Identification and Validation of Immune-Related Gene Prognostic Signature for breast cancer

Bing Zhou
The Third Hospital of Nanchang
Linyi Zhang
Nanchang University
Qinyu Wang
The Third Hospital of Nanchang
Changqin Pu
Nanchang University
Sixuan Guo
Nanchang University
Shuhui Lai
Nanchang University
Heming Zhang
Nanchang University
Wenwei Li
The Third Hospital of Nanchang
Zhibing Zhou
The Third Hospital of Nanchang
Yuexia Chen
The Third Hospital of Nanchang
Liping Zeng (✉ zeng13576018725@hotmail.com)
Nanchang University
Yao Zhou (✉ 331016157@qq.com)
The Third Hospital of Nanchang

Research Article

Keywords: breast cancer, immune-relevant genes, prognosis, overall survival

Posted Date: April 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-386385/v1

Abstract

Background

Although the outcome of breast cancer patients has been improved by advances in early detection, diagnosis and treatment. Due to the heterogeneity of the disease, prognostic assessment still faces challenges. The accumulated data indicate that there is a clear correlation between the tumor immune microenvironment and clinical outcomes.

Objective

Construct immune-related gene pairs to evaluate the prognosis of breast cancer and patient survival rate.

Methods

From the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) database, the Gene expression profiles and clinical data of breast cancer samples were downloaded. TCGA cohort were further divided into a training set (n = 764) and internal validation sets (n = 325). The GEO cohort was analyzed as an external validation cohort (n = 327). In the training set, differently expressed immune-relevant genes (IRGs) were screened firstly, and they were used to construct immune-relevant gene pairs (IRGPs). Then, the prognostic IRGPs were identified via univariate Cox regression analysis. Finally, least absolute shrinkage and selection operator (LASSO) Cox regression analysis was used to constituted the IRGP prognostic signature. Kaplan-Meier (KM) survival curves, receiver operating characteristic (ROC) curve analysis, univariate and multivariate Cox regression analysis were used to estimate the predictive value of the IRGP prognostic signature. And the IRGP prognostic signature was validated in the internal validation cohort and external validation cohort. We used gene set enrichment analysis (GSEA) to elucidate the biological functions of the IRGP prognostic signature.

Results

A total of 474 differently expressed IRGs and 2942 prognostic IRGPs were identified. Finally, we generated a IRGP prognostic signature consisting of 33 IRGPs. Subsequently, the 33 IRGPs grouped BRCA patients into high- and low-risk groups. Kaplan-Meier curves shown a significantly different overall survival in risk groups. Time-dependent ROC curves indicated that the IRGP prognostic signature possessed a high specificity and sensitivity in all the sets. Univariate and multivariate Cox regression analysis showed a statistical significance for the prognostic value of IRGP prognostic signature and the IRGP prognostic signature was a strong independent risk factor. The functional enrichment analysis indicated that low IRGP value was correlated with biological processes related to immune. Immune cell infiltration analysis indicated a significant difference in percentage of M2 macrophages between high- and low-risk groups.

Conclusion

The 33-IRGPs prognostic signature was developed to provide new insights for the identification of high-risk breast cancer and the evaluation of the possibility of immunotherapy in personalized breast cancer treatment.

1. Introduction

Breast cancer (BRCA) is the most frequently diagnosed malignant tumor in female, affects about 12% of women worldwide [1]. The incidence of BRCA was increased from 2004 to 2013 by about 2% per year among women [2]. TNM stage system, which takes both clinical and pathology information into consideration including tumor size (T), regional lymph nodes (N) and spread to distant metastatic sites (M), represents the most important prognostic factor for BRCA, and the higher the stage at diagnosis, the poorer the prognosis [3, 4]. Although the 5-year survival rate was more than 90% for stage I and II BRCA patients, it drops down to about 27% for stage IV [5]. Thus, early detection, screening, and advanced tailored treatment are extremely important to reduce the mortality of BRCA. However, due to the high degree of heterogeneity of breast cancer, the prognosis of patients with similar clinical features may vary greatly. Therefore, in order to stratify patients more accurately, in addition to clinical factors, other prognostic factors need to be considered.

Immune oncology has attracted more and more attention because of its excellent clinical application value in many kinds of malignant tumors [6]. Immune-based strategies point out a new direction for the treatment and prevention of breast cancer. Previous researches have shown that the immune cells and immune-related genes are attractive bio-target for regulating cancer prognosis [7–9]. Among them, because the construction of immune gene pairs can directly compare the expression of two genes in each sample, and there is no need for comparison between samples, the correction between samples is saved and the accuracy is improved [10]. It provides convenience for various subsequent studies. However, previous studies on breast cancer have mostly focused on the relationship between immune cell infiltration and tumor prognosis, and there are relatively few studies on immune-related gene pairs. Our research has done our best to fill the gap in the study of some breast cancer-related immune gene pairs.

In this study, we established a 33-IRGPs signature to predict the individual prognostic characteristics of BRCA by univariate Cox regression analysis and LASSO model. For validation of the signature, we investigated its accuracy and efficiency in determining the prognosis of BRCA patients with KM curves, ROC curve analysis, univariate and multivariate Cox regression analysis. The findings of this study showed and proved that the 33-IRGPs signature can be applied in the clinical prognosis of BRCA patients.

2. Materials And Methods

2.1 Transcriptome data acquisition and pre-processing

The BRCA transcriptome data and related clinical characteristics were downloaded from The Cancer Genome Atlas (TCGA) database [11]. In total 1089 BRCA samples and 113 adjacent normal samples were extracted from TCGA. Then, all selected patients were further separated into training and internal validation groups(7: 3) to apply the prognostic analyses based on cohorts. The models are identified and evaluated by “caret” package with its “createDataPartition” function [12]. GSE20685 Microarray data [13] from the Gene Expression Omnibus (GEO) database [14] was analyzed as external validation cohort which contained data of 327 BRCA patients. All selected clinical characteristics of BRCA patients were listed in Table 1.

Table1. Clinical characteristics for BRCA patients

Variables	Training cohort (764 patients)	Internal validation cohort (325 patients)	External validation cohort (327 patients)
follow-up (days) (mean, range)	1211 (1-8432)	1272 (1-8481)	2880 (146-5146)
Age(years) (mean, range)	58 (26-90)	58 (26-90)	48 (24-84)
Fraction genome altered (mean, range)	0.2978 (0-0.9971)	0.2937 (0-0.9463)	/
Sample initial weight (mean, range)	288.3 (20.0-1760.0)	328.7 (5.0-2190.0)	/
Neoadjuvant chemotherapy(n, %)			
yes	/	/	268 (81.96)
no	/	/	54 (16.51)
unknown	/	/	5 (1.53)
T stage(n, %)			
T1	199 (26.05)	80 (24.62)	101 (30.89)
T2	448 (58.64)	181 (55.69)	188 (57.49)
T3	88 (11.52)	50 (15.38)	26 (7.95)
T4	27 (3.53)	13 (4.00)	12 (3.67)
TX	2 (0.26)	1 (0.31)	/
N stage(n, %)			
N0	351 (45.94)	163 (50.15)	137 (41.90)
N1	259 (33.90)	100 (30.77)	87 (26.61)
N2	84 (10.99)	35 (10.77)	63 (19.26)
N3	54 (7.07)	23 (7.08)	40 (12.23)
NX	16 (2.10)	4 (1.23)	/
M stage(n, %)			
M0	632 (82.72)	274 (84.31)	319 (97.55)
M1	16 (2.10)	6 (1.84)	8 (2.45)

MX	116 (15.18)	45 (13.85)	/
Tumor stage(n, %)			
Stage I	127 (16.62)	54 (16.62)	/
Stage II	432 (56.54)	184 (56.61)	/
Stage III	174 (22.78)	74 (22.77)	/
Stage IV	14 (1.83)	6 (1.85)	/
Stage X	17 (2.23)	7 (2.15)	/
ER status (n, %)			
Negative	153 (20.03)	83 (25.54)	/
Positive	576 (75.39)	228 (70.15)	/
Indeterminate	35 (4.58)	14 (4.31)	/
PR status (n, %)			
Negative	237 (31.02)	106 (32.62)	/
Positive	491 (64.27)	203 (62.46)	/
Indeterminate	36 (4.71)	16 (4.92)	/
HER2 status (n, %)			
Negative	376 (49.22)	183 (56.31)	/
Positive	126 (16.49)	37 (11.38)	/
Indeterminate	262 (34.29)	105 (32.31)	/
Patient primary tumor site(n, %)			
Left	388 (50.79)	179 (55.08)	/
Right	376 (49.21)	146 (44.92)	/
BRCA subtype (n, %)			
Basal-like	119 (15.58)	53 (16.31)	/
HER2-enriched	46 (6.02)	26 (8.00)	/
Luminal B	147 (19.24)	44 (13.54)	/
Luminal A	354 (46.33)	152 (46.77)	/
Normal-like	88 (11.52)	48 (14.77)	/
Unknown	10 (1.31)	2 (0.61)	/

Immune subtype(n, %)			
C1	260 (34.03)	107 (32.92)	/
C2	273 (35.74)	115 (35.39)	/
C3	123 (16.10)	67 (20.62)	/
C4	71 (9.29)	21 (6.46)	/
C6	27 (3.53)	13 (4.00)	/
Unknown	10 (1.31)	2 (0.61)	

Abbreviation: BRCA, breast cancer.

