

Combination of Immune-Related Network and Subtype-Specific Analysis Define a Three- Gene Panel for Predicting Prognosis of Triple-Negative Breast Cancer

Jinguo Zhang (✉ Zhangjg2022@ustc.edu.cn)

The First Affiliated Hospital of USTC, University of Science and Technology of China

Shuaikang Pan

The First Affiliated Hospital of USTC, University of Science and Technology of China

Hongwei Jin

The First Affiliated Hospital of USTC, University of Science and Technology of China

Qingqing Sun

The First Affiliated Hospital of USTC, University of Science and Technology of China

Xinghua Han

The First Affiliated Hospital of USTC, University of Science and Technology of China

Research Article

Keywords: Triple-Negative Breast Cancer, Immune-related genes, Molecular subtypes, Prognosis, Immunotherapy

Posted Date: May 9th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1627964/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Recent breakthroughs in immune checkpoint inhibitors (ICIs) have shown promise in triple-negative breast cancer (TNBC). Due to the intrinsic heterogeneity among TNBC, clinical response to ICIs varies greatly among individuals. Thus, discovering rational biomarkers to select susceptible patients for ICIs treatment is warranted.

Methods: A total of 422 TNBC patients derived from The Cancer Genome Atlas (TCGA) database and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) dataset were included in this study. High immunogenic gene modules were identified using weighted gene coexpression network analysis (WGCNA). Immune-related genes (IRGs) expression patterns were generated by consensus clustering. We developed a three-gene signature named immune-related gene panel (IRGP) by Cox regression method. Afterward, the associations of IRGP with survival outcomes, infiltration of immune cells, drug sensitivity and the response to ICIs therapy were further explored.

Results: We found five high immunogenic gene modules. Two distinct IRGclusters and IRG-related genomic clusters were identified. The IRGP was constructed based on TAPBPL, FBP1, and GPRC5C genes. TNBC patients were subdivided into high- and low-IRGriskscore subgroups. TNBC patients with low IRGriskscore had a better survival outcome, higher infiltration of immune cells, lower TP53 mutation rate and more benefit from ICIs treatment than high IRGriskscore patients.

Conclusions: These findings offered novel insights into molecular subtype of TNBC and provided potential indicators for guiding ICIs treatment.

Background

According to estimates, female breast cancer (BC) has now displaced lung cancer as the leading cause of global cancer incidence, with an estimated 2.3 million new cases and 685,000 deaths in 2020 [1]. Among the subtypes of BC, triple-negative breast cancer (TNBC) is the most aggressive subtype that accounts for approximately 15–20% of all BC patients [2, 3]. Due to the absence of ideal treatment target, TNBC is not sensitive to endocrine therapy or molecular targeted therapy [4]. Chemotherapy is one of the primary systemic treatments of TNBC, but the efficacy of conventional postoperative adjuvant chemoradiotherapy differs widely and has considerable heterogeneity for different individuals. As high-throughput sequencing technologies have developed, more intrinsic subtypes of TNBC have been discovered and different subtypes have obviously different treatment guidelines and tumor biology [5]. Therefore, development of new treatment strategies and applicable molecular subtypes of TNBC have become an urgent challenge.

In recent years, immunotherapy has emerged as an attractive and effective pillar in cancer treatment, and increasing evidence supports a key role of the immune system in determining the response to standard therapy and long-term survival in BC patients [6, 7]. Recent studies suggested that TNBC was thought to be the most immunogenic BC subtype that characterized by the presence of tumor neoantigens and high

levels of lymphocytic infiltration [8]. According to the results of KEYNOTE-522 and IMpassion130, pembrolizumab and atezolizumab were approved by the FDA for early-stage or metastatic TNBC [9, 10]. However, TNBC patients that benefit from immune checkpoint inhibitors (ICIs) remain very limited. Furthermore, other immunotherapy strategies such as adoptive T-cell transfer, vaccination, or virotherapy have not yet demonstrated meaningful clinical activity [11]. Presently, biomarkers for predicting efficacy of immune therapy have gained increased attention. Remarkably, PD-L1 expression, tumor mutation burden and infiltrating lymphocytes (TILs) have long been proposed as reliable indicators for ICIs treatment [12]. As the factors that influence the efficacy of ICIs are multifaceted, many key questions remain unanswered. Therefore, exploring the expression patterns of immune-related pathways not only enhances our understanding of immune dysregulation but also provides potential opportunities for antitumor immunity in TNBC.

In this study, we aimed to investigate the aberrant immune network and explore a prognostic marker for predicting the immune microenvironment status in TNBC. Here, we comprehensively analyzed the immune network and identified immune-related subtypes based on the expression profiles of 422 TNBC patients. A gene panel to quantify the immune responses status for individual patients was established. Additionally, we linked our immune signature with immune cell-infiltrating characteristics and response to ICIs. Overall, our study systematically explored the immune-related subtypes and identified a robust indicator for predicting the prognosis of TNBC.

Methods

TNBC patients and public datasets

RNA sequencing data of 123 TNBC patients and 113 normal samples, somatic mutation data and corresponding clinicopathological information were obtained from The Cancer Genome Atlas (TCGA) database (www.portal.gdc.cancer.gov/projects). The mRNA expression data of METABRIC dataset on TNBC containing 299 samples and clinical information were accessed via cBioPortal (<http://www.cbioportal.org/>) [13]. Next, we extracted immune-related genes (IRGs) from ImmPort database [14] (www.immport.org/shared/home), InnateDB database [15] (www.innateDBdb.com/) and GSEA database (<http://www.gsea-msigdb.org/gsea/index.jsp>). After removal of duplicate genes, 4327 immune-related genes were acquired for further analysis. Patients without the required clinical data were excluded from all analyses.

Differential expression and enrichment analysis

The limma algorithm [16] was used to select the differentially expressed genes (DEGs) and differentially expressed immune-related genes (DEIGs) between TNBC sample and normal tissues in TCGA-TNBC cohort. The screening threshold was \log_2 fold change >1 and FDR < 0.05. Enrichment analysis for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was performed using the R package ClusterProfiler [17].

Identification of high immunogenicity modules by WGCNA

We used the weighted gene co-expression network analysis (WGCNA) to identify high immunogenicity modules in TNBC [18]. Firstly, similarity matrix was constructed by calculating the Pearson correlation coefficient between gene pairs based on the DEIGs expression data. Then, we used soft threshold of $\beta = 8$ to convert the similarity matrix into an adjacency matrix. A topological overlap matrix (TOM) was created from the adjacency matrix to describe the degree of association between genes. The dynamic pruning tree was designated to identify the modules using 1-TOM as the distance measure. Finally, we identified 7 modules by setting the merging threshold function at 0.25. The modules with significantly higher IRGs expression patterns in TNBC tumor tissues were selected for subsequent analysis. The expression of IRGs in these high immunogenicity modules was also extracted for further analysis.

Consensus molecular clustering of IRGs and gene set variation analysis (GSVA)

According to the expression of IRGs, unsupervised clustering analysis was used to identify diverse IRGs expression patterns. We categorized TNBC patients into different

molecular subtypes for subsequent analysis. The optimal number of clusters was established using the cophenetic coefficient, dispersion, and silhouette. All subtypes identification was conducted by utilizing the R package "ConsensusClusterPlus" with 1000 times repetitions [19]. In order to understand the biological processes involved in different IRGs patterns, we performed GSVA enrichment analysis using the R package "GSVA" [20]. The gene set of "c2.cp.kegg.v7.2.symbols" and "h.all.v7.4.symbols" were obtained from the MsigDB database (<http://www.gsea-msigdb.org/gsea/downloads.jsp>). A list of the top 20 enriched pathways was selected with adjusted $P < 0.05$. To estimate the abundance of 23 immune cell infiltrations in TNBC TIME, the single-sample gene-set enrichment analysis (ssGSEA) algorithm was performed as described before [21]. Gene sets for each immune cell type that we used were reported, previously [22]. The enrichment scores were limited to the range of 0 to 1.

