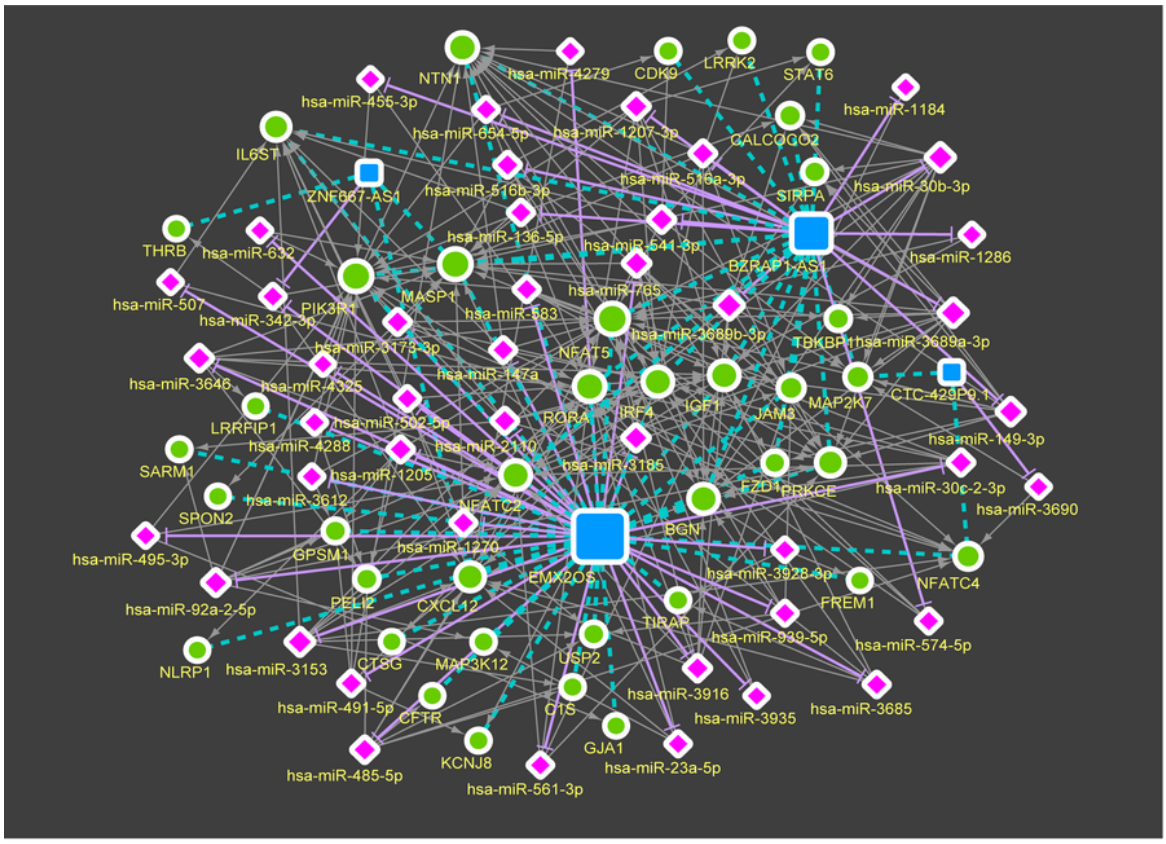

Figure 8

ceRNA network of 4-IRLs

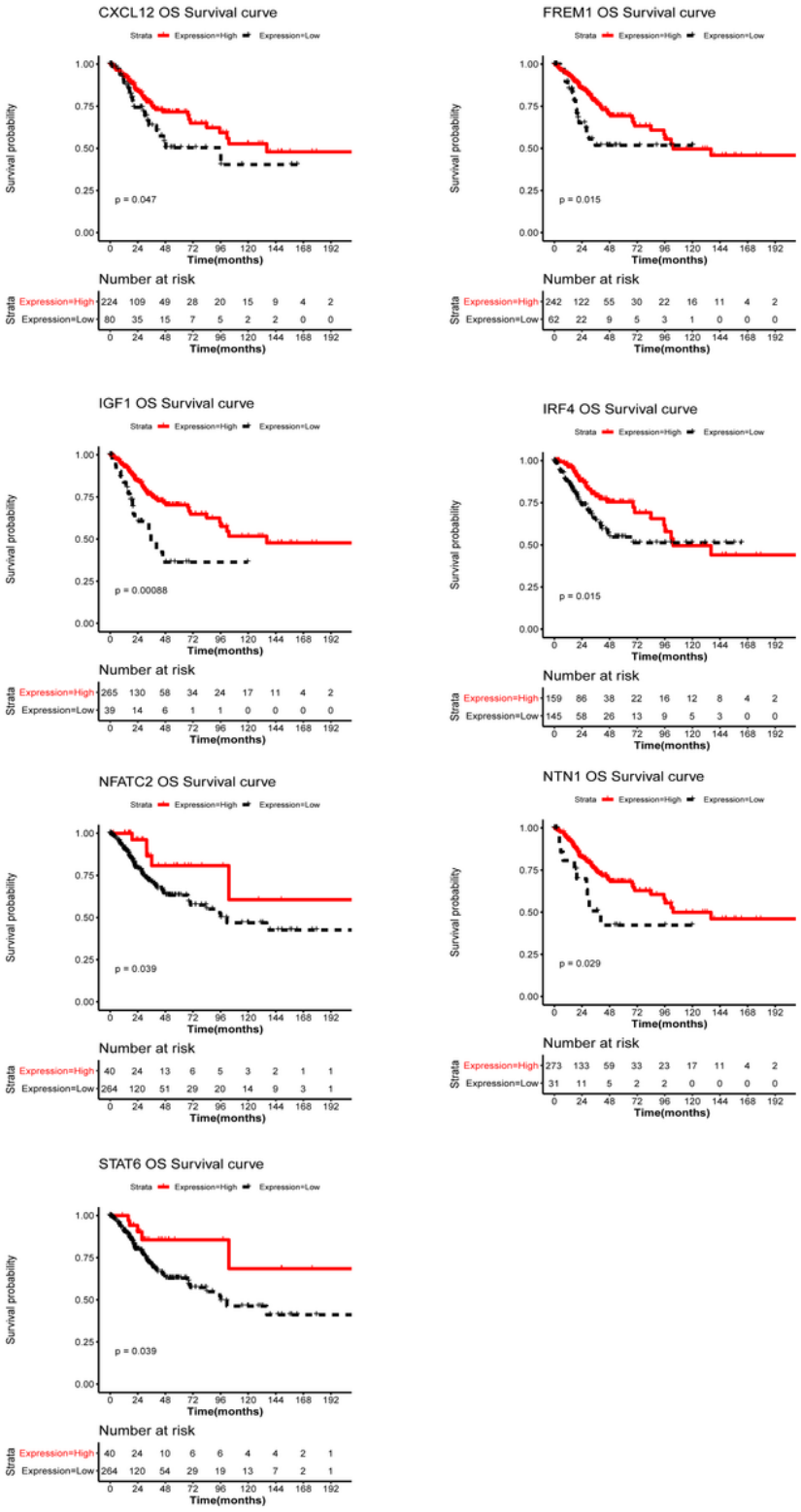

Figure 9

Kaplan-Meier curve of 7 mRNAs