2.2 Acquisition of somatic mutation data

A total of 4 subtypes of data files from the TCGA database were selected, from which the “Masked Somatic Mutation” data was processed by VarScan software. Somatic mutation data was processed into Mutation Annotation Format (MAF) file and visualized by “maftools” [15] R package, which can provide multiple analyzing models. There was no ethical conflict to declare because all the data in this research were from public databases.

2.3 Prognostic immune-relevant gene pairs identification

Immune-relevant genes (IRGs) from immported database (https://www.immport.org/home) were obtained to perform differentially expressed genes screening. The genes with different expression levels between tumor and normal samples of BRCA samples in TCGA cohort were analyzed by limma R packages. At the same time, P-value were adjusted by false discovery rate (FDR) that proposed by the Benjamini-Hochberg procedure to limit the false positives results. The genes were considered as differential expression with |log2(fold change)|>1 and FDR < 1.0E-5.

Besides, the immune-relevant genes with different expression levels were used to construct immune-relevant gene pairs (IRGPs). Each IRGP was calculated by comparing the expression levels of genes in a particular sample or sequence, and if IRG 1 > IRG 2, then IRGP equals to 1, otherwise IRGP equals to 0. By taking advantages of analyzing genes in a pairwise approach, standardization is no more required in individualized analysis. Besides, some IRGPs with unique values of 1 or 0 were excluded to avoid the biases. The prognostic IRGPs were identified via univariate Cox regression analysis in the TCGA cohort, where IRGPs with P-value < 1.0E-5 were selected for further analysis.

2.4 The IRGP prognostic signature construction and validation

TCGA samples were randomly separated into training (764 patients) and internal validation cohorts (325 patients). The endpoints analyzed in this study were overall survival (OS), defined as the interval between the date of diagnosis and death. Least absolute shrinkage and selection operator(LASSO) is a biased estimation tool for complex collinearity data, which can select variables and estimate parameters at the same time, and solve the problem of multicollinearity in regression analysis [16]. Thus, LASSO Cox regression analysis was applied to reduce dimensionality of IRGPs by R package glmnet. IRGPs represented by optimal values of the penalty parameter λ, which were determined by ten-fold cross-validations, constituted the IRGP prognostic signature in this study. On the basis of IRGP prognostic signature, the risk score for each BRCA sample was calculated according to the following formula: risk score = $\text{expression}_{\text{gene 1}} \times \beta_{\text{gene 1}} + \text{expression}_{\text{gene 2}} \times \beta_{\text{gene 2}} + \cdots + \text{expression}_{\text{gene x}} \times \beta_{\text{gene x}}$, (x = the number of IRGPs; β = coefficient value for each IRGP).

All BRCA samples were classified into high- and low-risk groups based on the cut-off values of median value of risk score. Kaplan-Meier (KM) survival curves were applied to calculate the over survival (OS) differences between two groups, and the statistical significance was obtained by log-rank test. The values of area under the curve (AUC) in 1, 3, and 5 years from the results of receiver operating characteristic (ROC) curve analysis were calculated and taken to evaluate the accuracy and sensitivity of the survival prediction results.

Then, the multivariate Cox regression model was used to evaluate whether the prognostic value of IRGPs was independent of clinical characteristics. Furthermore, the comparison between IRGP prognostic signature and clinical features was performed by forest plots to determine the effectiveness of the prognostic value. Wald test was performed to compare the statistical difference of IRGPs in clinical characteristics.

2.5 Clinical utility of IRGP prognostic signature

A composite nomogram was developed based on the IRGP prognostic signature and available clinical factors to predict the OS of BRCA patients with the rms package. Calculate the consistency index (C-index) to evaluate the discriminative ability of the nomogram. Besides, a calibration curve was drawn to compare the predicted probability and actual probability of OS. Each component of the nomogram is given points, and their sum represents the total points of a patient has obtained.

2.6 Estimation of immune infiltration

Immune infiltration assessment was performed by cibersort method to quantify the absolute abundance of 22 immune cell populations in heterogeneous tissues from transcriptome data [17]. The cibersort method was applied by CIBERSORT package to convert mRNA data into the levels of infiltrating immune cells. And standard annotation files was used to prepare gene expression profiles before immune infiltration analysis.

2.7 Functional enrichment analysis

Gene Set Variation Analysis (GSVA) is a method of gene set enrichment analysis, which estimates the changes of pathways and biological process activities in a sample population in an unsupervised manner[18]. The gene set file of “c5.all.v7.0.symbols.gmt”, downloaded from the “Molecular Signatures Database” (https://www.gsea-msigdb.org/gsea/), were employed for GSVA using “GSVA” R packages. Each group of data was taken a normality test for test the data distribution by the means of shapiro, and homogeneity test for variance in multiplicate samples by the means of bartlett. The result is normal distribution and comparable when $P > 0.05$, and one- way ANOVA was used to compare differences between groups. Otherwise the data was analyzed by Kruskal-Wallis test. The gene sets in the “Molecular Signatures Database” of “c5.all.v7.0.symbols.gmt”, “c2.all.v7.0.symbols.gmt”, and “h.all.v7.0.symbols.gmt” were used to explored the potential mechanisms of the IRGP prognostic signature by gene set enrichment analysis (GSEA) with the fgsea package[19]. The number of random sample permutations was set at 10000, and the significance threshold was set at $p < 0.05$.

3. Results

3.1 Prognostic immune-relevant gene pairs identification

We obtained 1382 common IRGs overlapped between TCGA cohort and GSE20685 cohort. A total of 135 up-regulated and 339 down-regulated IRGs were identified between 1089 BRCA tumor and 113 normal samples in TCGA cohort (Fig. 1A). Then we acquired 683795 common IRGPs based on these 1382 common IRGs between TCGA cohort and GSE20685 cohort by gene pairwise calculation. Univariate Cox regression analysis was performed for the 683795 common IRGPs, of which 2942 IRGPs showed significant prognostic potential (P-value < 1.0E-5), containing 1102 IRGs. Among these 2942 IRGPs, 1692 IRGPs were risky and 1250 IRGPs were protective (Fig. 1B). There were 356 overlapped IRGs between differently expressed IRGs and prognostic IRGs, and they constituted 302 IRGPs (Fig. 1C). The workflow chart for data collection and analysis was shown in Fig. S1.

3.2 Construction of molecular subgroups using prognostic immune-relevant gene pairs (IRGPs)

Firstly, unsupervised clustering methods was performed to group 1089 tumor samples into different subgroups based on 302 prognostic IRGPs. And optimal cluster number was calculated by ConsensusClusterPlus package. Three distant patient clusters, termed as the subtype1, subtype2 and subtype3, were finally identified (Fig. 2A-B). Through the log-rank test, the Kaplan-Meier curve shown that the overall survival difference between the three clusters is significant (Fig. 2C). The patients in subtype2 had worst survival outcome, whereas the survival time in subtype3 group was the longest. Samples in subtype3 exhibited a more negative ER status or PR status and lower fraction of genome altered, whereas samples in subtype1 and subtype2 exhibited a more positive ER status or PR status and higher fraction of genome altered (Table S1). With regards to biological behavior (Fig. 2D), pathways involved in humoral immune response and macrophage proliferation were activated in subtype3, which might cause the longer survival time for subtype3 patients. Whereas patients in subtype2 showed a lower immune activity,

which might lead to the worst survival outcome for subtype2 patients. Subsequent immune cell infiltration analysis shown that subtype3 patients was obviously infiltrated by naive B cell, memory resting CD4 T cells and CD8 T cells, while subtype2 patients showed an obvious increase trend in infiltration by M0 macrophages and M2 macrophages (Fig. 2E). Finally, the distribution of these three subtypes and other established breast cancer molecular subtypes was compared. The result demonstrated that subtype3 patients were mainly concentrated in the Luminal A subtypes with the best prognosis, while subtype2 patients were mainly concentrated in the C4 subtypes with the worst prognosis (Fig. 2F, Table S1).

3.3 Construction and validation of the IRGP prognostic signature

The IRGPs in the risk model were selected by applying LASSO Cox regression analysis. All samples in TCGA cohort were regrouped into training and internal validation groups randomly for prognostic analyses. No significant difference was observed when comparing patient characteristics between the two groups (Table S2). The IRGP prognostic signature consisting of 33 IRGPs was generated through the LASSO model (Table 2). The 33 IRGPs grouped BRCA patients into high- and low-risk groups based on the median value of risk score. In training, internal validation and external validation cohorts, Kaplan-Meier curves indicated that patients in low-risk group have significantly longer survival time than that in high-risk group (Fig. 3A-C). Figure 3A shown the time-dependent ROC curve of the IRGP prognostic signature in the training cohort. The area under the curve (AUC) for 1-year, 3-years, and 5-years OS predictions were 0.932, 0.859, and 0.817, respectively. The IRGP prognostic signature also possessed a high specificity and sensitivity in internal validation cohort (with AUC of 0.789, 0.732 and 0.746 for 1-, 3- and 5-year OS prediction) and external validation cohort (with AUC of 0.840, 0.652 and 0.638 for 1-, 3- and 5-year OS prediction) (Fig. 3B-C).