Generation of IRGs gene signature and functional enrichment analysis

To screen the IRG-related DEGs among different IRGs clusters, we utilized the "limma" method with the filtering criteria of adjusted $P < 0.05$. The univariate Cox regression model was applied to assess the prognostic value of IRG-related DEGs. Then the IRGs gene clusters were also achieved with consensus clustering algorithm. To identify functional enrichment of IRG-related DEGs, annotation analysis of GO and KEGG pathway was carried out with R package "ggplot2," "clusterProfiler," "org.Hs.eg.db" and "enrichplot". The top 30 biological terms were used with $P < 0.05$.

Construction and validation of the immune-related gene panel (IRGP) by LASSO regression.

We combined TCGA-TNBC cohort and METABRIC-TNBC cohort to create a final meta-cohort. To establish a prognostic gene panel, we used the least absolute shrinkage and selection operator (LASSO) method to evaluate the associations between IRGs expressions and overall survival (OS) [23]. In this way, each TNBC sample could be calculated with a risk score by the following formula: (coef 1 × expression of gene

1)+(coef 2×expression of gene 2)+...(coef n×expression of gene n). According to the median of risk score, patients with TNBC were divided into high- and low-risk subgroups for subsequent study. Survival analysis in high- and low-risk subgroups was conducted by the R “survival” and “survminer” packages. R package “survivalROC” was used to produce receiver operating characteristic (ROC) curves. Analyses of dimensionality reduction were performed using principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE).

Comprehensive analysis of immune characteristics and ICIs therapy in different IRGP subgroups

To identify immune properties of TNBC samples, mRNA expression data were calculated by the CIBERSORT algorithm to estimate the abundance of 22 types of immune cells [24]. The associations between the abundance of 22 types of immune cells and clinicopathological factors were presented as a landscape map. To clarify the immune function in high- and low-risk subgroups, we identified the IRGP signatures enrichment level as ssGSEA score. For the effects of IRGP signatures on gene mutation, a mutation profile was visualized using the R/Bioconductor 'maftools' package in two IRGP subgroups [25]. In calculating the correlations between IRGP score and PD-1, PD-L1 and CTLA-4 expression, spearman's correlation analysis was performed. To predict chemotherapeutic response, the "pRophetic" R package was implemented to analyze the expression data of each patient [26]. All drug response data were collected from the Genomics of Drug Sensitivity in Cancer (GDPS) database (<https://www.cancerrxgene.org/>) [27]. Immunophenoscore (IPS) data were retrieved from online database (<https://tcia.at/>) for TCGA-TNBC [22]. Analysis of IPS between low and high IRGs riskscore was performed according to the status of CTLA-4 and PD-1 expression.

Statistical analysis

In this study, all statistical analyses were performed using R-4.1.1. Continuous variables were compared between the two groups by using the independent t-test. The chi-square test was used to test categorical data. Statistical comparisons of TIDE score between groups were performed using the Wilcoxon test. A Kaplan-Meier survival analysis based on univariate log-rank test was performed. Statistical significance was defined as a *P* value below 0.05.

Results

Identification of high immunogenicity modules in TNBC by WGCNA

First, we performed differential mRNAs expression analysis between 123 TNBC tumors and 113 normal control in TCGA database. As a result, 6927 differentially expressed mRNAs were screened, of which 3584 genes were upregulated and 3343 genes were downregulated in the TNBC samples compared with normal samples (Fig. S1). Taking the intersection of DEGs with IRGs, 1139 differentially expressed IRGs were obtained, of which 690 genes were upregulated and 449 were downregulated in TNBC samples (Fig. 1a). Functional annotation and pathway analysis revealed that these genes were involved in the process of T cell activation, negative regulation of immune system and chemokine signaling pathway (Fig. 1b-

1c). To obtain the high immunogenicity modules in TNBC, WGCNA analysis was performed based on the expression of DEIGs. The outliers' sample was removed by sample clustering (Fig. S2). The optimal soft-thresholding power of 6 was selected determined by the scale-free network (Fig. 1d). A hierarchical cluster analysis was applied to detect co-expression clusters and 7 co-expression modules were found using dynamic tree-cutting methods (Fig. 1e). Among the 7 co-expression modules, DEIGs in five modules (magenta, green, pink, turquoise, grey) were higher expression in TNBC tumors than normal (Fig. 1f). Since these IRGs from the five modules were more enrichment in TNBC samples, indicating TNBC patients showed a higher expression of these IRGs. Finally, the expression of IRGs in the five high immunogenicity modules was extracted for further analysis.

Construction of distinct IRGs expression patterns in TNBC

We combined TCGA-TNBC and METABRIC-TNBC cohort to generate a final metacohort. Consensus clustering analysis was applied to classify TNBC samples into different IRGs patterns and two IRGs clusters were identified (Fig. 2a and Fig. S3). We term it as IRGcluster A-B and survival analysis showed that TNBC patients with IRGcluster B had a better overall survival (OS) than IRGcluster A (Fig. 2b). ssGSEA scores were calculated to quantify the abundances of 23 immune-infiltrating cells in different IRGs clusters. IRGcluster B subtype presented a higher abundance of activated CD8 T cell, activated dendritic cell, natural killer cells and so on (Fig. 2c). Heatmap demonstrated that the expression of IRGs was higher in IRGcluster B subtype compared to IRGcluster A (Fig. 2d). GSEA was conducted to explore the biological process in two IRGs clusters. As shown in Figure 2e, IRGcluster-B was enriched in KRAS signaling, TNFA signaling and interferon gamma response. Furthermore, immune-related pathways including natural killer cell mediated cytotoxicity, T cell receptor signaling, and antigen processing and presentation were highly enriched in IRGcluster-B subgroup (Fig. 2f). Together these results demonstrated that ICGcluster B was as a high immunogenicity TNBC subtype characterized by abundant immune cells infiltration and activated immune-related pathways.

Generation of IRGs gene signatures and construction of the immune-related gene panel (IRGP)

We further screened the DEGs between two IRGclusters. In total, 530 DEGs were recognized, and functional enrichment analysis demonstrated that these genes were involved in leukocyte mediate immunity, MHC protein complex and various immune processes (Fig. S4). Then, we performed univariate cox regression analysis and 266 prognostic IRGs were associated with OS in TNBC (Table S1). Unsupervised consensus clustering further classified TNBC into two IRG gene clusters (Fig. 3a). Kaplan-Meier curve illustrated that IRG-genecluster B was correlated with better OS compared with IRG-genecluster A (Fig. 3b). To construct an Immune-Related Gene Panel (IRGP) for predicting the OS of TNBC patients, we applied LASSO regression based on the expression of 266 prognostic IRGs and a three-gene signature was identified (Fig. 3c and Fig. 3d). The riskscore was calculated by the formula $IRGP = \text{expression level of TAPBPL} * (-0.41) + \text{expression level of FBP1} * (-0.22) + \text{expression level of GPRC5C} * (0.26)$. Interestingly, a lower riskscore was observed in IRGcluster B and IRG-genecluster B (Fig. 3e and

Fig. 3f). Alluvial diagram illustrated that IRGcluster B had a high overlap with the IRG-genecluster B subtype, which showed a lower IRGriskscore with a favorable survival prognosis (Fig. 3g).

