Am J Cancer Res 2022;12(8):3913-3931
www.ajcr.us /ISSN:2156-6976/ajcr0144796

Original Article

A novel risk model based on immune response predicts clinical outcomes and characterizes immunophenotypes in triple-negative breast cancer

Xunxi Lu[1,2*], Zongchao Gou[3,4*], Luoting Yu[4], Hong Bu[1,2]

[1]Department of Pathology, West China Hospital, Sichuan University, Chengdu 610041, China; [2]Institute of Clinical Pathology, West China Hospital, Sichuan University, Chengdu 610041, China; [3]Department of Breast Surgery, West China Hospital, Sichuan University, Chengdu 610041, China; [4]State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu 610041, China. [*]Equal contributors.

Received June 16, 2022; Accepted August 3, 2022; Epub August 15, 2022; Published August 30, 2022

Abstract: Triple-negative breast cancer (TNBC) is highly heterogeneous in prognosis. The current TNM staging system shows its limitation in accurate risk evaluation. Immune response and immune cell abundances in the tumor immune microenvironment (TIME) are critical for cancer progression, clinical outcome and therapeutic response in TNBC. However, there is a lack of an effective risk model based on the overall transcriptional alterations relevant to different immune responses. In this study, multiple bioinformatics and statistical approaches were used to develop an immune-related risk (IRR) signature based on the differentially expressed genes between the immune-active and immune-inactive samples. The IRR model showed great performance in risk stratification, immune landscape evaluation and immunotherapy response prediction. Compared with the low-IRR group, the high-IRR group exhibited a poorer prognosis, less cytotoxic cell infiltration, higher M2/M1 ratio and upregulated glycolytic activity. Moreover, the high-IRR group showed more resistance to immunotherapy than the low-IRR group. Our study reveals that the IRR model may be a promising tool to help clinicians assess risk and optimize treatment for TNBC patients.

Keywords: Triple-negative breast cancer, prognosis, tumor immune microenvironment, immunotherapy

Introduction

Triple-negative breast cancer (TNBC) is characterized by a lack of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) expression, accounting for 10-15% of all breast cancer subtypes [1]. TNBC is highly heterogeneous and presents higher genetic instability, frequent copy number alternation and complex structural rearrangement [2]. Compared with the other subtypes, TNBC is more aggressive and prone to early relapse and distant metastasis. To date, TNBC mainly depends on adjuvant chemotherapy due to the lack of therapeutic targets; thus, its prognosis is poorer than that of the other subtypes.

It is widely accepted that the American Joint Committee on Cancer (AJCC) tumor-node-metastasis (TNM) classification provides the standard guideline to stratify cancer patients into different risks. However, significant differences in the prognosis of TNBC patients within the same pathological TNM stage are still observed, suggesting the limitations of the TNM staging system [3]. Immune activation and the tumor immune microenvironment (TIME) in TNBC play a considerable role in TNBC development. Many published works have confirmed the prognostic value of the immune response and immune cell infiltration for TNBC [4, 5]. For example, abundant tumor-infiltrating lymphocytes (TILs) are associated with a good prognosis in TNBC. Recently, high-throughput sequencing technologies have been developing rapidly, followed by the emergence of many machine learning methods relevant to immune cell quantification [6]. Some studies attempt to construct prognostic models using prognosis-related immune cell abundances quantified by immunohistochemistry, immunofluorescen-

ce or machine algorithms [7-9]. Some have developed prognostic signatures based on predefined gene sets of cancer-related hallmarks [10-12]. Nevertheless, none of these models was developed based on the overall molecular alterations relevant to immune responses that are associated with distinct prognoses.

With the increase in immunotherapy trials for TNBC in recent years, it has been proven that TNBC patients can benefit from immune checkpoint inhibitor (ICI) treatment [13-15]. However, the exploration of effective biomarkers for predicting the efficacy of immunotherapy remains a challenge. Due to a relatively low tumor mutation burden (TMB) and low incidence of mismatch repair, PD-L1 expression is the only recognized and extensively used biomarker for immunotherapy in TNBC [16]. PD-L1 has some limitations as well. For instance, some PD-L1-negative patients also show a response to ICI therapy, but some PD-L1-positive patients do not [17]. Therefore, it is necessary to explore more novel predictors to identify TNBC patients who have the potential to benefit from immunotherapy.

In this study, we divided TNBC patients into immune-active and immune-inactive groups based on immune response-associated cells and pathways, and then we constructed a gene signature based on the different genes in the two groups. The predictive power of this risk model was validated in external GEO datasets. We also investigated the differences in TIME components, transcriptomes, genomes and therapeutic responses between the high- and low-risk groups. We also examined the performance of the model for immunotherapeutic efficacy prediction. We believe that this model can provide clinicians with a novel reference for prognosis evaluation and clinical decision-making.