Table 2
Genes of IRGP prognostic signature

Gene	Coef
ACO1|CCL24	-1.53461
ADCYAP1R1|SLIT2	0.159353
AMH|IGHA2	0.378844
BMP6|NMB	0.178603
CAT|ROBO3	-2.92446
CCL24|LTBP4	0.031074
CCL24|NFKBIE	2.32E-14
CCL24|PIK3R2	0.087811
CCL24|RABEP2	1.25E-09
CCL24|THRA	2.87E-14
CHGA|PCSK2	-0.26989
FABP5|FGF7	-0.42612
FABP7|LALBA	-0.0183
FGF16|IGHA1	2.434383
IGHA1|LRP1	-0.08221
IGHA1|NR2F6	-0.09532
IGHA1|PLTP	-0.124
IGHA1|RBP2	-0.00955
IGHA1|TNFRSF9	-4.39E-12
IGHA2|PENK	-0.2354
IGHE|TNFRSF18	0.007074
IGHE|TOR2A	0.854685
IL1RL2|MIA	0.039159
IL1RL2|THRA	1.16766
IL27|KLRC1	0.009705
IL27|TSHR	0.092078

Abbreviation: IRGP, immune-relevant gene pair.

Gene	Coef
IL31RA|TGFBR3	0.863127
INHBA|PIK3R2	0.094279
LGR4|MASP2	-1.10075
LGR4|NPPA	-0.28221
LTB|SCTR	-1.28011
OASL|PRLH	-1.27046
PTGDR2|PTN	0.226948
Abbreviation: IRGP, immune-relevant gene pair.	

Univariate Cox regression analysis showed a statistical significance for risk score determined by IRGP prognostic signature (Table 3). In order to explore the independence of IRGP prognostic signature in survival prediction, a multivariate Cox regression analysis was performed in the TCGA training cohort and internal validation cohort, including risk score, age, M stage, N stage, T stage, tumor stage, immune subtype, and BRCA subtypes. Whereas only age, M stage, N stage and T stage were available for GEO external validation cohort, we integrated risk score and these clinical features in multivariate cox regression analysis. The prognostic values of IRGP prognostic signature was significant compared with other clinical characteristics in all three cohorts, and indicated that IRGP prognostic signature was a strong independent risk factor (Table 3). The predictive power of the IRGP prognostic signature was further tested in various subgroups stratified by TNM stage, age, fraction genome altered, sample initial weight, ER status, HER2 status, PR status, tumor site, immune subtype, and BRCA subtype in the TCGA entire cohort. The forest plot shown that in almost all subgroups, the higher the risk score, the more obvious the patients with the worse prognosis (Fig. 4).

Table 3
Univariate and multivariate survival analyses of the IRGP prognostic signature and clinical variables.

	Univariate analysis			Multivariate analysis		
Variables	HR	95% CI	P-value	HR	95% CI	P-value
Training cohort (764 patients)						
riskScore	6.10	4.65–8.00	5.94E-39	8.42	5.66–12.52	7.50E-26
Age	1.03	1.01–1.05	1.96E-04	1.01	0.99–1.03	4.10E-01
M stage (vs. M stage0)						
M1 stage	4.90	2.60–9.24	8.77E-07	0.17	0.02–1.46	1.07E-01
N stage (vs. N stage0)						
N1 stage	1.37	0.85–2.21	1.89E-01	0.83	0.43–1.59	5.71E-01
N2 stage	2.17	1.13–4.15	1.92E-02	0.86	0.29–2.53	7.79E-01
N3 stage	3.71	1.75–7.89	6.40E-04	1.87	0.6–5.87	2.81E-01
T stage (vs. T stage1)						
T2 stage	1.29	0.78–2.11	3.20E-01	0.55	0.25–1.20	1.33E-01
T3 stage	1.20	0.6–2.4	5.99E-01	0.56	0.20–1.54	2.58E-01
T4 stage	4.15	1.97–8.72	1.76E-04	0.25	0.08–0.84	2.55E-02
Tumor stage (vs. Stage I)						
Stage II	1.43	0.74–2.77	0.291923	2.30	0.77–6.89	0.134892
Stage III	2.54	1.26–5.13	0.009344	3.22	0.76–13.65	0.111932
Stage IV	18.78	7.91–44.61	2.99E-11	269.64	19.39–3749.84	3.08E-05
Immune subtype (vs. C1)						
C2	1.26	0.78–2.03	0.33989	1.11	0.62–1.99	0.717914
C3	1.42	0.80–2.52	0.226986	1.87	0.91–3.86	0.090167
C4	2.38	1.18–4.79	0.015639	0.53	0.22–1.25	0.146363
C6	0.78	0.19–3.28	0.734812	1.24	0.27–5.78	0.781076
BRCA subtype (vs. Normal-like)						
Luminal A	0.27	0.17–0.45	5.21E-07	0.37	0.2–0.7	0.002306

Abbreviation: IRGP, immune-relevant gene pair.

	Univariate analysis			Multivariate analysis		
Luminal B	0.50	0.27–0.9	0.021979	0.42	0.2–0.86	0.01837
HER2-enriched	0.34	0.13–0.89	0.02762	0.18	0.06–0.54	0.002221
Basal-like	0.44	0.24–0.83	0.011065	0.19	0.07–0.52	0.001287
Internal validation cohort (325 patients)						
riskScore	1.88	1.22–2.88	0.003927	2.74	1.40–5.39	0.003356
Age	1.04	1.02–1.07	0.000203	1.06	1.03–1.09	4.65E-05
M stage (vs. M stage0)						
M1 stage	5.02	1.95–12.91	0.000832	9.43	0.30–301.15	0.204258
N stage (vs. N stage0)						
N1 stage	3.16	1.5–6.65	0.002512	4.06	1.27–13.02	0.018388
N2 stage	3.60	1.45–8.98	0.005954	3.37	0.43–26.12	0.24478
N3 stage	5.09	1.89–13.73	0.001287	2.19	0.25–19.11	0.478551
T stage (vs. T stage1)						
T2 stage	1.50	0.66–3.4	0.329302	1.26	0.29–5.53	0.755855
T3 stage	2.90	1.18–7.1	0.019906	0.96	0.13–6.89	0.96494
T4 stage	4.22	1.26–14.1	0.019263	0.57	0.06–5.46	0.627337
Tumor stage (vs. Stage I)						
Stage II	1.63	0.61–4.38	0.332773	1.00	0.13–7.56	0.999351
Stage III	3.50	1.27–9.63	0.01517	2.71	0.16–46.98	0.493153
Stage IV	8.92	2.56–31.08	0.000595	NA	NA	NA
Immune subtype (vs. C1)						
C2	0.70	0.36–1.39	0.313671	0.77	0.31–1.92	0.570762
C3	0.57	0.24–1.39	0.217516	1.22	0.33–4.42	0.766659
C4	1.18	0.35–4.06	0.788934	3.00	0.67–13.50	0.152069
C6	0.69	0.16–3.01	0.623159	1.31	0.21–7.97	0.772321
BRCA subtype (vs. Normal-like)						
Luminal A	0.35	0.17–0.71	0.003892	0.47	0.17–1.32	0.152461

Abbreviation: IRGP, immune-relevant gene pair.

	Univariate analysis			Multivariate analysis		
Luminal B	0.45	0.17–1.24	0.124116	0.16	0.04–0.58	0.005464
HER2-enriched	1.39	0.5–3.84	0.522492	3.18	0.59–17.19	0.178663
Basal-like	0.28	0.1–0.78	0.015176	0.30	0.07–1.24	0.096199
External validation cohort (327 patients)						
riskScore	1.60	1.25–2.05	0.000232	1.75	1.32–2.33	0.000103
Age	0.99	0.97–1.01	0.483386	0.99	0.97–1.01	0.460612
M stage (vs. M stage0)						
M1 stage	5.20	2.39–11.33	3.22E-05	1.26	0.41–3.9	0.683614
N stage (vs. N stage0)						
N1 stage	2.40	1.24–4.66	0.009625	2.53	1.29–4.98	0.006904
N2 stage	5.10	2.74–9.48	2.77E-07	4.82	2.43–9.56	6.98E-06
N3 stage	5.10	2.55–10.23	4.38E-06	5.28	2.38–11.72	4.35E-05
T stage (vs. T stage1)						
T2 stage	1.14	0.66–1.94	0.642067	0.71	0.4–1.26	0.237128
T3 stage	4.80	2.44–9.43	5.29E-06	1.73	0.78–3.82	0.177103
T4 stage	4.43	1.95–10.08	0.000389	1.35	0.43–4.23	0.603462
Abbreviation: IRGP, immune-relevant gene pair.						

3.4 Landscape of mutation profiles in breast cancer

The somatic mutation profiles of 968 BRCA patients with 4 types of data based on different processing software were downloaded from TCGA. Besides, we visualized the results of 442 high-risk and 526 low-risk BRCA patients using the " maftools" package, based on mutation data with VCF format. The result shown that the frequency of somatic mutation in high risk BRCA patients was slightly higher than low risk BRCA patients. The waterfall plot shows the mutation information for each gene in each sample, and the different types of mutations are annotated by different colors at the bottom (Fig. 5A, Fig. S2A). In summary, these mutations were further classified based on the different taxonomic categories, among which missense mutation accounted for the majority (Fig. 5B, Fig. S2B), only single nucleotide polymorphism occurred (Fig. 5C, Fig. S2C), and C > T was the most common of single nucleotide variants (SNV) in BRCA patients (Fig. 5D, Fig. S2D). Besides, the number of altered bases in each sample was counted and mutation type with different colors in box plot for BRCA patients were illustrated (Fig. 5E-F, Fig. S2E-F). Last, we exhibited the top 10 mutated genes with ranked percentages, including PIK3CA

(28%), TP53 (29%), TTN (14%), MUC16 (8%), MUC4 (7%), KMT2C (5%), RYR3 (5%), DMD (4%), CDH1 (4%), USH2A (4%) in high risk BRCA patients and PIK3CA (30%), TP53 (25%), TTN (11%), MUC16 (6%), MUC4 (6%), CDH1 (5%), KMT2C (4%), USH2A (4%), RYR2 (3%), NEB (3%) in low risk BRCA patients (Fig. 5G, Fig. S2G). The overall distribution of six different conversions in summarized boxplot (Fig. 5H, Fig. S2H) and the fraction of conversions in each sample in stacked barplot (Fig. 5I, Fig. S2I) shown C > T was the most common of SNV in BRCA patients. And each single nucleotide variant was classified into either one of the four types of transversions or two types of transitions. The single nucleotide transitions were more common than the single nucleotide transversions (Fig. 5J, Fig. S2J).