Validation of the capacity of IRGP

Based on the median IRGriskscore, TNBC cohort was divided into low-risk and high-risk subgroup. TNBC patients were randomly classified as two datasets: the training

cohort (N = 211) and testing cohort (N = 210). Kaplan–Meier survival curves showed that TNBC patients with lower IRGriskscore had a longer survival time than those with higher IRGriskscore in both the training and testing cohort (Fig. 4a, b). ROC curves with AUC at 1 year, 3 year and 5 year also demonstrated the moderate accuracy of our IRGP (Fig. 4c, d). Distribution of risk scores between high risk and low risk groups was shown in Fig. 4e and Fig. 4f. The scatter plots demonstrated that an increase of deaths was accompanied by increasing risk score in both the training cohort and testing cohort (Fig. 4g, h). As shown in Fig. 4i, j, the expression of TAPBPL and FBP1 was lower in high-risk group, while a higher expression of GPRC5C was observed in high-risk group.

Prognostic value and genomic features in different IRGP subgroups

PCA and t-SNE analysis showed that IRGP subgroups presented obvious segregation in training cohort (Fig. 5a, b). A similar result was also observed in testing cohort (Fig. 5c, d). To investigate whether the effect of IRGP on OS was an independent factor for TNBC patients, univariate and multivariate Cox analyses for variables including age, tumor stage, tumor size, node status and IRGriskscore were performed. Univariable Cox regression analysis revealed that IRGriskscore was a high-risk factor for TNBC [HR =2.188, 95% CI (1.360, 3.522), $p < 0.001$, Fig. 5e]. Multivariate Cox analysis indicated that IRGP could serve as an independent prognostic factor [HR =2.216, 95% CI (1.384, 3.548), $p < 0.001$, Fig. 5f]. Then, we explored the genomic features in different IRGP subgroups of TCGA-TNBC cohort. The waterfall plot revealed that the low IRGriskscore group presented a decreased mutation burden compared with the high IRGriskscore group (Fig. 5g, h). Among the top 20 most significant mutated gene, TP53 (84% vs. 75%) had higher mutation rates in the high IRGriskscore group.

Immune characteristics of different IRGP subgroups

Immune checkpoints gene expression was often considered as indicators for ICIs treatment. Firstly, the expression of PD-1, PD-L1 and CTLA-4 was significantly higher in low IRGriskscore group compared to high IRGriskscore group (Fig. 6a-6c). IRGriskscore also had a significant negative correlation with the expression of PD-1, PD-L1 and CTLA-4 (Fig. 6d-6f). Immune cells landscape suggested that the proportion of CD8 T cells, B naïve cell, plasma cells, M1 macrophages and T helper cells were more abundant in the low IRGriskscore subgroup, while M2 macrophages were more abundant in the high IRGriskscore subgroup (Fig. 6g). Correlation analysis further suggested that IRGriskscore presented a positive correlation with M0 macrophages, M2 macrophages and T regulatory cells (Fig. 6h-6j). However, a negative correlation was found between IRGriskscore and activated NK cells, CD4 memory T cells, CD8

cells and M1 macrophages (Fig. 6k-6n). All in all, TNBC patients with low IRGriskscore had a higher abundance with tumor-infiltrating immune cells.

Relationship of IRGP subgroups with immune subtypes and IPS

As for the correlation between IRGP and clinicopathological parameters, we found from Fig. 7a that our IRGriskscore was significantly correlated with patients tumor size. Thorsson et al. reported the immune landscape of pan-cancer of TCGA database and defined six immune subtypes [28]. As shown in Fig. 7b, TNBC patients mainly belonged to the subtype of C1 (Wound Healing) and C2 ((IFN- γ Dominant). More C2 subtypes were enriched in low IRGriskscore group ($p = 0.001$, chi-square test). Our IRGP signature contains three immune genes: FBP1, GPRC5C and TAPBPL. As in Fig. 7c, FBP1 had a strong correlation with infiltration of resting dendritic cell, resting mast cells and T helper cells. The expression of GPRC5C was strongly associated with the infiltration of M2 macrophages, resting mast cells and activated CD4 memory T cells. Furthermore, a strong positive correlation was seen between the expression of TAPBPL and M1 macrophages, activated NK cells, activated CD4 memory T cells and CD8 T cells. We then used IPS to assess the potential clinical efficacy of ICIs in different IRGP groups. The higher IPS prediction score indicated a better response to ICIs therapy. There was no significant difference for PD-1 and CTLA-4 negative expression patients (Fig. 7d). However, low IRGriskscore might be a robust indicator for ICIs therapy in TCGA-TNBC cohort for patients with PD-1 or CTLA-4 positive expression (Fig. 7e-7g). Although TNBC patients with high IRGriskscore might not be the most appropriate population for ICIs therapy, our drug sensitivity analysis demonstrated that high IRGriskscore subgroup were more sensitivity to PARP inhibitors (ABT.888 and AZD.2281), doxorubicin, gemcitabine and methotrexate (Fig. S4).

Discussion

Although women with TNBC make up 12–17% of all breast cancer cases, it is clinically recognized as the the worst outcomes among all BC subtypes [29]. With rapid growth in multi-omics approach, many newly discovered types of TNBC have been identified [30, 31]. Genomically, TNBC is heterogeneous and characterized by distinct molecular subtypes that lead to different biological behavior and treatment responses [32]. Recent advances in antitumor immunotherapy by targeting the PD-1/PD-L1 checkpoint have led to significant clinical improvements in TNBC and HER2 positive breast cancers [33]. The clinical efficacy of cancer immunotherapy is associated with individual tumor immune microenvironment and “hot” tumor tend to have a better ICIs response than “cold” one [34]. Thus, identification of TNBC subtypes with highly immunogenic will facilitate a rational design of anti-tumor immunotherapy. Recently, the involvement of gene signatures has been described in several studies based on multi-omics data. Yan et al. developed a risk autophagy-related prediction model that can predict the survival status of TNBC patients[35]. Yang et al. constructed a risk scoring system to predict the immune activity and potential therapeutic response of TNBC [36]. Our previous work revealed an eight immune-related panel in the immunomodulatory subtype of TNBC [37]. Given the complexity of tumor immune microenvironment, we first explored the TNBC subtypes according to the immune genes’ expression patterns. Beyond this, a risk scoring panel was established and we further investigated the association of IRGs riskscore with

immune infiltration and response to anti-tumor targeted immunotherapy. This study aimed to explore the role of immune-related genes network in a systems biology manner and investigated the immunotherapeutic value of the immune-related signature in TNBC.

Various molecular subtypes of TNBC have been identified, including basal-like 1 (BL1), basal-like 2 (BL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem-like (MSL) and luminal androgen receptor (LAR) [38]. The IM subtype manifests high levels of immune signal transduction pathways and immune-related genes [39]. The expression levels of PD-L1, PD-1 and CTLA4 were reported to be higher in IM subtype than other subtypes [31]. The FUTURE trial showed that anti PD-1 strategy could be a promising treatment for refractory metastatic TNBCs with IM subtype [40]. To find the high immunogenic subtype, we firstly screened the differentially expressed immune-related genes and functional annotation analysis further confirmed these DEGs indeed played an essential role in regulation of lymphocyte activation, T cell activation and various immune-related process. Then we identified the high expression ICGs modules by WGCNA and further categorized TNBC into two subtypes based on the high immunogenicity TNBC modules. Further analysis revealed that IRGcluster-B had a superior survival compared to IRGcluster-A. Abundant immune cell infiltration was observed in IRGcluster-B patients with elevated immune genes expression levels. GSVA analysis also confirmed that IRGcluster-B subgroup was markedly enriched in immune-related biological processes including interferon gamma signaling, T cell receptor signaling pathway and so on. Thus, TNBC patients with IRGcluster-B subtype might belong to immune “hot” tumor and more likely to benefit from immunotherapy.