Materials and methods

Data collection

The genomic data, transcriptome profiles and clinical information from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database were downloaded from https://www.cbioportal.org. Normalized gene expression profiles of GSE103091, GSE21653 and GSE20685 were obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). We extracted TNBC patients from all breast cancer patients in the METABRIC (n=205), GSE103091 (n=107) and GSE21653 (n=73) cohorts according to immunohistochemical (IHC) staining of ER, PR and HER2. However, for the GSE20685 cohort, basal-like patients (n=49) were identified by PAM50 subtype analysis due to the lack of IHC information. In addition, genes with an expression level of 0 in more than 30% of the samples in all of the above cohorts were excluded, and only patients with a follow-up time of more than one month were included in the study. The METABRIC dataset was used as the training cohort, and three GEO datasets were used for external validation. The GSE106977 cohort that received neoadjuvant chemotherapy was used to assess the association between the risk score and chemotherapy response. The TNBC cohort (GSE194040), metastatic melanoma cohorts (GSE35640, GSE78220) and metastatic urothelial carcinoma cohort (IMvigor210) were used to examine the distribution of the risk score in non-responders and responders to immunotherapy.

Evaluation of immune cell infiltration

To estimate the abundances of immune cells, we calculated the ESTIMATE score, stromal score, immune score and tumor purity by the ESTIMATE algorithm provided in the R package "estimate" [18]. Furthermore, we quantified the relative abundances of immune and stromal cell subsets using three algorithms named TIMER [19], xCell [20] and CIBERSORT [21]. The CYT score, which is an RNA-based measure of immune cytolytic activity in tumors, is defined as the geometric mean of two key cytolytic effectors, granzyme A (GZMA) and perforin (PRF1) [22].

Establishment of the immune-related risk (IRR) score

We referred to the published articles to establish a set of 29 immune signatures containing 680 genes, including immune cell subsets, immune-related pathways and functions (Table S1) [23, 24]. Single-sample gene set enrichment analysis (ssGSEA) was performed to estimate enrichment scores in each sample using the R package "GSVA" [25]. Meanwhile, unsupervised hierarchical clustering was conducted

to classify the samples into two subgroups. The differentially expressed genes (DEGs) were identified by the R package "limma" according to the cutoff of |Log2FC| > 0.5 and *p* value < 0.01. Subsequently, to screen out reliable prognostic markers, candidate genes with a *p* value < 0.05 in univariate Cox regression were input into the least absolute shrinkage and selection operator (LASSO) Cox regression analysis, and tenfold cross-validation was performed to avoid overfitting. Finally, 12 robust genes were identified to construct an immune-related risk (IRR) score based on their normalized expression values and LASSO Cox coefficients. The calculation formula was as follows:

$$IRR = \sum Coefficient(mRNA_i) \times Expression(mRNA_i)$$

Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted with the "enrichGO" and "enrichKEGG" functions in the R package "clusterProfiler". Significantly enriched functions or pathways (*p* value < 0.05) were visualized by the R package "GOplot". For GSEA, hallmark gene sets were retrieved from the Molecular Signatures Database (MSigDB), and enrichment levels of corresponding gene sets were evaluated by the "GSEA" function provided in the R package "clusterProfiler". ssGSEA was achieved using the "ssGSEA" method implemented in the R package "GSVA" based on the gene sets obtained from the MSigDB or publicly published articles.

Analysis of genomic data

Somatic mutation data of the METABRIC samples were organized into mutation annotation format (MAF) files and analyzed using the R package "maftools". Significantly mutated genes (*p* value < 0.05) between the high- and low-IRR groups were identified and visualized by the "forestPlot" and "onOncoplot" functions in the "maftools" package. Tumor mutation burden (TMB) was defined as the total number of nonsynonymous somatic mutations per million bases, with 38 Mb as the number of variants or estimate of the exome size.

HRD score

Allelic imbalance extending to the telomere (ntAI) was defined as the number of regions with allelic imbalance longer than 11 Mb that extend to one of the subtelomeres but do not cross the centromere. Loss of heterozygosity (LOH) was defined as the number of LOH regions shorter than the whole chromosome and longer than 15 Mb. Large-scale state transition (LST) was defined as chromosome breakpoints between adjacent regions longer than 10 Mb after filtering shorter than 3 Mb. The homologous recombination deficiency (HRD) score was calculated as the sum of ntAI, LST, and LOH scores [26].

Estimation of therapeutic sensitivity

Immunotherapeutic response was inferred with the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm (http://tide.dfci.harvard.edu), immunophenoscore (IPS) calculation and submap analysis (https://cloud.genepattern.org/gp) [27-29]. Normally, a lower TIDE score and a higher IPS score are positively correlated with an active immune response. The response to chemotherapy drugs in the METABRIC samples was predicted using the public Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www. cancerrxgene.org). The half maximal inhibitory concentration (IC50) of the anticancer drug was estimated by the R package "oncoPredict".

Statistical analysis

Kaplan-Meier plots were applied to describe survival curves, and the log-rank test was used to determine significance. Univariate and multivariate Cox regression analyses were conducted to calculate the hazard ratio (HR) and identify significant predictors associated with survival. Time-dependent receiver operating characteristic (tROC) analysis was used to examine the predictive capacity of variables via the R package "timeROC". A scoring nomogram was generated, and a calibration curve was plotted using the R package "rms". Statistical differences in the distribution of continuous variables between two groups were examined by the Wilcoxon test, and those of three or more groups were compared by the Kruskal-Wallis test. Correlation analysis was conducted with Spearman's correlation. The association between the IRR score and therapy response was examined using the chi-square test. A two-sided *p* value < 0.05 was considered statistically significant. All statistical analyses were performed using R software (version 4.1.1).