3.5 Establishment of a nomogram predicting OS in BRCA patients

In order to develop a clinically relevant quantitative method to predict the mortality rate of patients, we constructed a nomogram integrating IRGP prognostic signature derived score and clinical information to predict survival probability of BRCA patients in TCGA cohort and GSE20685 cohort (Fig. 6A and Fig.S3A). The concordance index (C-index) of the nomogram was 0.862 in TCGA cohort and 0.758 in GSE20685 cohort, which indicating a good discriminatory ability of the nomogram. The calibration plots showed that the derived nomograms performed well compared to the performance of an ideal model at 1-year, 3-years and 5-years (Fig. 6B-D and Fig. S3B-D). Decision curve analysis shows that the clinical utility of nomograms greatly exceeds TNM staging system (Fig. 6E and Fig. S3E).

3.6 Identification of IRGP prognostic signature related biological pathways and processes

The relationship between IRGP prognostic signature derived scores of clinical characteristics and molecular subtypes were further investigated in the entire TCGA cohorts (Fig. 7A). In terms of clinical features, IRGP prognostic signature was increased in patients with more advanced TNM stage and patients who had died due to the disease. Furthermore, while age influenced the value of IRGP prognostic signature, that value of IRGP prognostic signature didn't varied between sample initial weight and tumor site. In terms of molecular characteristics, it is explored that more genomic changes relative to higher values of IRGP prognostic signature, and patients in molecular subtypes C4 and Luminal B exhibited significantly higher values of IRGP prognostic signature than others.

Next, we performed GSEA to elucidate the biological functions of the IRGP prognostic signature (Fig. 7B), which revealed that genes highly expressed in the low- IRGP risk score group were associated with immune-related gene set, including regulation of lymphocyte activation, phagocytosis and immune response regulating cell surface receptor signaling pathway. While the high- IRGP-related genes showed significant enrichment in multiple biological processes such as the DNA activity-related gene set, including helicase activity, catalytic activity acting on DNA, ATP dependent 5_3 DNA helicase activity. Immune cell infiltration analysis indicated that patients with high IRGP prognostic signature scores seemed to infiltrated by M0 and M2 macrophages, while the patients in low-risk group were infiltrated by naive B cell, memory resting CD4 T cells and CD8 T cells (Fig. 7C).

4. Discussion

Immunity is closely linked to tumors. For example, Li B et al inferred the abundance of 6 immune cell types (B cells, CD4 T cells, CD8 T cells, neutrophils, macrophages, and dendritic cells), and discovered significant associations between cell abundance and prognosis in 23 cancers. For example, in addition to prolonging the survival of patients, CD8 T cells may also play an important role in preventing tumor recurrence, such as in melanoma, colorectal cancer, and cervical cancer [20]. In addition, increasing evidence points to the importance of biomarkers (especially genes) in determining cancer outcomes, which provides new opportunities for integrating this information into treatment algorithms [10]. Many previous studies have shown that immune-related genes can be used as prognostic indicators for breast cancer. However, researches were often limited by the singularity of genes and differences between samples [21, 22]. Therefore, new methods are urgently needed to improve the accuracy of breast cancer prediction.

Immune-related gene pairs were widely used in tumor analysis, and great progress has been made in many cancers, such as melanoma, ovarian cancer and pancreatic cancer [23–25]. However, further researches were needed in breast cancer, and this study would fill the gap in this regard. Our study focused on predicting BRCA survival using the prognostic characteristics of IRGP composed of 33 IRGP combined with clinical information. We believed that our research would provide a new perspective for clinical decision-making and prognostic monitoring in breast cancer.

The study was originally designed to use a variety of bioinformatics tools and databases to demonstrate that 33-IRGPs features can be applied to the clinical prognosis of patients with BRCA. On this basis, in this study, we established 33-IRGPs characteristics to predict the individual prognostic characteristics of BRCA through univariate Cox regression analysis and LASSO model. In order to verify this feature, we studied the accuracy and efficiency of KM curve, ROC curve analysis, univariate Cox regression and multivariate Cox regression analysis in predicting the prognosis of BRCA patients. In order to develop a clinically related quantitative method for predicting the probability of death, we constructed a line chart combining IRGP prognostic characteristic derived scores and clinical information to represent the survival rate of BRCA patients in the TCGA cohort and validation in the GSE20685 cohort. The consistency analysis index shows that the line chart has good discriminant ability. Nomogram analysis shows that the clinical effectiveness of the IRGP prognostic signature is significantly higher than that of the TNM stage, which implied that IRGPs ware strong independent indicators of prognosis prediction. GSEA showed that low IRGP related genes were related to immune related gene sets, including regulating lymphocyte activation, phagocytosis and immune response regulating cell surface receptor signal pathways, while high IRGP related genes were significantly enriched in the DNA activity-related gene set, including helicase activity, catalytic activity acting on DNA, atp dependent 5_3 DNA helicase activity. DNA helicase plays a significant role in cancer. Not only are there single-gene helicase diseases that have a strong propensity for cancer, but it is also well known that helicase variants are related to specific cancers such as breast cancer. At the same time, DNA helicase is usually overexpressed in cancer tissues, and the reduction of helicase gene expression leads to inhibited proliferation and growth of cancer cells, as well

as induction of DNA damage and apoptosis. The important role of helicases in DNA damage and replication stress response and DNA repair pathways confirms their vitality in cancer biology and suggests their potential value in anti-cancer therapy[26].

At the same time, we also analyzed the somatic mutations in patients with breast cancer, and we found that most of the missense mutations occurred in the following genes: Last, we exhibited the top 10 mutated genes with ranked percentages, including PIK3CA (28%), TP53 (29%), TTN (14%), MUC16 (8%), MUC4 (7%), KMT2C (5%), RYR3 (5%), DMD (4%), CDH1 (4%), USH2A (4%) in high risk BRCA patients and PIK3CA (30%), TP53 (25%), TTN (11%), MUC16 (6%), MUC4 (6%), CDH1 (5%), KMT2C (4%), USH2A (4%), RYR2 (3%), NEB (3%) in low risk BRCA patients. The PIK3CA gene encodes the catalytic subunit of phosphatidylinositol 3-kinase (PI3K) and is among the most frequently mutated genes in solid tumor malignancies. Cancer-associated mutations in PIK3CA promote signaling via the PI3K pathway and stimulate tumor cell growth [27]. Through previous studies, we have learned that an increased prevalence of PIK3CA mutations in women with CRC, and an increased risk of recurrence and poorer prognosis associated with PIK3CA mutations [28–30]. The TP53 protein is a DNA-bound transcription factor that has the potential to bind to hundreds of different promoter elements in the genome to regulate gene expression[31]. Tumors with TP53 mutations are usually characterized by poor differentiation, increased invasiveness and high metastatic potential, which are associated with poor prognosis [32, 33]. Combined with our conclusions about somatic mutation analysis, the mutations of PIK3CA and TP53 affect the prognosis of patients with breast cancer to a great extent. So we can determine the reliability of our functional enrichment results [31, 34] and our IRGPs may play some role in the prognosis of breast cancer.

At present, several immune-related therapies for tumors have achieved good results in clinical trials. For instance, by blocking macrophage chemokines (such as CXCL12) and preventing macrophages from entering tumors, the development and proliferation of cancer cells can be inhibited [35]. Although the focus of this research is not on the mechanism of immune cells, it still provides strong evidence that tumor-related immune genes may become potential targets for cancer treatment. Our research focuses on immune-related genes and uses strict standard-level screening to obtain genes that may be prognostic targets for BRCA [6].

The current study has some limitations. First, due to the retrospective nature of this study, the patient population was heterogeneous. Secondly, because gene expression data are required for Cox regression as categorical variables, the optimal cut-off value needs to be further verified in future studies. In conclusion, the IRGPS gene map is a powerful tool for predicting breast cancer survival and guiding treatment. In addition, prospective clinical trials are needed to validate our findings.

Declarations

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The raw data involved in the current study are publicly available in TCGA (https://portal.gdc.cancer.gov/) database and GEO (https://www.ncbi.nlm.nih.gov/geo/) database.