To investigate the major difference at molecular level between these two subtypes, we further constructed IRG-related genomic clusters and two gene clusters were identified. Interestingly, IRG-genecluster B have a better survival outcome than IRG-genecluster A, which was similar to IRGs cluster patterns. To evaluate different IRGs expression for individual patients, a quantification scoring model termed IRGriskscore

was adopted by LASSO algorithm. Previous studies have also attempted to construct an immune-related gene prognostic index to distinguish the molecular and immune characteristics in various solid tumors [41–43]. For the associations of IRGriskscore with IRGcluster and IRG-genecluster, we found the IRGcluster-B and IRG-genecluster B subtype exhibited a lower IRGriskscore. To verify the accuracy and efficacy of our proposed model, we then assessed the survival outcomes, diagnostic value, risk curve, univariate and multivariate analyses in training and validation dataset. As a result, there was little variation in the training and validation dataset, suggesting our model had certain universality, efficacy, and accuracy. Overall, the TNBC patients with low IRGriskscore was associated with improved prognosis.

Here, TAPBPL, FBP1 and GPRC5C was applied to build our IRGs signature. These genes were identified as prognostic immune signatures in TNBC for the first time. TAP binding protein like (TAPBPL), as a novel T cell co-inhibitory factor, was reported to localize on chromosome position 12p13.3 near an MHC paralogous locus [44]. Study showed that TAPBPL was expressed on the surface of APCs and played a critical role in negatively controlling T-cell functions [45]. Targeting TAPBPL-mediated inhibitory pathway might be effective strategy for the patients who were resistant to PD-1/PD-L1 or CTLA-4 antitumor

therapy [45]. As a rate-limiting enzyme in gluconeogenesis, fructose-1,6-bisphosphatase (FBP1) controls the rate of the reaction and works as an important tumor suppressor in human malignancies [46]. Previous study demonstrated that overexpression of FBP1 conferred sensitivity to cisplatin via modulating STAT3 in ovarian cancer [47]. Another study reported that FBP1 inhibited tumorigenesis of cholangiocarcinoma partly via Wnt/ β -catenin pathway [48]. The study of Cong et al. revealed that aberrant expression of FBP1 in NK cells resulted in their dysfunction via inhibiting glycolysis and impairing viability [49]. To date, the role of FBP1 in anti-tumor immune response is not well-understood. GPRC5C belongs to the type 3 G protein-coupled receptor family [50]. It was reported that GPRC5C was found in inrenal proximal tubules and participated in regulating renal acid-base homeostasis [51]. In pancreatic β -cells, down-regulation of GPRC5C resulted in a decreased of glucose-stimulated insulin release and cAMP content [52]. However, the function of GPRC5C in cancer remain unclear.

Recent evidence demonstrated that infiltrating lymphocytes (TILs) was observed in HER2 + BC and TNBC patients [53]. Therefore, BC has ceased to be considered as an immunological quiescent tumor type. Previous research indicated that abundance of TILs in the tumor microenvironment was related to better response to ICI therapies [54, 55]. In the present study, patients with low IRGriskscore displayed a higher level of TILs compared to high IRGriskscore subgroup. IRGriskscore was negatively associated with infiltrating of activated NK, CD4 and CD8 cells, suggesting that our IRGriskscore might be a good predictive marker for ICI therapies. Currently, atezolizumab and pembrolizumab were approved in combination with chemotherapy among patients with unresectable locally or metastatic TNBC expressing PD-L1 [56, 57]. Our results showed that high level of PD1, PD-L1 and CTLA4 was observed in TNBC patients with low IRGriskscore, which strongly implicated that TNBC patients with low IRGriskscore might be suitable for anti-tumor immunotherapy. For the effect of IRGP signature on gene mutation rate in TNBC, our findings demonstrated that patients with low IRGriskscore had unusually low mutation rates compared to high IRGriskscore group, specifically for TP53. Previous studies reported that TP53 mutation status were associated with depressed immune signatures and the efficiency of ICIs therapy in solid tumors [58–60]

The study of Thorsson et al. identified six immune subtypes based on the differences in lymphocyte signatures across TCGA pan-cancer [61]. The subtype of C1 (wound healing) had increased expression of angiogenic genes and C2 (IFN- γ dominant) subtype showed a strong CD8 signal and the highest M1/M2 macrophage polarization [61]. In our analysis, we found TCGA-TNBC cohort was primarily categorized as C1 and C2 subtype. Most of low IRGriskscore TNBC patients belong to C2 subtype, suggestive of high lymphocyte expression signature. Based on previous studies, IPS have been identified as useful biomarkers for predicting patient response to immunotherapy [22]. IPS was able to quantify the determinants of tumor immunogenicity and had predictive value in cancer patients treated with the CTLA-4 and PD-1 blockers [22]. In this study, we also evaluated the predictive value of the TNBC IRGriskscore using IPS. Low IRGriskscore had the potential to be a robust predictor for PD-1 or CTLA-4 therapy and its predictive value depended on the PD-1 or CTLA-4 expression. Meanwhile, drug sensitivity analysis revealed that TNBC patients with high IRGriskscore was associated with high sensitivity of PARP

inhibitors, doxorubicin, gemcitabine and methotrexate, which will probably guide clinical practice for chemotherapeutic agents' decision-making.

To our knowledge, we firstly identified immune-related subtypes of TNBC using WGCNA combined with consensus clustering. There are several notable limitations to our study. First, our immune signature was constructed based on TCGA and METABRIC database. Thus, prospective cohort of TNBC patients receiving ICIs therapy are needed to validate our findings. Second, the effect of our gene signature on immune infiltration should be experimentally verification in the future. Thirdly, comparison of predictive value between different models also should be conducted.

Conclusions

Schematic summary of the proposed model was presented as Fig. 8. In conclusion, we identified the DEIGs in TNBC and found five high immunogenicity modules by WGCNA. We developed two distinct IRGclusters, two IRG-related genomic clusters and IRGriskscore to assess the characteristics of immune cells infiltration. The three-gene panel was strongly correlated with patient's prognosis, the abundances of immune cell infiltration, immunophenotyping of TNBC, chemotherapy and immunotherapy sensitivity. Our study offered novel insights into molecular subtype of TNBC and provided potential indicators for guiding ICIs treatment.

Abbreviations

APCs: Antigen presenting cells; BC: Breast cancer; BL1: basal-like 1; BL2: basal-like 2; CTLA-4: cytotoxic T-lymphocyte-associated antigen-4; DEGs: Differentially expressed genes; FBP1: Fructose-1,6-bisphosphatase; FDA: Food and Drug Administration; FDR: false discovery rate; GDPS: Genomics of Drug Sensitivity in Cancer; GO: Gene Ontology; GSVA: Gene Set Variation Analysis; GSEA: Gene set enrichment Analysis; ICIs: Immune checkpoint inhibitors; IM: Immunomodulatory ; IPS: Immunophenoscore; IRGs: Immune-related genes; IRGP: Immune-related gene panel; KEGG: Kyoto Encyclopedia of Genes and Genomes; LAR: luminal androgen receptor; LASSO: least absolute shrinkage and selection operator; M: mesenchymal; METABRIC: Molecular Taxonomy of Breast Cancer International Consortium; MHC: major histocompatibility complex; MSL: mesenchymal stem-like; OS: overall survival; PARP: poly (ADP-ribose) polymerase; PCA: principal component analysis; PD-1: programmed cell death receptor 1; PD-L1: programmed cell death 1 ligand 1; ROC: Receiver operating characteristic; ssGSEA: single-sample gene-set enrichment analysis; STAT3: signal transducer and activator of transcription 3; TAPBPL: TAP binding protein like; TCGA: The Cancer Genome Atlas; TILs: Tumor infiltrating lymphocytes; TNBC: Triple-negative breast cancer TOM: Topological overlap matrix; t-SNE: t-distributed stochastic neighbor embedding; WGCNA: Weighted gene co-expression network analysis

Declarations

Acknowledgments

We are thankful for all research participants and clinicians involved in this research project. We acknowledged TCGA database and cBioPortal database for providing their platforms and contributors for uploading their meaningful datasets.