Results

Immune profiling identified immune-active and immune-inactive clusters

The workflow diagram illustrating the design of our study is shown in Figure S1. We first evaluated the immune cell infiltration of each TNBC sample by ssGSEA according to the 29 immune signatures. Then, we performed unsupervised hierarchical clustering based on the Euclidean distance of the 29 signature scores in 205 TNBC patients. Finally, all the samples were divided into two clusters. Cluster 1 was characterized by low immune cell infiltration and a weak immune response, so we treated it as an immune-inactive group. Cluster 2 was characterized by high immune cell infiltration and a relatively strong immune response, and we classified it as an immune-active group (**Figure 1A**). Compared with the immune-active group, the immune-inactive group exhibited lower immune, stromal and ESTIMATE scores and higher tumor purity (**Figure 1B**). Moreover, Kaplan-Meier analyses indicated that the immune-inactive group had a significantly worse prognosis in terms of overall survival (OS), breast cancer-specific survival (BCSS), and relapse-free survival (RFS) (Figure S2A-C) than the immune-active group.

The IRR signature was constructed based on 12 robust prognostic genes

To explore immune response-associated factors, we identified 640 DEGs between the immune-active and immune-inactive groups, including 557 upregulated genes and 83 downregulated genes in the immune-active group (**Figure 1C**). GO analysis demonstrated that upregulated DEGs were mainly enriched in immune-related biological processes. However, downregulated DEGs were enriched in some oncogenic biological processes involved in keratinocyte differentiation, cell-cell junctions and embryonic organ development (Figure S2D). Similarly, KEGG enrichment analysis also indicated the upregulation of immune response pathways and the downregulation of carcinogenesis-related pathways (Figure S2E). Subsequently, 243 candidate DEGs significantly associated with prognosis in the univariate Cox regression analysis were input into the LASSO Cox regression analysis. Tenfold cross-validation was applied to overcome overfitting. With the λ value of the minimum criteria selected, 12 genes (CA9, C7orf68, PARM1, MATK, PLCL2, HLA-DRB4, RAMP3, TMEM176A, COBL, STAMBPL1, KLRD1, PHF15) with corresponding LASSO coefficients were incorporated into the IRR signature construction (Figure S3A, S3B). Therefore, the final risk model was as follows: IRR score = 0.18546 × CA9 + 0.05457 × C7orf68 - 0.05881 × PARM1 - 0.14190 × MATK - 0.19422 × PLCL2 - 0.19703 × HLA-DRB4 - 0.39867 × RAMP3 - 0.54037 × TMEM176A - 0.69667 × COBL - 0.77352 × STAMBPL1 - 0.78868 × KLRD1 - 0.95601 × PHF15. The distribution of LASSO coefficients of the 12 genes was displayed using a lollipop plot (**Figure 1D**). Correlation analyses demonstrated that these 12 genes were strongly correlated with each other (**Figure 1E**). Furthermore, we divided the 205 METABRIC samples into high- and low-IRR groups according to the median IRR score. The expression profile of the 12 genes indicated that CA9 and C7orf68 were more highly expressed in the high-IRR group, whereas the others presented higher expression in the low-IRR group (**Figure 1F**). Principal component analysis (PCA) showed that patients in different risk groups were separated into two distinct clusters using the 12 hub genes in the METABRIC and three external GEO datasets (Figure S3C-F), which confirmed the uniqueness and discriminative capacity of the 12 hub genes for TNBC patients.

The IRR score served as an indicator of prognosis in the METABRIC and three GEO datasets

To further evaluate the prognostic value of the IRR score, we investigated the survival differences in the high- and low-IRR groups using the Kaplan-Meier method and analyzed the distribution of the IRR score in different clinical subgroups. The Sankey diagram illustrated the flow from the two immune clusters to different risk groups and vital survival statuses (Figure S4A). Moreover, we observed that the IRR Z score was significantly elevated in patients who died during follow-up and in patients with higher TNM stages (**Figure 2A**). Kaplan-Meier analyses demonstrated that a higher IRR score indicated poorer OS, BCSS, and RFS in the METABRIC dataset (**Figure 2B-D**). In the subgroup analyses concerning age and stage for the METABRIC cohort, patients with higher IRR scores also exhibited worse OS than those with