Author Contributions

All authors contributed substantially to the preparation of this manuscript. B Zhou, S. X. Guo, S. H. Lai, L. P. Zeng and Y Zhou were responsible for protocol design. B Zhou, L. Y. Zhang, S. X. Guo, C. Q. Pu, S. H. Lai and L. P, Zeng were responsible for data acquisition. Q. Y. Wang, W. W. Li, Z. B. Zhou, Y. X. Chen and Y. Zhou were responsible for data analysis. All authors interpreted the data. L. Y. Zhang, S. X. Guo, C. Q. Pu, H. M. Zhang and L. P. Zeng wrote the manuscript. All authors revised and finalized the manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Cedolini C, Bertozzi S, Londero AP, Bernardi S, Seriau L, Concina S, Cattin F, Risaliti A: **Type of breast cancer diagnosis, screening, and survival**. *Clin Breast Cancer* 2014, **14**(4):235–240.
2. Siegel RL, Miller KD, Jemal A: **Cancer Statistics, 2017**. *CA Cancer J Clin* 2017, **67**(1):7–30.
3. Edge SB, Byrd DR, Compton CC, Fritz AG, Trotti A: **American Joint Committee on Cancer (AJCC) Cancer Staging Manual**. New York: Springer; 2010.
4. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, Meyer L, Gress DM, Byrd DR, Winchester DP: **The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging**. *CA: a cancer journal for clinicians* 2017, **67**(2):93–99.
5. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, Jemal A, Siegel RL: **Breast cancer statistics, 2019**. *CA: a cancer journal for clinicians* 2019, **69**(6):438–451.

6. Li J, Liu C, Chen Y, Gao C, Wang M, Ma X, Zhang W, Zhuang J, Yao Y, Sun C: **Tumor Characterization in Breast Cancer Identifies Immune-Relevant Gene Signatures Associated With Prognosis**. *Front Genet* 2019, **10**:1119.
7. Elinav E, Nowarski R, Thaiss CA, Hu B, Jin C, Flavell RA: **Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms**. *Nat Rev Cancer* 2013, **13**(11):759–771.
8. Topalian SL, Taube JM, Anders RA, Pardoll DM: **Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy**. *Nat Rev Cancer* 2016, **16**(5):275–287.
9. Nagarsheth N, Wicha MS, Zou W: **Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy**. *Nat Rev Immunol* 2017, **17**(9):559–572.
10. Li B, Cui Y, Diehn M, Li R: **Development and Validation of an Individualized Immune Prognostic Signature in Early-Stage Nonsquamous Non-Small Cell Lung Cancer**. *JAMA Oncol* 2017, **3**(11):1529–1537.
11. Balbin OA, Malik R, Dhanasekaran SM, Prensner JR, Cao X, Wu YM, Robinson D, Wang R, Chen G, Beer DG *et al*: **The landscape of antisense gene expression in human cancers**. *Genome Res* 2015, **25**(7):1068–1079.
12. Kuhn M: **Caret: Classification and regression training**. 2013, **1**.
13. Kao KJ, Chang KM, Hsu HC, Huang AT: **Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization**. *BMC Cancer* 2011, **11**:143.
14. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M *et al*: **NCBI GEO: archive for functional genomics data sets–update**. *Nucleic Acids Res* 2013, **41**(Database issue):D991-995.
15. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP: **Maftools: efficient and comprehensive analysis of somatic variants in cancer**. *Genome Res* 2018, **28**(11):1747–1756.
16. Duan J, Soussen C, Brie D, Idier J, Wan M, Wang YP: **Generalized LASSO with under-determined regularization matrices**. *Signal Processing* 2016, **127**:239–246.
17. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: **Robust enumeration of cell subsets from tissue expression profiles**. *Nat Methods* 2015, **12**(5):453–457.
18. Hänzelmann S, Castelo R, Guinney J: **GSVA: gene set variation analysis for microarray and RNA-seq data**. *BMC Bioinformatics* 2013, **14**:7.
19. 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):15545–15550.
20. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, Jiang P, Shen H, Aster JC, Rodig S *et al*: **Comprehensive analyses of tumor immunity: implications for cancer immunotherapy**. *Genome Biol* 2016, **17**(1):174.
21. Zhang Y, Dong X, Wang Y, Wang L, Han G, Jin L, Fan Y, Xu G, Yuan D, Zheng J *et al*: **Overexpression of LncRNA BM466146 Predicts Better Prognosis of Breast Cancer**. *Front Oncol* 2020, **10**:628757.

22. Huang Z, Xiao C, Zhang F, Zhou Z, Yu L, Ye C, Huang W, Li N: **A Novel Framework to Predict Breast Cancer Prognosis Using Immune-Associated LncRNAs**. *Front Genet* 2020, **11**:634195.
23. Huang RZ, Mao M, Zheng J, Liang HQ, Liu FL, Zhou GY, Huang YQ, Zeng FY, Li X: **Development of an immune-related gene pairs index for the prognosis analysis of metastatic melanoma**. *Sci Rep* 2021, **11**(1):1253.
24. Zhang B, Nie X, Miao X, Wang S, Li J, Wang S: **Development and verification of an immune-related gene pairs prognostic signature in ovarian cancer**. *J Cell Mol Med* 2021.
25. Li Y, Yao P, Zhao K, Ye Z, Zhang H, Cao J, Zhang S, Xing C: **Individualized prognostic signature for pancreatic carcinoma validated by integrating immune-related gene pairs (IRGPs)**. *Bioengineered* 2021, **12**(1):88–95.
26. Dhar S, Datta A, Brosh RM, Jr.: **DNA helicases and their roles in cancer**. *DNA Repair (Amst)* 2020, **96**:102994.
27. Cai Y, Yousef A, Grandis JR, Johnson DE: **NSAID therapy for PIK3CA-Altered colorectal, breast, and head and neck cancer**. *Adv Biol Regul* 2020, **75**:100653.
28. Benvenuti S, Frattini M, Arena S, Zanon C, Cappelletti V, Coradini D, Daidone MG, Pilotti S, Pierotti MA, Bardelli A: **PIK3CA cancer mutations display gender and tissue specificity patterns**. *Hum Mutat* 2008, **29**(2):284–288.
29. Kato S, Iida S, Higuchi T, Ishikawa T, Takagi Y, Yasuno M, Enomoto M, Uetake H, Sugihara K: **PIK3CA mutation is predictive of poor survival in patients with colorectal cancer**. *Int J Cancer* 2007, **121**(8):1771–1778.
30. Miyaki M, Iijima T, Yamaguchi T, Takahashi K, Matsumoto H, Yasutome M, Funata N, Mori T: **Mutations of the PIK3CA gene in hereditary colorectal cancers**. *Int J Cancer* 2007, **121**(7):1627–1630.
31. Silwal-Pandit L, Langerød A, Børresen-Dale AL: **TP53 Mutations in Breast and Ovarian Cancer**. *Cold Spring Harb Perspect Med* 2017, **7**(1).
32. Olivier M, Hollstein M, Hainaut P: **TP53 mutations in human cancers: origins, consequences, and clinical use**. *Cold Spring Harb Perspect Biol* 2010, **2**(1):a001008.
33. Silwal-Pandit L, Vollan HK, Chin SF, Rueda OM, McKinney S, Osako T, Quigley DA, Kristensen VN, Aparicio S, Børresen-Dale AL *et al*: **TP53 mutation spectrum in breast cancer is subtype specific and has distinct prognostic relevance**. *Clin Cancer Res* 2014, **20**(13):3569–3580.
34. Mosele F, Stefanovska B, Lusque A, Tran Dien A, Garberis I, Droin N, Le Tourneau C, Sablin MP, Lacroix L, Enrico D *et al*: **Outcome and molecular landscape of patients with PIK3CA-mutated metastatic breast cancer**. *Ann Oncol* 2020, **31**(3):377–386.
35. Gholamin S, Mitra SS, Feroze AH, Liu J, Kahn SA, Zhang M, Esparza R, Richard C, Ramaswamy V, Remke M *et al*: **Disrupting the CD47-SIRPα anti-phagocytic axis by a humanized anti-CD47 antibody is an efficacious treatment for malignant pediatric brain tumors**. *Sci Transl Med* 2017, **9**(381).