Authors' contributions

X H and J Z were responsible for the research design. J Z and S P were responsible for data collection and analysis. J Z, S P, H J and Q S were responsible for manuscript preparation and revision. All authors read and approved the final manuscript.

Funding

This work was supported Scientific Research Start-up Funds of The First Affiliated Hospital of USTC (RC2021122).

Availability of data and materials

Publicly available datasets or databases were analyzed in the present study. This data can be found here: TCGA database (<https://portal.gdc.cancer.gov>) and cBioPortal database (<http://www.cbioportal.org/>).

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

Not applicable.

Competing interests

The authors have declared that no competing interest exists.

Author details

¹ Department of Medical Oncology, The First Affiliated Hospital of USTC, Division of Life Science and Medicine, University of Science and Technology of China, Hefei, China.

² School of Medical Oncology, Wan Nan Medical College, China.

³ School of Medical Oncology, Anhui Medical University, Hefei, China

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians*. 2021;71(3):209–49.
2. Wong KK. DNMT1: A key drug target in triple-negative breast cancer. *Seminars in cancer biology*. 2021;72:198–213.
3. Yin L, Duan JJ, Bian XW, Yu SC. Triple-negative breast cancer molecular subtyping and treatment progress. *Breast cancer research: BCR*. 2020;22(1):61.
4. Zhu Y, Zhu X, Tang C, Guan X, Zhang W. Progress and challenges of immunotherapy in triple-negative breast cancer. *Biochimica et biophysica acta Reviews on cancer*. 2021;1876(2):188593.
5. Jovanović B, Mayer IA, Mayer EL, Abramson VG, Bardia A, Sanders ME, et al. A Randomized Phase II Neoadjuvant Study of Cisplatin, Paclitaxel With or Without Everolimus in Patients with Stage II/III Triple-Negative Breast Cancer (TNBC): Responses and Long-term Outcome Correlated with Increased Frequency of DNA Damage Response Gene Mutations, TNBC Subtype, AR Status, and Ki67. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2017;23(15):4035–45.
6. Emens LA. Breast Cancer Immunotherapy: Facts and Hopes. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2018;24(3):511–20.
7. Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. *Nature*. 2011;480(7378):480–9.
8. Loi S, Michiels S, Adams S, Loibl S, Budczies J, Denkert C, et al. The journey of tumor-infiltrating lymphocytes as a biomarker in breast cancer: clinical utility in an era of checkpoint inhibition. *Annals of oncology: official journal of the European Society for Medical Oncology*. 2021;32(10):1236–44.
9. Emens LA, Molinero L, Loi S, Rugo HS, Schneeweiss A, Diéras V, et al. Atezolizumab and nab-Paclitaxel in Advanced Triple-Negative Breast Cancer: Biomarker Evaluation of the IMpassion130 Study. *Journal of the National Cancer Institute*. 2021;113(8):1005–16.
10. Cetin B, Gumusay O. Pembrolizumab for Early Triple-Negative Breast Cancer. *The New England journal of medicine*. 2020;382(26):e108.
11. Tarantino P, Antonarelli G, Ascione L, Curigliano G. Investigational immunomodulatory drugs for enhancement of triple negative breast cancer (TNBC) immunotherapy: early phase development. *Expert opinion on investigational drugs*. 2021:1–15.
12. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Annals of oncology: official journal of the European Society for Medical Oncology*. 2019;30(1):44–56.
13. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486(7403):346–52.
14. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Scientific data*.

2018;5:180015.

15. Breuer K, Froushani AK, Laird MR, Chen C, Sribnaia A, Lo R, et al. InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic acids research*. 2013;41(Database issue):D1228-33.
16. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology*. 2004;3:Article3.
17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*. 2012;16(5):284–7.
18. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics*. 2008;9:559.
19. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics (Oxford, England)*. 2010;26(12):1572–3.
20. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC bioinformatics*. 2013;14:7.
21. Zhang J, Han X, Lin L, Chen J, Wang F, Ding Q, et al. Unraveling the Expression Patterns of Immune Checkpoints Identifies New Subtypes and Emerging Therapeutic Indicators in Lung Adenocarcinoma. *Oxidative medicine and cellular longevity*. 2022;2022:3583985.
22. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell reports*. 2017;18(1):248–62.
23. Zhang J, Shan B, Lin L, Dong J, Sun Q, Zhou Q, et al. Dissecting the Role of N6-Methyladenosine-Related Long Non-coding RNAs Signature in Prognosis and Immune Microenvironment of Breast Cancer. *Frontiers in cell and developmental biology*. 2021;9:711859.
24. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature methods*. 2015;12(5):453–7.
25. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome research*. 2018;28(11):1747–56.
26. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS one*. 2014;9(9):e107468.
27. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids research*. 2013;41(Database issue):D955-61.
28. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. *Immunity*. 2018;48(4):812 – 30.e14.
29. Derakhshan F, Reis-Filho JS. Pathogenesis of Triple-Negative Breast Cancer. *Annual review of pathology*. 2022;17:181–204.

30. Gong Y, Ji P, Yang YS, Xie S, Yu TJ, Xiao Y, et al. Metabolic-Pathway-Based Subtyping of Triple-Negative Breast Cancer Reveals Potential Therapeutic Targets. *Cell metabolism*. 2021;33(1):51–64.e9.
31. Jiang YZ, Ma D, Suo C, Shi J, Xue M, Hu X, et al. Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. *Cancer cell*. 2019;35(3):428 – 40.e5.
32. Bianchini G, De Angelis C, Licata L, Gianni L. Treatment landscape of triple-negative breast cancer - expanded options, evolving needs. *Nature reviews Clinical oncology*. 2022;19(2):91–113.
33. Zhang J, Pan S, Jian C, Hao L, Dong J, Sun Q, et al. Immunostimulatory Properties of Chemotherapy in Breast Cancer: From Immunogenic Modulation Mechanisms to Clinical Practice. *Frontiers in immunology*. 2021;12:819405.
34. Gaikwad S, Agrawal MY, Kaushik I, Ramachandran S, Srivastava SK. Immune checkpoint proteins: Signaling mechanisms and molecular interactions in cancer immunotherapy. *Seminars in cancer biology*. 2022.
35. Yan C, Liu Q, Jia R. Construction and Validation of a Prognostic Risk Model for Triple-Negative Breast Cancer Based on Autophagy-Related Genes. *Frontiers in oncology*. 2022;12:829045.
36. Yang A, Wu M, Ni M, Zhang L, Li M, Wei P, et al. A risk scoring system based on tumor microenvironment cells to predict prognosis and immune activity in triple-negative breast cancer. *Breast cancer (Tokyo, Japan)*. 2022.
37. Zhang J, Wang L, Xu X, Li X, Guan W, Meng T, et al. Transcriptome-Based Network Analysis Unveils Eight Immune-Related Genes as Molecular Signatures in the Immunomodulatory Subtype of Triple-Negative Breast Cancer. *Frontiers in oncology*. 2020;10:1787.
38. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of clinical investigation*. 2011;121(7):2750–67.
39. Lehmann BD, Pietenpol JA. Clinical implications of molecular heterogeneity in triple negative breast cancer. *Breast (Edinburgh, Scotland)*. 2015;24 Suppl 2(0 2):S36-40.
40. Jiang YZ, Liu Y, Xiao Y, Hu X, Jiang L, Zuo WJ, et al. Molecular subtyping and genomic profiling expand precision medicine in refractory metastatic triple-negative breast cancer: the FUTURE trial. *Cell research*. 2021;31(2):178–86.
41. Cao K, Liu M, Ma K, Jiang X, Ma J, Zhu J. Prediction of prognosis and immunotherapy response with a robust immune-related lncRNA pair signature in lung adenocarcinoma. *Cancer immunology, immunotherapy: CII*. 2021.
42. Chen Y, Li ZY, Zhou GQ, Sun Y. An Immune-Related Gene Prognostic Index for Head and Neck Squamous Cell Carcinoma. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2021;27(1):330–41.
43. Feng D, Zhang F, Li D, Shi X, Xiong Q, Wei Q, et al. Developing an immune-related gene prognostic index associated with progression and providing new insights into the tumor immune microenvironment of prostate cancer. *Immunology*. 2022.