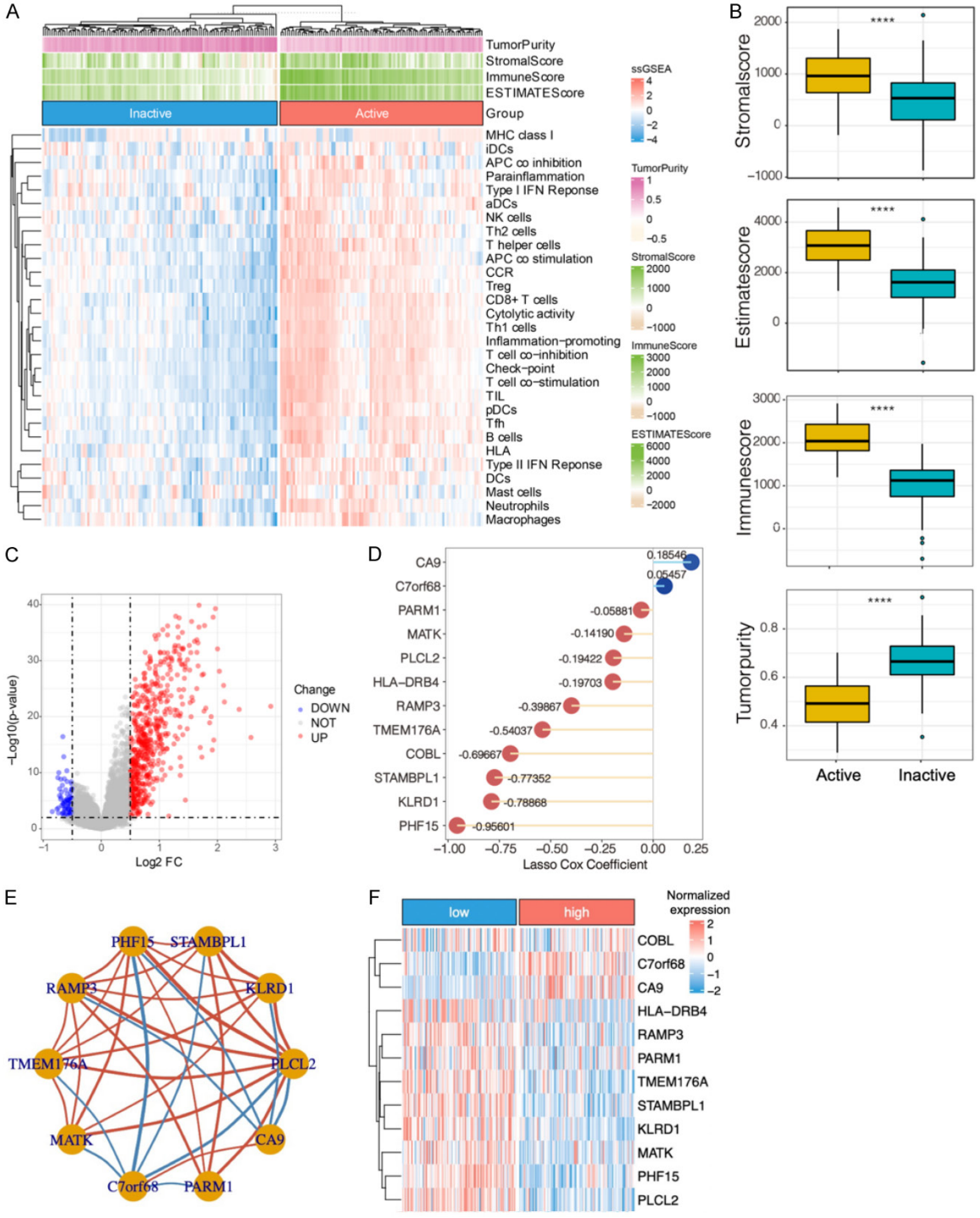


Figure 1. Construction of the immune-related risk gene signature. A. Heatmap depicting the enrichment levels of immune cell types and immune-associated pathways in the immune-active and immune-inactive clusters. B. Boxplots showing differences in immune, stromal and ESTIMATE scores and tumor purity between the immune-active and immune-inactive groups. C. Volcano plot of DEGs in the immune-active group versus the immune-inactive group (|Log2FC| > 0.5; *p* value < 0.01). D. Lollipop plot showing the LASSO coefficients of 12 robust markers included in the prognostic model. E. Correlation of the 12 hub genes. Red and blue lines represent the positive and negative correlations between genes, respectively. The thicker the line is, the stronger the correlation. F. Normalized expression profile of the 12 hub genes in the high- and low-IRR groups.

lower IRR scores (Figure S4B-F). The distribution of the IRR score in patients with different clinical outcomes and its prognostic value were externally validated in three independent GEO cohorts (Figure S5A-G). The areas under the ROC curves of the IRR score for 3-year, 5-year, and 10-year survival were 0.68, 0.70, and 0.72, respectively (**Figure 2E**). In the external GEO cohorts, the areas under the ROC curves of the IRR score ranged from approximately 0.6 to 0.8, indicating good prognostic efficiency of the IRR score (Figure S5H-J). In addition, univariate and multivariate Cox regression analyses demonstrated that the IRR score was an independent prognostic factor for the METABRIC and three GEO datasets (**Figure 2F**; Table S2). Then, we built a nomogram based on the IRR score together with age and stage for predicting 3-, 5-, and 10-year OS (**Figure 2G**). In the calibration analysis, the nomogram performance was extremely close to that of an ideal model, suggesting a high accuracy of prediction (**Figure 2H**). Decision curve analysis (DCA) showed that the nomogram had a potential clinical application with more net benefit than either the treat-none scheme or the treat-all scheme (**Figure 2I**). Furthermore, tROC analysis showed that compared with age and stage, the nomogram presented much more accuracy and stability for 3-, 5-, and 10-year OS prediction (**Figure 2J**).