Figures

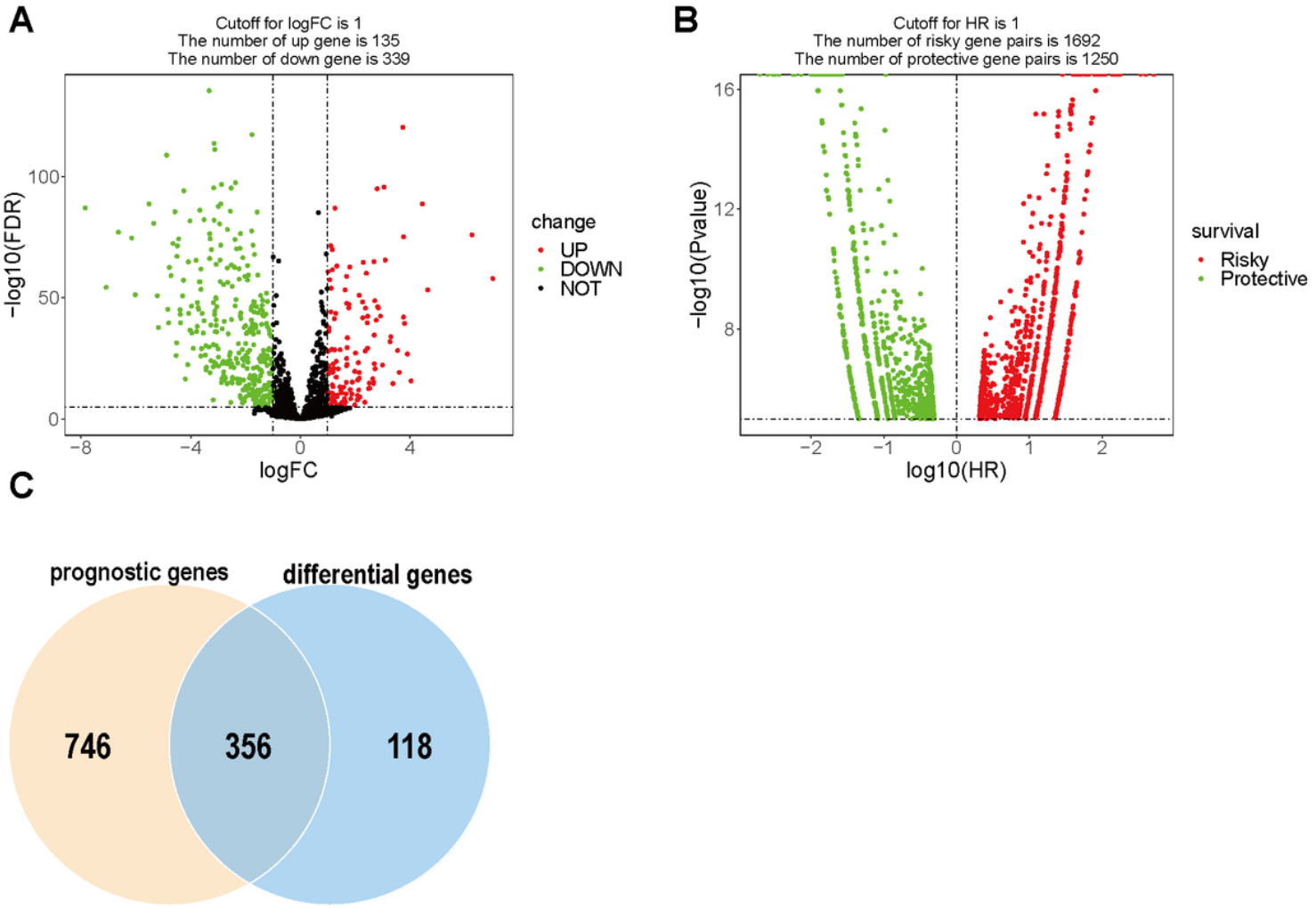

Figure 1

Candidate IRGPs identification. (A) Volcano plot for differentially expressed IRGs. (B) Volcano plot for prognostic IRGPs. (C) Venn diagram comparing differentially expressed IRGs and prognostic IRGPs. IRGs, immune-relevant genes; IRGPs, immune-relevant gene pairs.

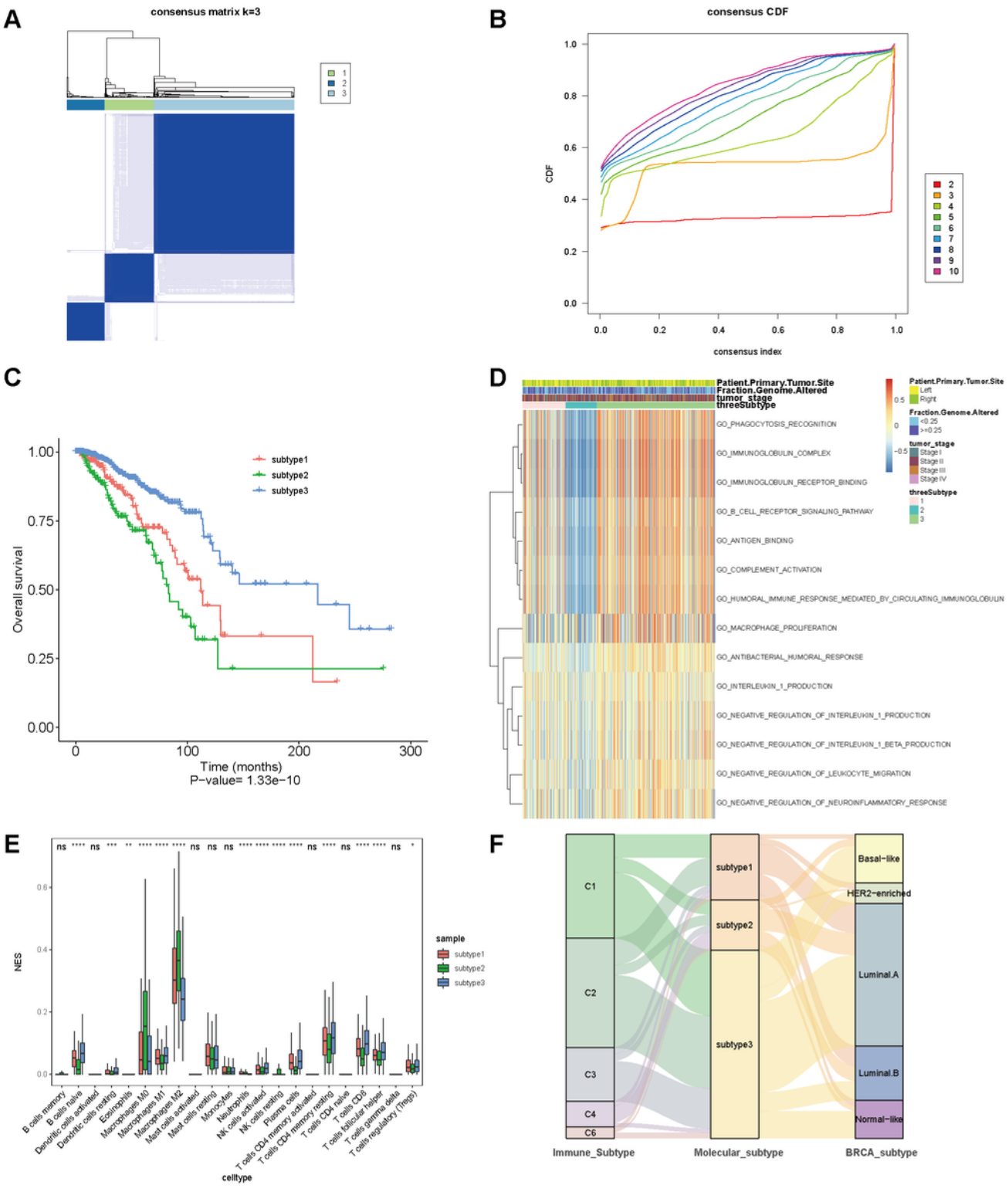

Figure 2

Consensus clustering of IRGs in BRCA. (A) Consensus matrices of BRCA patients for k=3; (B) Consistent cumulative distribution function (CDF) graph of BRCA patients for k=3; (C) Differences in patient overall survival with three molecular subgroups; (D) Heatmap showing the activation status of the biological processes in different molecular subgroups; (E) Box plot of the comparison of immune cell infiltration between the three molecular subgroups; (F) Sankey chart displaying the distribution of the three

molecular subgroups in C1–C6 subtypes and BRCA subtypes. IRGs, immune-relevant genes; BRCA, breast cancer.

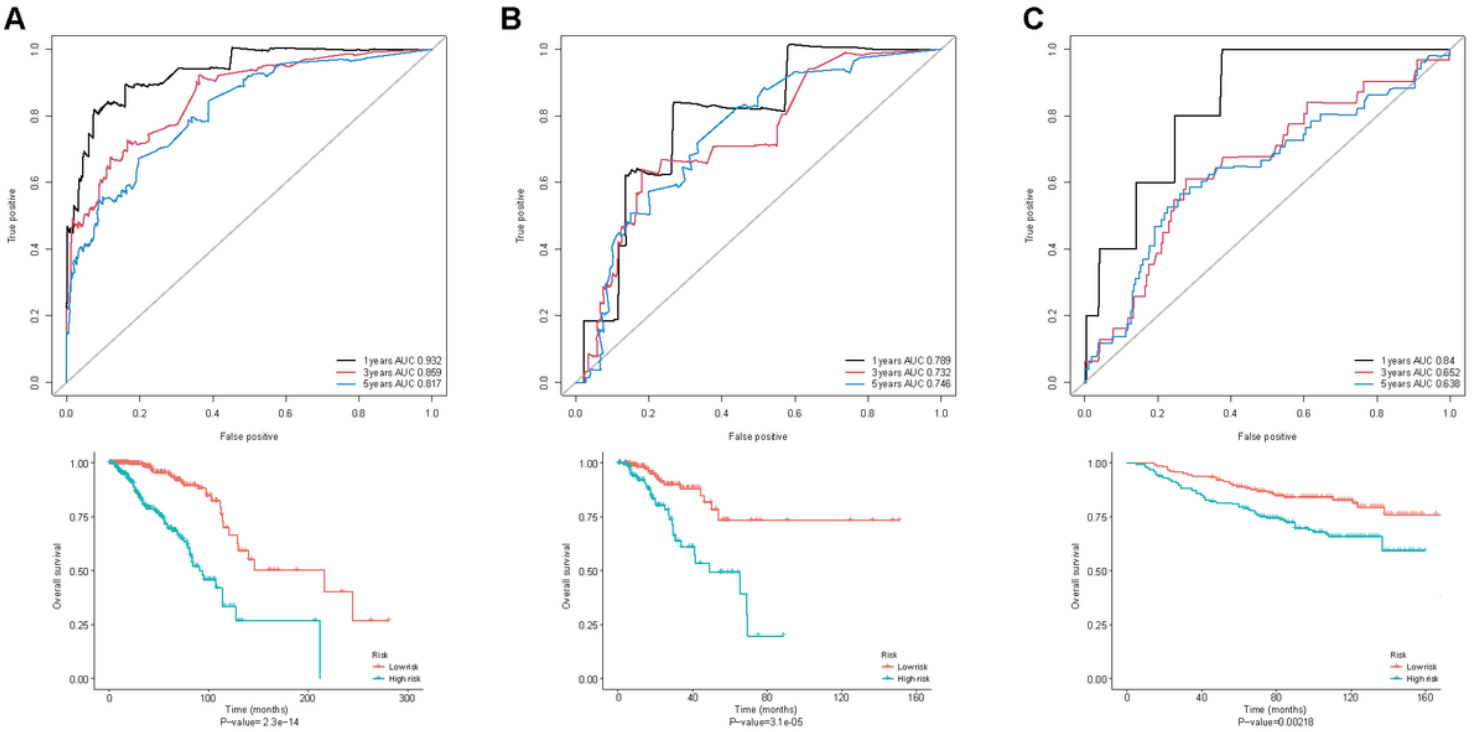

Figure 3

Construction and validation of the IRGP prognostic signature. (A–C) Kaplan–Meier curves (left) and ROC curves (right) of overall survival according to IRGP prognostic signature risk score in training cohort (A), internal validation cohort (B), and external validation cohort (C); IRGP, immune-relevant gene pair; ROC, receiver operating characteristic.