44. Neerincx A, Boyle LH. Properties of the tapasin homologue TAPBPR. *Current opinion in immunology*. 2017;46:97–102.
45. Lin Y, Cui C, Su M, Silbart LK, Liu H, Zhao J, et al. Identification of TAPBPL as a novel negative regulator of T-cell function. *EMBO molecular medicine*. 2021;13(5):e13404.
46. Yang X, Zhao L, Pei J, Wang Z, Zhang J, Wang B. CELF6 modulates triple-negative breast cancer progression by regulating the stability of FBP1 mRNA. *Breast cancer research and treatment*. 2020;183(1):71–82.
47. Li H, Qi Z, Niu Y, Yang Y, Li M, Pang Y, et al. FBP1 regulates proliferation, metastasis, and chemoresistance by participating in C-MYC/STAT3 signaling axis in ovarian cancer. *Oncogene*. 2021;40(40):5938–49.
48. Zhao W, Yang S, Chen J, Zhao J, Dong J. Forced overexpression of FBP1 inhibits proliferation and metastasis in cholangiocarcinoma cells via Wnt/ β -catenin pathway. *Life sciences*. 2018;210:224–34.
49. Cong J, Wang X, Zheng X, Wang D, Fu B, Sun R, et al. Dysfunction of Natural Killer Cells by FBP1-Induced Inhibition of Glycolysis during Lung Cancer Progression. *Cell metabolism*. 2018;28(2):243 – 55.e5.
50. Robbins MJ, Michalovich D, Hill J, Calver AR, Medhurst AD, Gloger I, et al. Molecular cloning and characterization of two novel retinoic acid-inducible orphan G-protein-coupled receptors (GPRC5B and GPRC5C). *Genomics*. 2000;67(1):8–18.
51. Rajkumar P, Cha B, Yin J, Arend LJ, Păunescu TG, Hirabayashi Y, et al. Identifying the localization and exploring a functional role for Gprc5c in the kidney. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*. 2018;32(4):2046–59.
52. Amisten S, Mohammad Al-Amily I, Soni A, Hawkes R, Atanes P, Persaud SJ, et al. Anti-diabetic action of all-trans retinoic acid and the orphan G protein coupled receptor GPRC5C in pancreatic β -cells. *Endocrine journal*. 2017;64(3):325–38.
53. Kumar A, Swain CA, Shevde LA. Informing the new developments and future of cancer immunotherapy: Future of cancer immunotherapy. *Cancer metastasis reviews*. 2021;40(2):549–62.
54. Kirtane K, Elmariah H, Chung CH, Abate-Daga D. Adoptive cellular therapy in solid tumor malignancies: review of the literature and challenges ahead. *Journal for immunotherapy of cancer*. 2021;9(7).
55. Park J, Kwon M, Kim KH, Kim TS, Hong SH, Kim CG, et al. Immune Checkpoint Inhibitor-induced Reinvigoration of Tumor-infiltrating CD8(+) T Cells is Determined by Their Differentiation Status in Glioblastoma. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2019;25(8):2549–59.
56. Cortes J, Cescon DW, Rugo HS, Nowecki Z, Im SA, Yusof MM, et al. Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for previously untreated locally recurrent inoperable or metastatic triple-negative breast cancer (KEYNOTE-355): a randomised, placebo-controlled, double-blind, phase 3 clinical trial. *Lancet (London, England)*. 2020;396(10265):1817–28.

57. Emens LA, Adams S, Barrios CH, Diéras V, Iwata H, Loi S, et al. First-line atezolizumab plus nab-paclitaxel for unresectable, locally advanced, or metastatic triple-negative breast cancer: IMpassion130 final overall survival analysis. *Annals of oncology: official journal of the European Society for Medical Oncology*. 2021;32(8):983–93.
58. Pan YH, Zhang JX, Chen X, Liu F, Cao JZ, Chen Y, et al. Predictive Value of the TP53/PIK3CA/ATM Mutation Classifier for Patients With Bladder Cancer Responding to Immune Checkpoint Inhibitor Therapy. *Frontiers in immunology*. 2021;12:643282.
59. Zhang Y, Lin A, Li Y, Ding W, Meng H, Luo P, et al. Age and Mutations as Predictors of the Response to Immunotherapy in Head and Neck Squamous Cell Cancer. *Frontiers in cell and developmental biology*. 2020;8:608969.
60. Lyu H, Li M, Jiang Z, Liu Z, Wang X. Correlate the TP53 Mutation and the HRAS Mutation with Immune Signatures in Head and Neck Squamous Cell Cancer. *Computational and structural biotechnology journal*. 2019;17:1020–30.
61. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. *Immunity*. 2019;51(2):411–2.

Figures

Figure 1

Identification of differentially expressed immune-related genes (DEIGs) and high immunogenicity modules in TNBC. a Heatmap of differentially expressed immune-related genes between TNBC tumor samples (red) and normal samples (blue) using the “limma” method. b Circos graph for GO enrichment of DEIGs. c Circos graph for KEGG pathways analysis of DEIGs. d Network topology for different soft-thresholding powers. A soft-thresholding power of 8 was selected to achieve maximal model fit. e Clustering dendrogram of DEIGs. Each short vertical line corresponds to a gene. The branches are modules of highly interconnected groups of gene expression. Seven modules were identified, and the lower panel shows colors designated for each module. f The module–trait relationships between the identified modules and clinical status (normal and tumor). Rows are module eigengene (ME) regards to each module, and the columns indicate traits. Red represents high adjacency and blue represents low adjacency. The upper number in each cell indicates the correlation coefficient of each module in the trait, and the lower number is the corresponding P-value.

Figure 2

Identification of distinct IRGs subtypes in TNBC. a Consensus clustering matrix for $k = 2$ based on IRGs expression. b Kaplan-Meier curves of OS for two TNBC subtypes. The numbers of patients in IRGcluster-A and IRGcluster-B are 196 and 225, respectively. c The abundance of infiltrating immune cells in two IRGs subtypes. The lines in the boxes were the median value. The top and bottom ends of the boxes indicated interquartile range of values (* $P < 0.05$; *** $P < 0.001$). d Unsupervised clustering of IRGs expression to divide TNBC patients into two IRGs subtypes. The IRG clusters, projects, age, stage, tumor size and node status were used as patient annotations. e-f Heat map showed the GSVA score of hallmark signature and KEGG pathways in different IRGcluster. Red means activated pathways, and blue means inhibited pathways.