High-IRR patients had TIME characteristics distinct from those of low-IRR patients

To further investigate the underlying mechanisms leading to the different outcomes in the low- and high-IRR groups, we analyzed the association between TIME components and IRR score through a series of algorithms. As illustrated in **Figure 3A**, the IRR score exhibited a negative correlation with the stromal, immune and cytolytic activity (CYT) scores but a positive correlation with tumor purity, which indicated less immune activation in the high-IRR group. We then compared the differences in the proportions of immune cell types inferred by the TIMER, CIBERSORT and xCell algorithms between the high- and low-IRR groups. The CIBERSORT results showed that macrophage cells generally occupied the largest proportion of all immune cell types in the METABRIC samples (Figure S6A). In addition, many immune cell types were discriminately distributed between the high- and low-IRR groups. The low-IRR group had more immune cell infiltration than the high-IRR group, especially effector T cells, NK cells and DCs (**Figures 3B**, **3C**, S6B). The IRR score was negatively correlated with CD8 T cells and active NK cells, which are the main effector cells that kill tumors in the TIME (**Figure 3D**). Remarkably, the M2 macrophage/(M2 macrophage + M1 macrophage) ratio of the high-IRR group was significantly higher than that of the low-IRR group (**Figure 3E**), indicating an M2 phenotype in the high-IRR group, but we did not observe a significant difference in the Th2/(Th1 + Th2) ratio between the two groups (**Figure 3F**). Moreover, among these hub genes, COBL showed only a weakly positive correlation with M1 macrophages and follicular helper T cells and a weakly negative correlation with activated memory CD4 T cells. CA9 and C7orf-68 were mainly positively correlated with macrophages, activated mast cells and neutrophils but negatively correlated with CD4/CD8 T cells. Overall, the other genes mainly exhibited a moderately negative correlation with M0/M2 macrophages and Treg cells but a positive correlation with effector T cells (**Figure 3G**). The correlation analysis provided additional proof for the protective or risk roles of these robust genes in TNBC progression. Taken together, this evidence demonstrated significant differences in intrinsic tumor immunogenicity and TIME components between the two risk groups.

Upregulation of the glycolysis pathway and suppression of innate immunity might contribute to the suppressive TIME

Tumor immunogenicity, oncogenic pathway activity and altered immune checkpoint expression are commonly acknowledged causes of immune escape. To clarify the intrinsic immune escape mechanisms associated with IRR, we compared the differences between the high- and low-IRR groups in these aspects. We found that the IRR score showed significantly positive correlations with the ntAI, LOH and HRD scores but not with the LST score or TMB. Although the significance level of the correlation between TMB and IRR score did not reach the standard (p value < 0.05), the TMB level in the high-IRR group was significantly higher than that in the low-IRR group (**Figure 4A**). Subsequently, we performed GSVA based on the KEGG gene sets in the METABRIC and three GEO cohorts. The volcano plots indicated that most pathways were upregulated in the low-IRR groups (Figure S7A). Finally, an intersection including eighteen