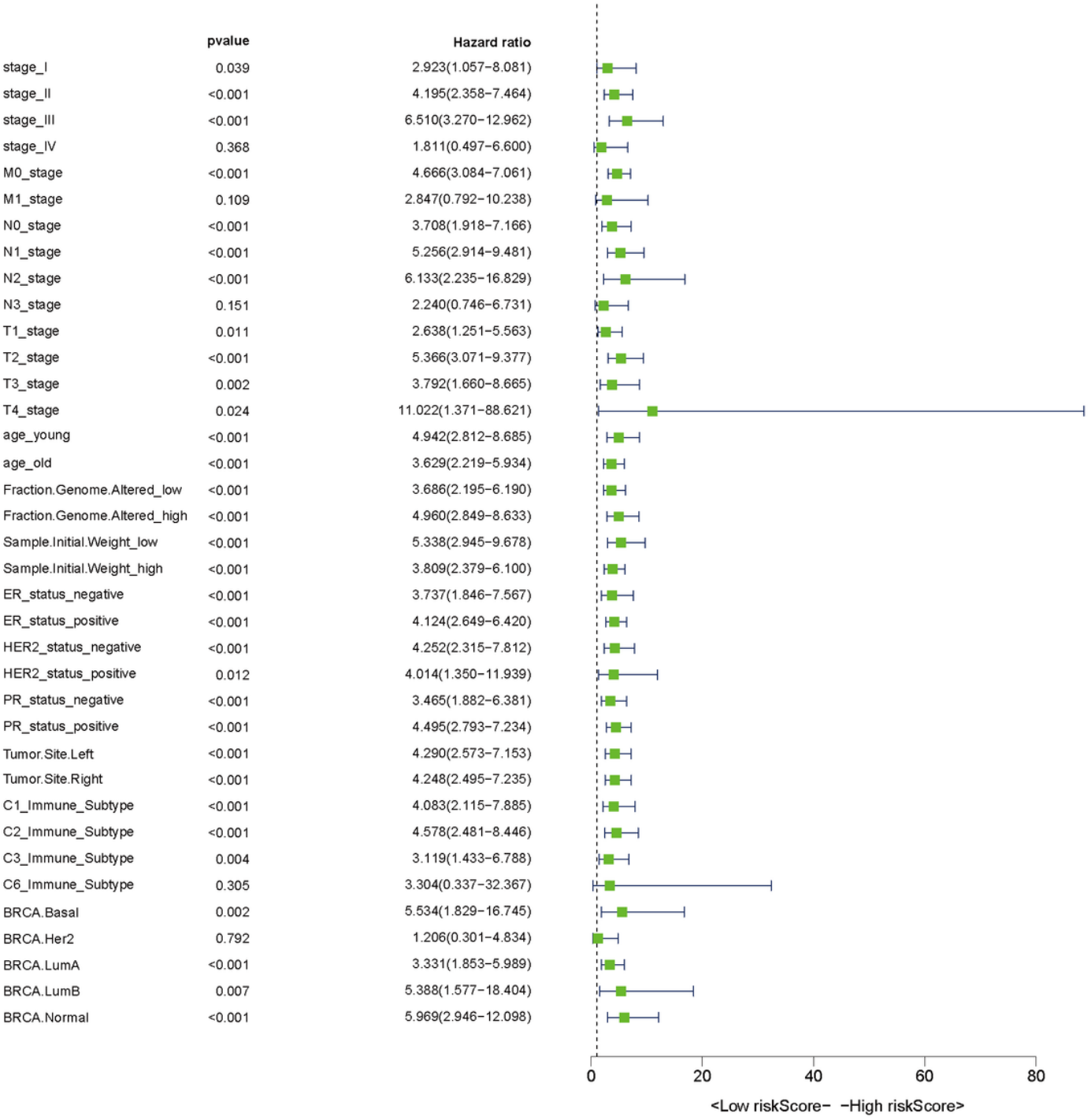

Figure 4

Forest plots of the associations between IRGP prognostic signature and overall survival in various subgroups in TCGA cohort. Unadjusted HRs (boxes) and 95% confidence intervals (horizontal lines) are depicted. IRGP, immune-relevant gene pair.

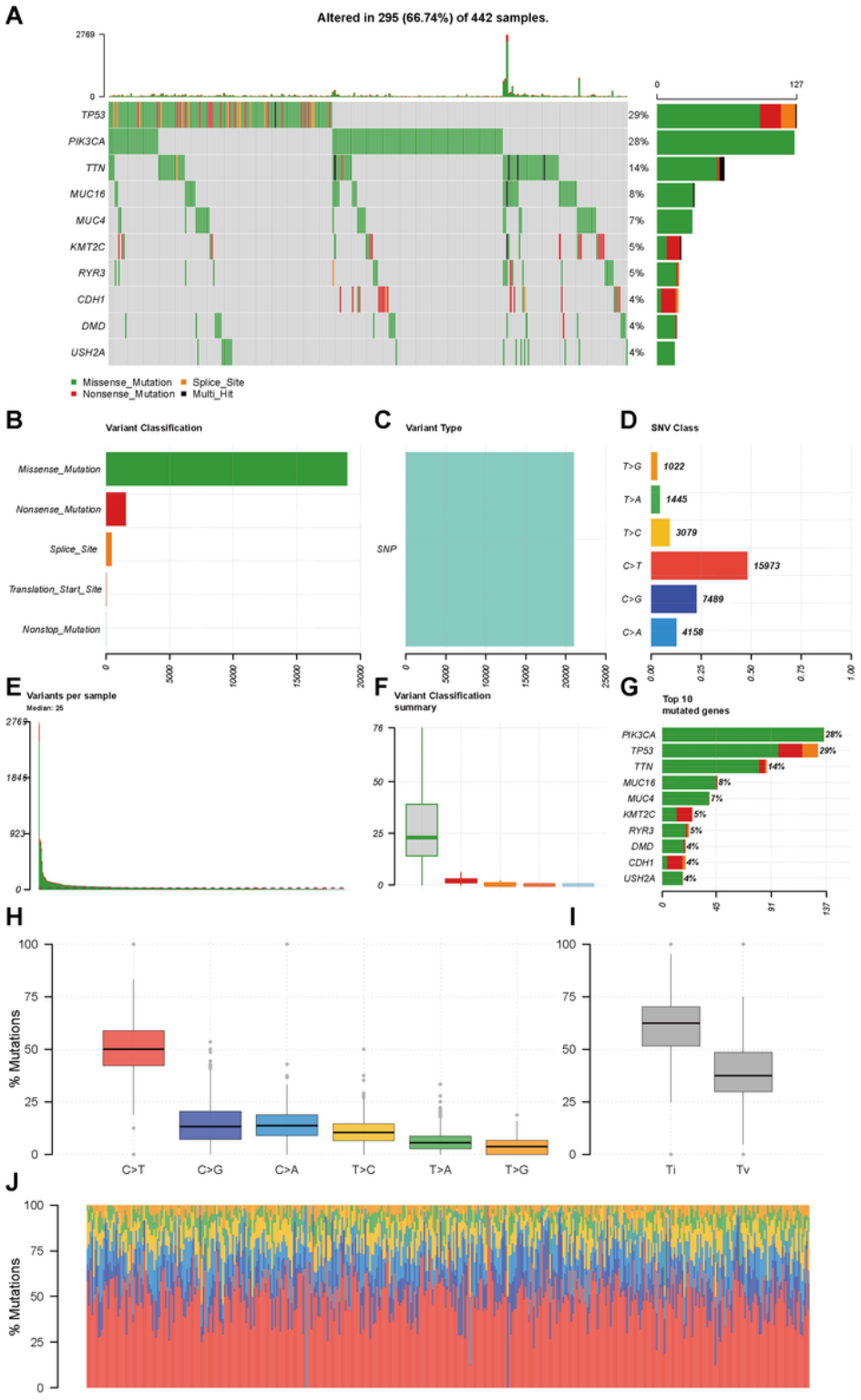

Figure 5

Landscape of mutation profiles in high risk BRCA patients. (A) Waterfall plot of mutation information for each gene in each sample, in which various colors with annotations at the bottom represented the different mutation types. The barplot above the legend exhibited the mutation burden; (B,C,D) Classification of mutation types according to different categories; (E,F) Tumor mutation burden in specific samples; (G) The top 10 mutated genes in high risk BRCA patients; (H) The boxplot of overall

distribution of six different conversions; (I) The stacked barplot of fraction of conversions in each sample; (J) The boxplot of single nucleotide transitions and single nucleotide transversions. BRCA, breast cancer; SNP, single nucleotide polymorphism; SNV, single nucleotide variants.