Figure 3

Construction of IRGs gene signatures and the immune-related gene panel (IRGP). a The identification of IRG-geneCluster by consensus clustering matrix for $k = 2$. b Survival curves of the IRG-geneCluster were plotted by the Kaplan-Meier plotter. The numbers of patients in IRG-geneCluster A and IRG-geneCluster B subtypes are 213 and 208. c Elucidation for LASSO coefficient profiles of prognostic IRGs. d The least absolute shrinkage was performed and construction of selection operator (LASSO) regression model. e The correlation between IRGs risk score and IRGs subtypes. f The correlation between IRGs risk score and IRG-geneCluster. g Alluvial diagram of IRGs clusters in groups with different IRG-geneCluster, IRGs risk score, and OS.

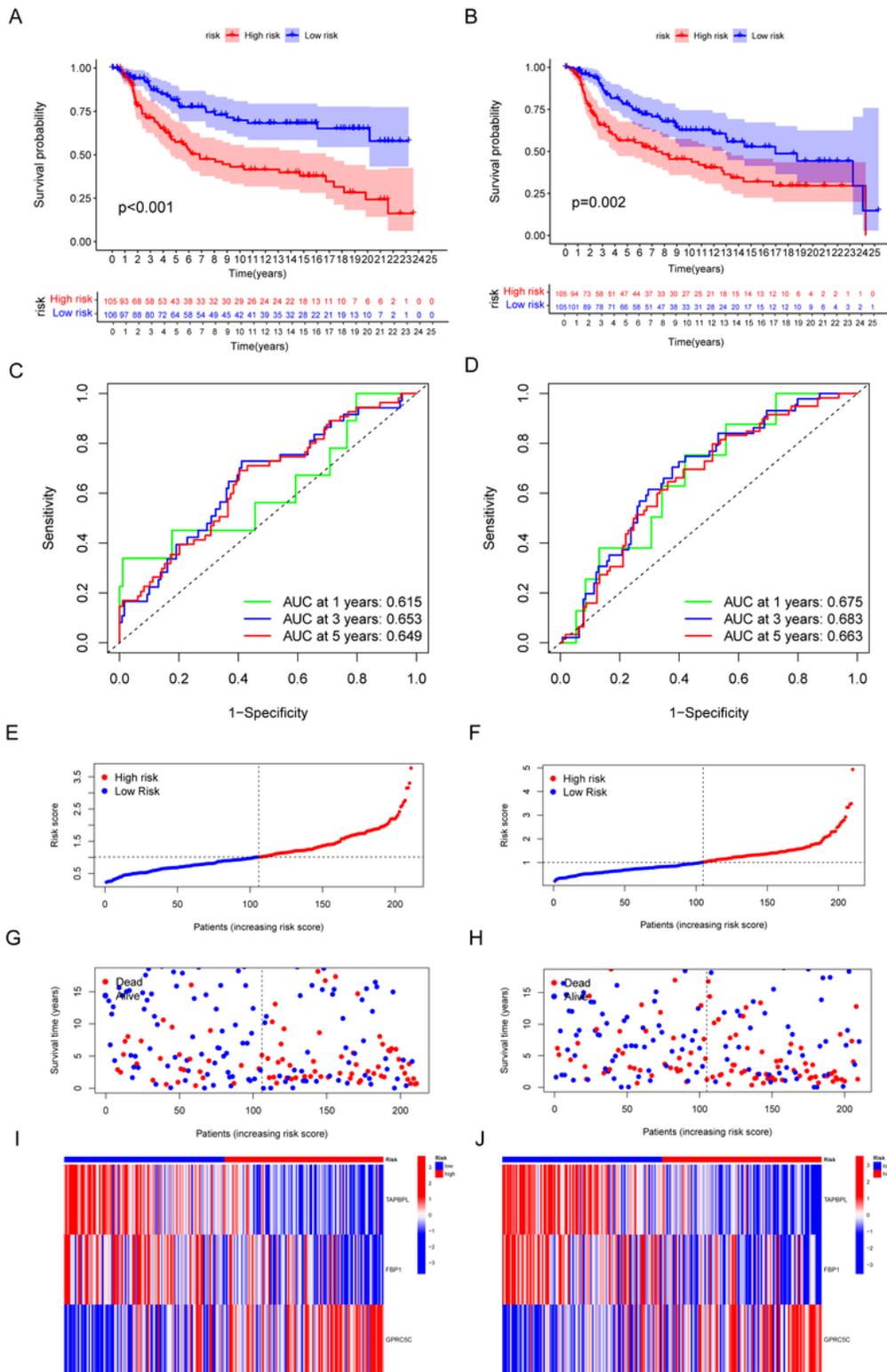


Figure 4

Validation of the efficiency of IRGP. a-b Kaplan–Meier curves for OS of TNBC patients according to the risk stratification in the training cohort and testing cohort. c-d ROC analysis for OS prediction at 1 year, 3 year and 5 year in TNBC patients in the training cohort and testing cohort. e-f Distribution of risk score for the training group and testing group. g-h Distribution of survival time of patients in the training and

testing groups. i-j Heatmap depicting the expression of TAPBPL, FBP1 and GPRC5C between high-risk group and the low-risk group in training and testing groups.

Figure 5

Prognostic value and genomic features in different IRGP subgroups. a-b PCA analysis and t-SNE analysis of TNBC patients in different IRGP subgroups for the training cohort. c-d PCA analysis and t-SNE analysis of TNBC patients based on IRGP model for the testing cohort. e Forest plot for IRGP model score and clinical features in TNBC patients by univariate analyses. f The multivariate Cox forest plot of IRGP model score and clinical characteristics. g-h The waterfall plot of showed the somatic mutation rate in low IRGP riskscore and high IRGP riskscore. Each column indicated an individual sample. The upper barplot was the tumor mutation burden for an individual sample. The right histogram generalized the percentage of each variant type. The number on the right indicated the mutation frequency for each gene.