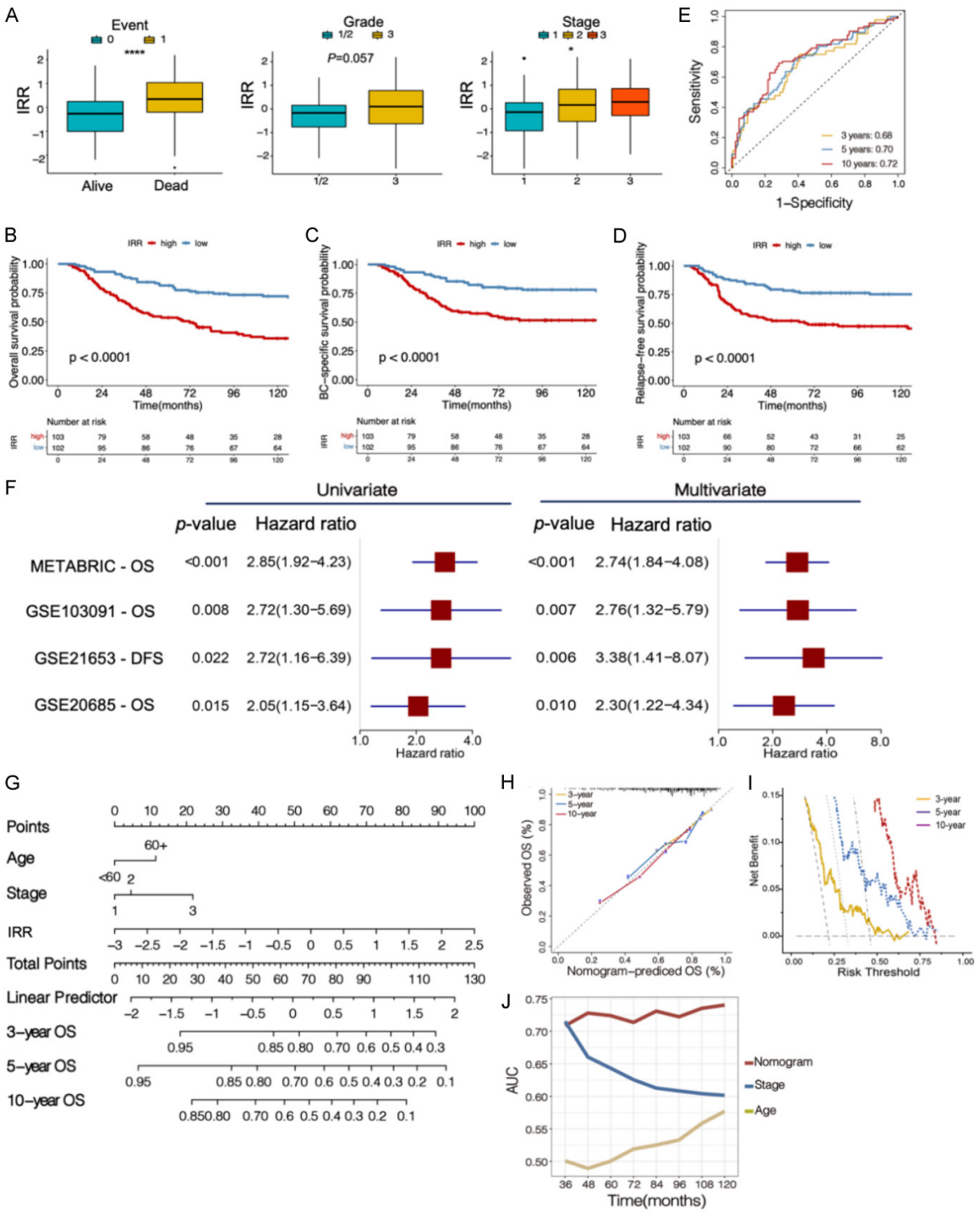


Figure 2. Evaluation and validation of the prognostic significance of the IRR signature. (A) Distribution of the IRR score in different risk, grade and stage subgroups. (B-D) Kaplan-Meier survival curves of OS (B), BCSS (C) and RFS (D) in the high- and low-IRR groups. (E) ROC curves of the IRR score with regard to 3-, 5- and 10-year OS. (F) Forest plots of univariate and multivariate Cox regression analyses in the METABRIC and GEO datasets. (G) A nomogram was established for predicting the probabilities of 3-, 5- and 10-year OS. (H) Calibration curves for assessing the agreement between nomogram-predicted 3-, 5- and 10-year OS and observed OS. (I) DCA curves for evaluating the potential of nomogram application clinically at 3, 5 and 10 years. (J) Time-dependent AUC value for the nomogram, stage and age in the METABRIC dataset.