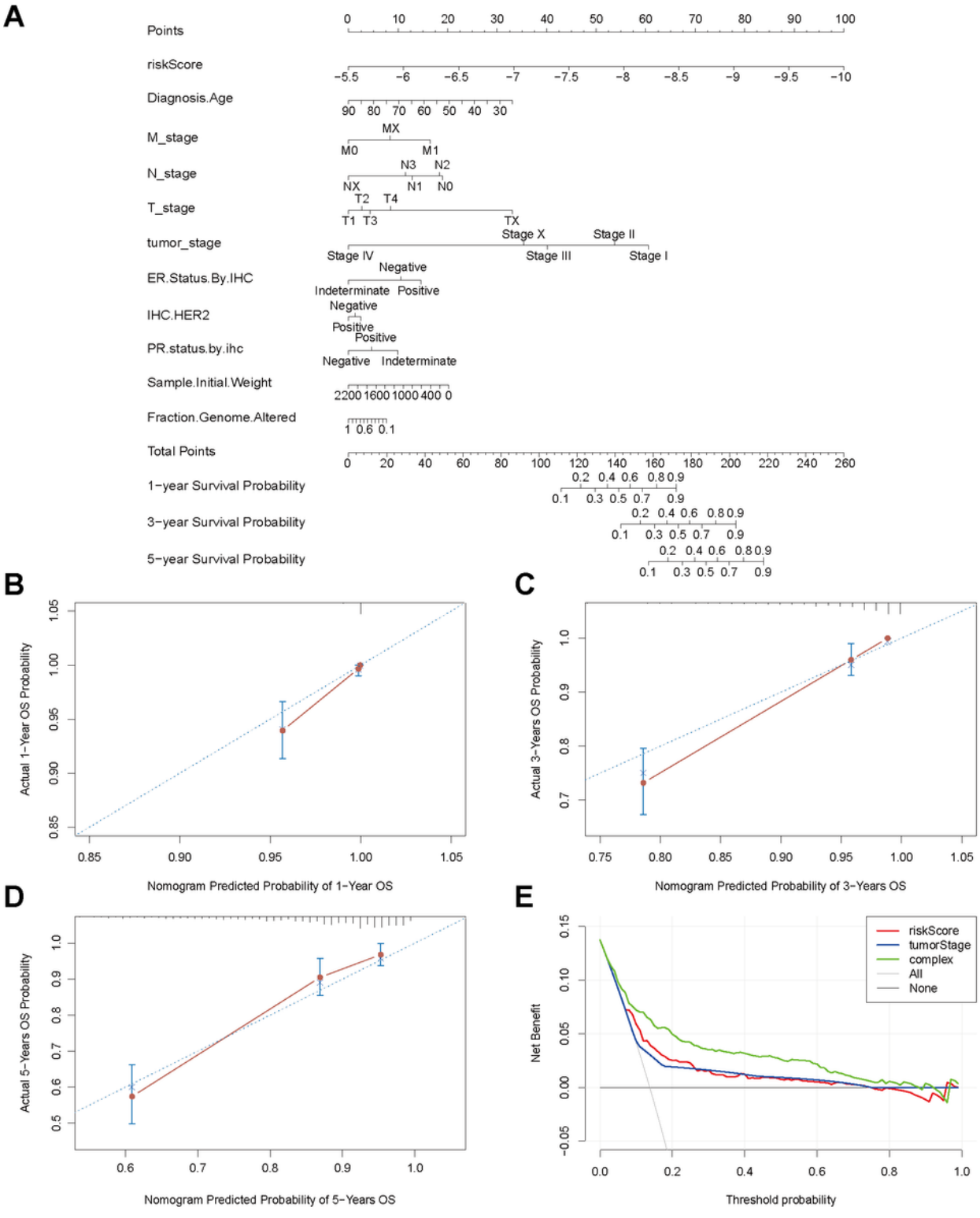

Figure 6

Establishment of a nomogram predicting OS for BRCA patients in TCGA cohort. (A) Nomograms for predicting the probability of patient mortality based on IRGP prognostic signature and clinical variables;

(B-D) Plots depict the calibration of nomograms based on IRGP prognostic signature in terms of agreement between predicted and observed 1-year (B), 3-years (C), and 5-years (D) outcomes. Nomogram performance is shown by the plot, relative to the 45-degree line, which represents the ideal prediction; (E) Decision curve analyses of the nomograms based on IRGP prognostic signature. BRCA, breast cancer; IRGP, immune-relevant gene pair.

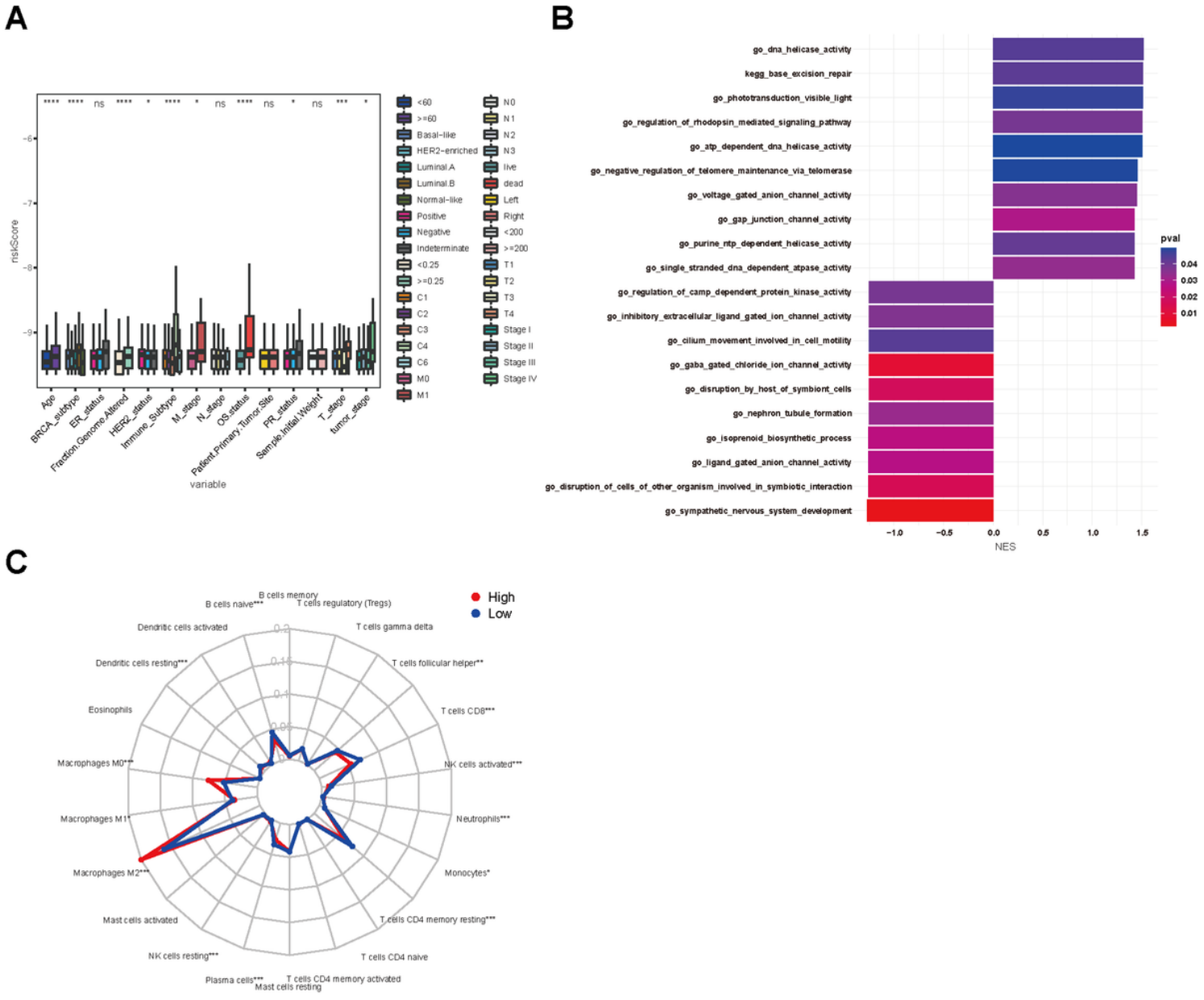

Figure 7

Clinical significance and biological function of IRGP prognostic signature. (A) IRGP prognostic signature values in different clinical subgroups. Boxes represent 25-75% of values, black lines in boxes represent median values, whiskers represent 1.5 interquartile ranges, and black dots represent outliers; (B) The

barplot of GSEA result; (C) Radar map showing the correlations between IRGP prognostic signature risk group and immune cell infiltration. IRGP, immune-relevant gene pair; GSEA, gene set enrichment analysis.

Supplementary Files

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

- Supplementaryfigureslegends.docx
- FigS1.tif
- FigS2.tif
- FigS3.tif
- TableS1.docx
- TableS2.docx