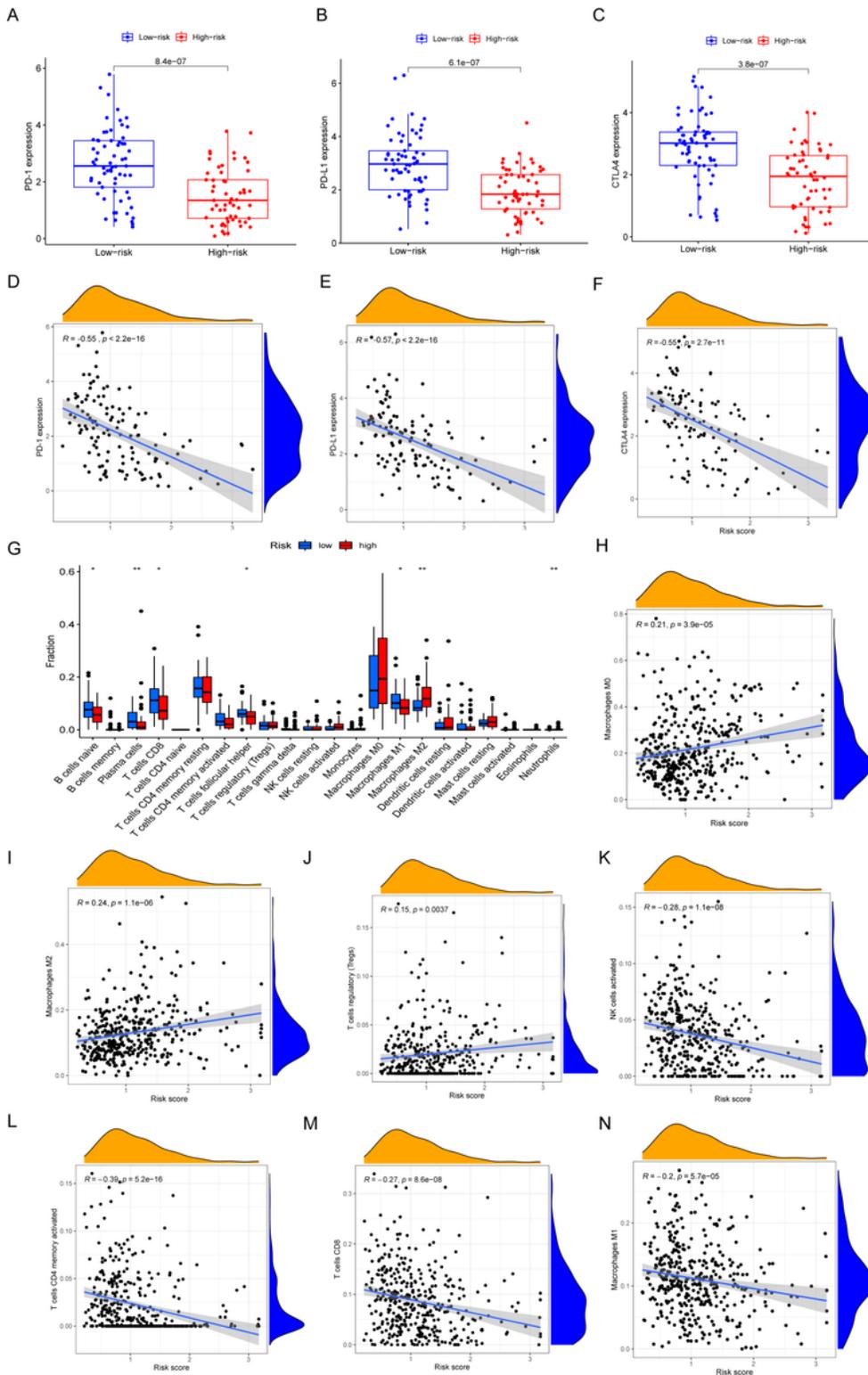


Figure 6

The immune characteristics of different IRGP subgroups. a-c The expression of PD-1, PD-L1, and CTLA-4 in high- and low-risk score TNBC groups. d-f IRGP signature negatively related to the expression of PD-1, PD-L1, and CTLA-4. g The abundance of each infiltrating immune cells in different IRGP subgroups. The lines in the boxes were the median value. The top and bottom ends of the boxes showed interquartile range of values (* $P < 0.05$; ** $P < 0.01$). h-j IRGP signature positively related to infiltrating of M0

macrophages, M2 macrophages and T regulatory cells. k-n IRGP signature negatively correlated with infiltrating of activated NK cells, CD4 memory T cells, CD8 cells and M1 macrophages.

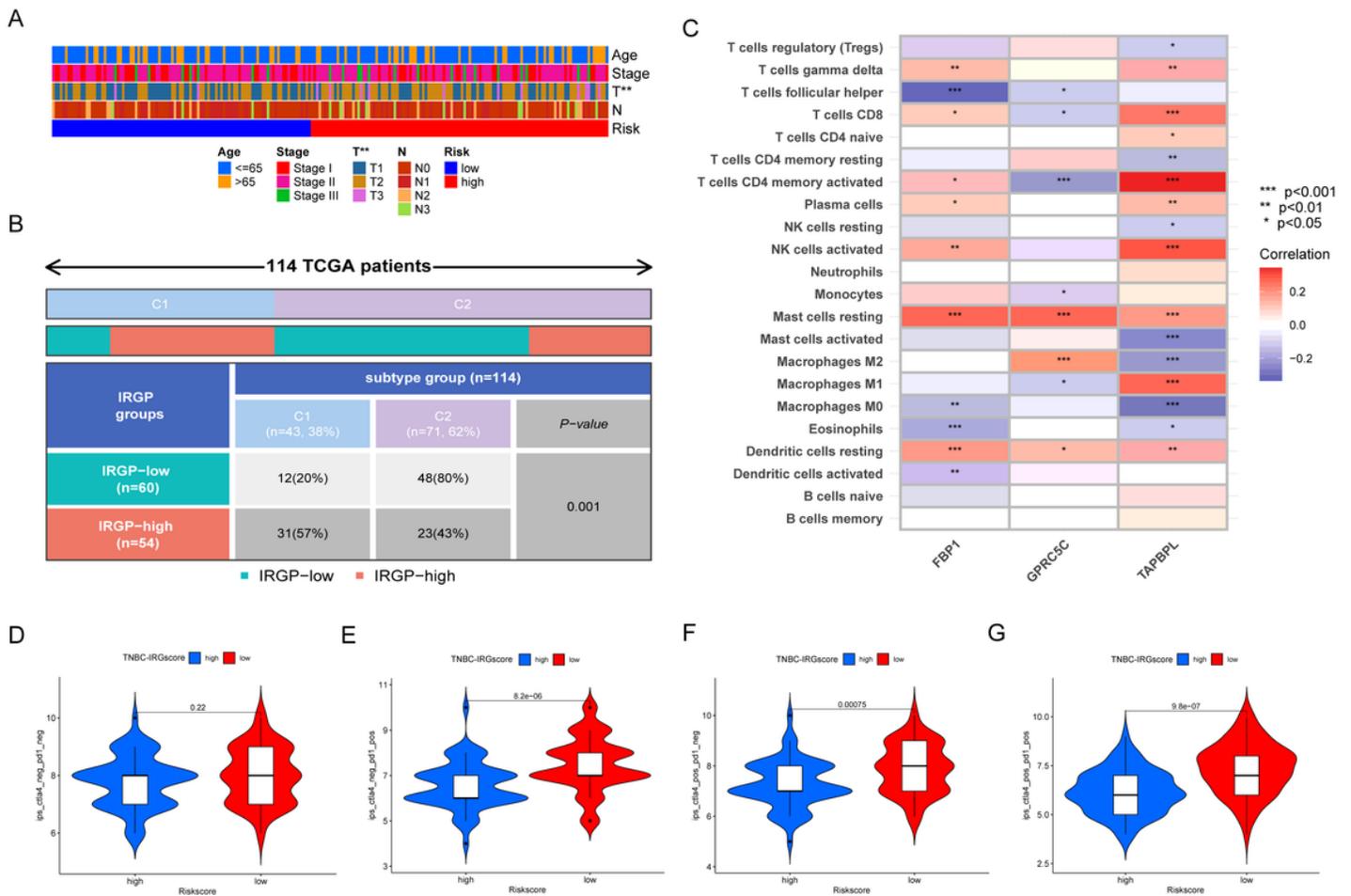


Figure 7

Predictive value of IRGP model in immune subtypes and IPS. a Heatmap showing the relationship between IRGP model and clinicopathological parameter. b Heatmap showing the relationship between IRGP model and immune subtypes. c Heatmap showing the correlation of FBP1, GPRC5C and TAPBPL with the infiltrating of 22 immune cells. d-g The relative distribution of IPS identified by the status of CTLA-4 or PD-1 was compared between high IRGriskscore versus low IRGriskscore in TCGA-TNBC cohort.

Figure 8

The flowchart of the current study.

Supplementary Files

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

- [S1.pdf](#)
- [S2.pdf](#)
- [S3.tif](#)
- [S4.tif](#)
- [S5.tif](#)
- [TABLES1.docx](#)