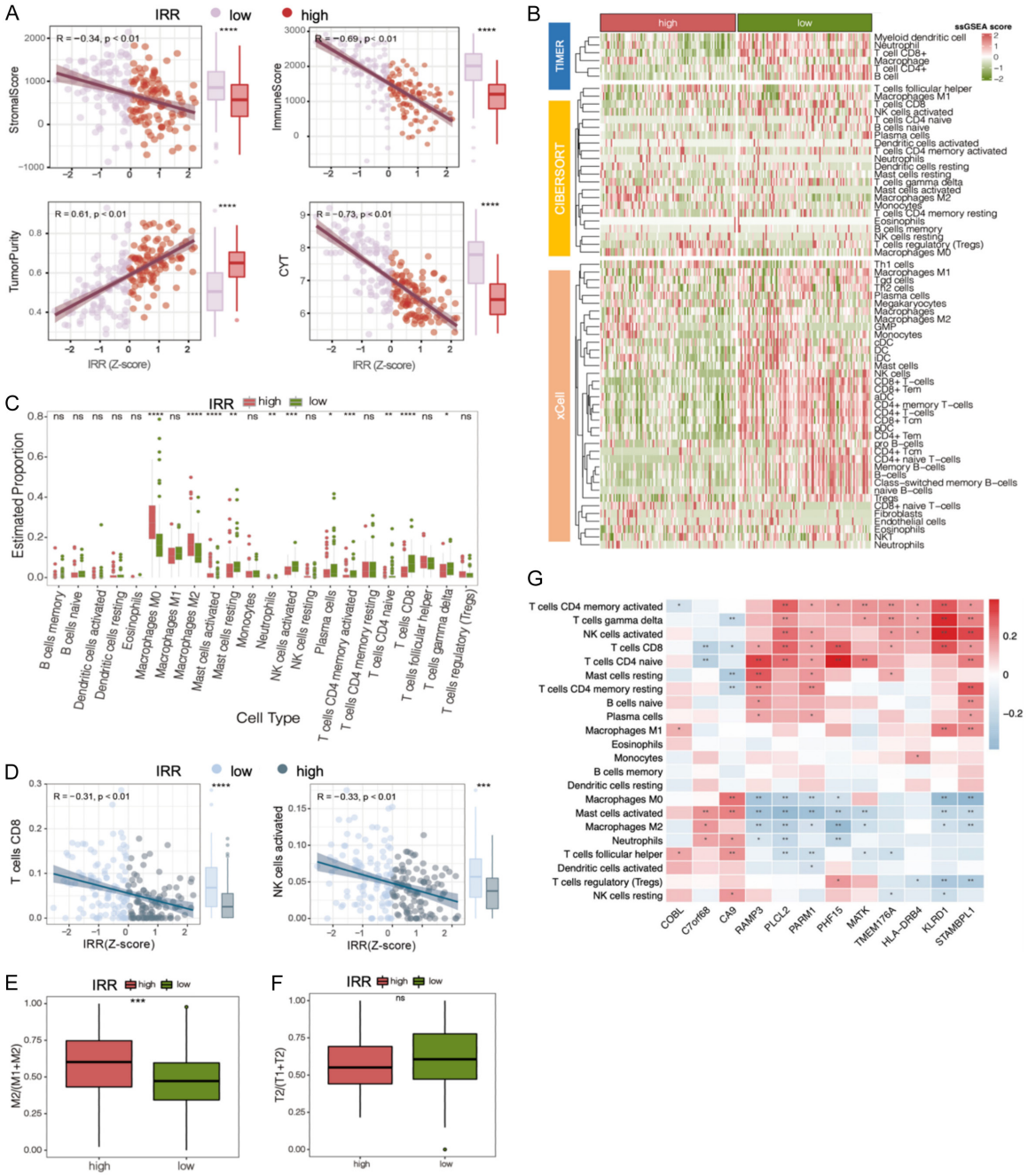


Figure 3. The landscape of TIME components and their correlations with IRR. (A) Association of the IRR score with the immune score, stromal score, tumor purity and CYT score in the METABRIC cohort. (B) Heatmap displaying immune and stromal cell infiltration levels in the high- and low-IRR groups. (C) Comparison of the estimated proportions of 22 immune cell types in the high- and low-IRR groups. (D) Association of the IRR score with CD8 T cells (left) and activated NK cells (right). (E, F) Comparison of the ratio of M2 macrophages/(M1 + M2 macrophages) (E) and Th2 cells/(Th1 + Th2) (F) in the high- and low-IRR groups. (G) Correlation matrix of robust prognostic genes and the abundance of 22 immune cell types estimated by the CIBERSORT algorithm.

upregulated pathways in the low-IRR groups of the four datasets was identified using a Venn diagram (Figure S7B). Most of the eighteen pathways were involved in immune activation and response (Figure S7C, S7D). The GSEA results also validated the discovery in the GSVA that many immune-related pathways were notably upregulated in the low-IRR group

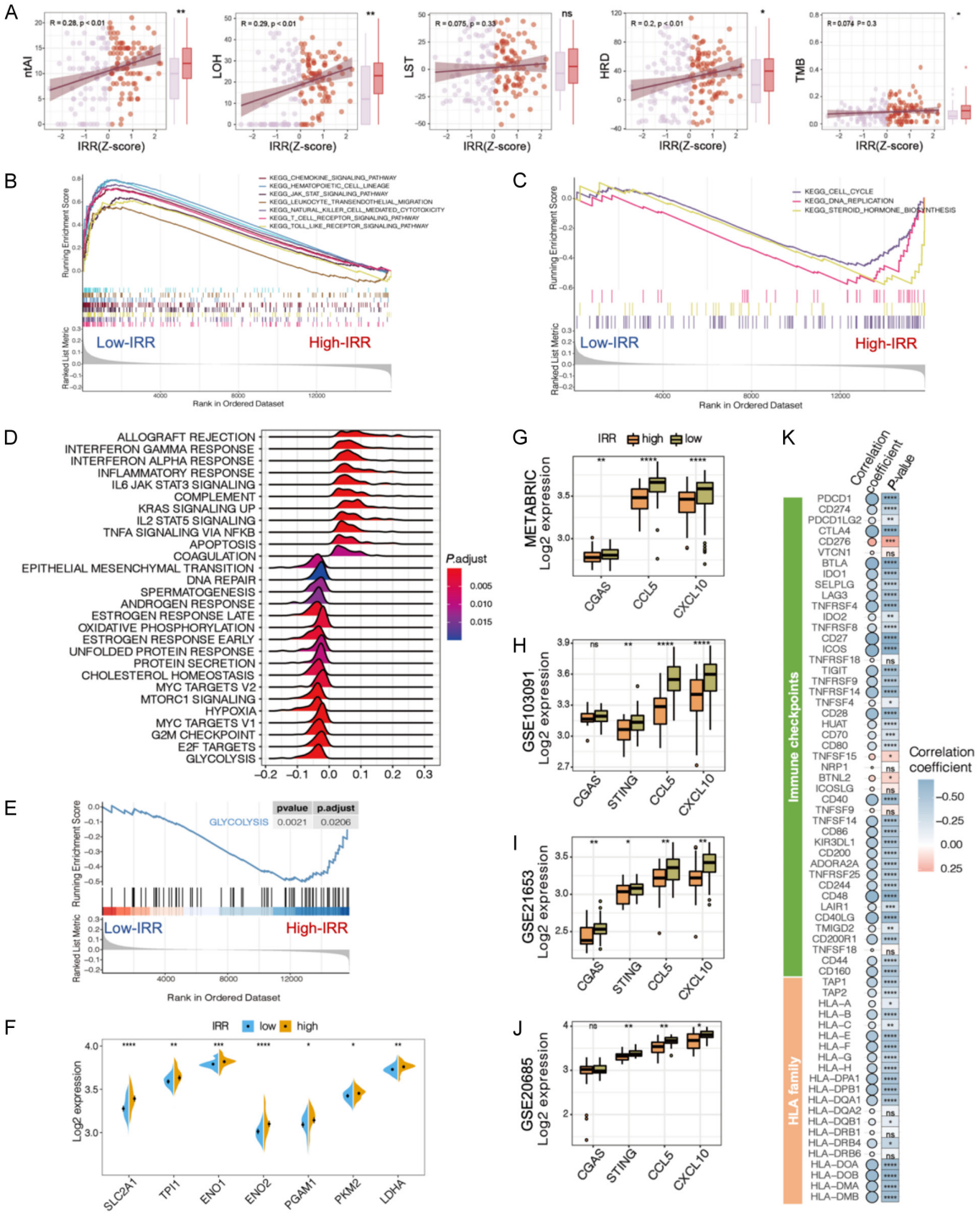


Figure 4. Underlying immune escape mechanisms in high-IRR TNBC patients. (A) Correlation analyses for ntAI, LOH, LST, HRD and TMB with the IRR score. (B, C) GSEA showed KEGG pathways enriched in the low-IRR (B) and high-IRR groups (C). (D) Ridge plots showing HALLMARK signature differences identified by GSEA between the low-IRR and high-IRR groups. (E) GSEA showed that the glycolysis pathway was enriched in the high-IRR group versus the low-IRR group. (F) Comparison of the expression levels of glycolytic pathway-associated genes in the high- and low-IRR groups. (G-J) The expression differences in cGAS-STING pathway-associated genes in the high- and low-IRR groups. (K) Correlation analysis of the IRR score with the expression levels of immune checkpoint and HLA family genes.

(**Figure 4B**). Some pathways involved in the cell cycle, DNA replication and steroid hormone biosynthesis were upregulated in the high-IRR group (**Figure 4C**). Likewise, GSEA according to

the HALLMARK gene sets showed that immune-related pathways were more enriched in the low-IRR group. Meanwhile, in the high-IRR group, we observed more enrichment in oncogenic pathways, such as epithelial-mesenchymal transition, MYC, hypoxia, G2/M checkpoint and glycolysis pathways (**Figure 4D**). As shown in **Figure 4E**, the glycolysis pathway was upregulated in the high-IRR group. We then examined the expression levels of glycolysis-associated genes in different risk groups. The half-violin plot indicated that all the genes showed significantly higher mRNA levels in the high-IRR group than in the low-IRR group (**Figure 4F**). In addition, the Hippo pathway and Notch pathway were significantly more enriched in the high-IRR group than in the low-IRR group of the METABRIC dataset (Figure S8A, S8B). Considering that the initiation and maintenance of the T-cell response are closely associated with the innate immune response, we further examined the key initiation molecules and downstream effector factors involved in innate immunity. As shown in the boxplots (**Figure 4G-J**), the expression levels of cGAS were elevated in the low-IRR group in all cohorts, and STING mRNA levels in the low-IRR groups were also higher than those in the high-IRR groups of the three GEO cohorts. CCL5 and CXCL10, typical effector factors in innate immunity, were significantly upregulated in the low-IRR groups compared with those in the high-IRR groups in three GEO cohorts. Furthermore, we analyzed the correlations of the IRR score with immune checkpoints and HLA family molecules (**Figure 4K**). For immune checkpoints, only CD276, TNFSF15 and BTNL2 showed a positive correlation with the IRR score. CD276 and TNFSF15 expression levels were higher in the high-IRR group than in the low-IRR group. Most of the other immune checkpoints were negatively correlated with the IRR score and upregulated in the low-IRR group (Figure S9A). Most of the HLA family molecules related to antigen presentation were negatively correlated with the IRR score and upregulated in the low-IRR group (Figure S9B). These results indicated that the high IRR samples showed lower immune-related pathway enrichment, more oncogenic pathway upregulation and lack of innate immunity activation, to some extent leading to low immune infiltration and an immunosuppressive microenvironment.

Somatic mutation landscapes of high-IRR and low-IRR patients

To explore the genomic alterations in different risk groups, we analyzed somatic mutations in the METABRIC samples. The Maftools analysis results demonstrated a relatively higher mutant frequency in the high-IRR samples than in the low-IRR samples. The top twenty highly mutated genes in the high- and low-IRR groups were displayed using waterfall plots (**Figure 5A**, **5B**). TP53, which is frequently mutated in TNBC, was the most highly mutated gene in both groups. Next, we observed twelve co-occurrence events in the high-IRR group and nine entirely different co-occurrence events in the low-IRR group. Interestingly, there was only one unique mutually exclusive case in the high-IRR group (PIK3CA-TP53), which might indicate their competitive effects and selective advantages in keeping one mutation between them (**Figure 5C**). With a threshold of p value < 0.05, the top five differentially mutated genes were HERC2, TP53, PIK3R1, ARID1A and KMT2C, but only HERC2 and TP53 showed significantly higher mutation frequencies in the low-IRR group (**Figure 5D**, **5E**). Since there was only one sample with the HERC2 mutation in the low-IRR group, we could not compare the difference in samples with HERC2 mutation in the two groups in detail. Next, we investigated TP53 mutation, the most frequently mutated and significantly more mutated gene in the high-IRR group. We created a lollipop plot to depict the mutation sites for TP53 (**Figure 5F**). Compared with TP53-wild-type samples, the IRR score was significantly elevated in TP53-mutant samples (**Figure 5G**). Furthermore, we found that for the TP53-mutant patients, the high-IRR group had a poorer prognosis than the low-IRR group (**Figure 5H**), suggesting an excellent discriminatory ability of the IRR score with the existence of TP53 mutation.

Patients with a low IRR showed more sensitivity to adjuvant therapy and were more likely to benefit from immunotherapy

Drug sensitivities forecasted by the GDSC database indicated that the estimated IC50 indexes of seven drugs, including docetaxel, epirubicin, gemcitabine, cisplatin, olaparib, 5'-flurouracil, and cyclophosphamide, showed a significant increase in the high-IRR group compared with

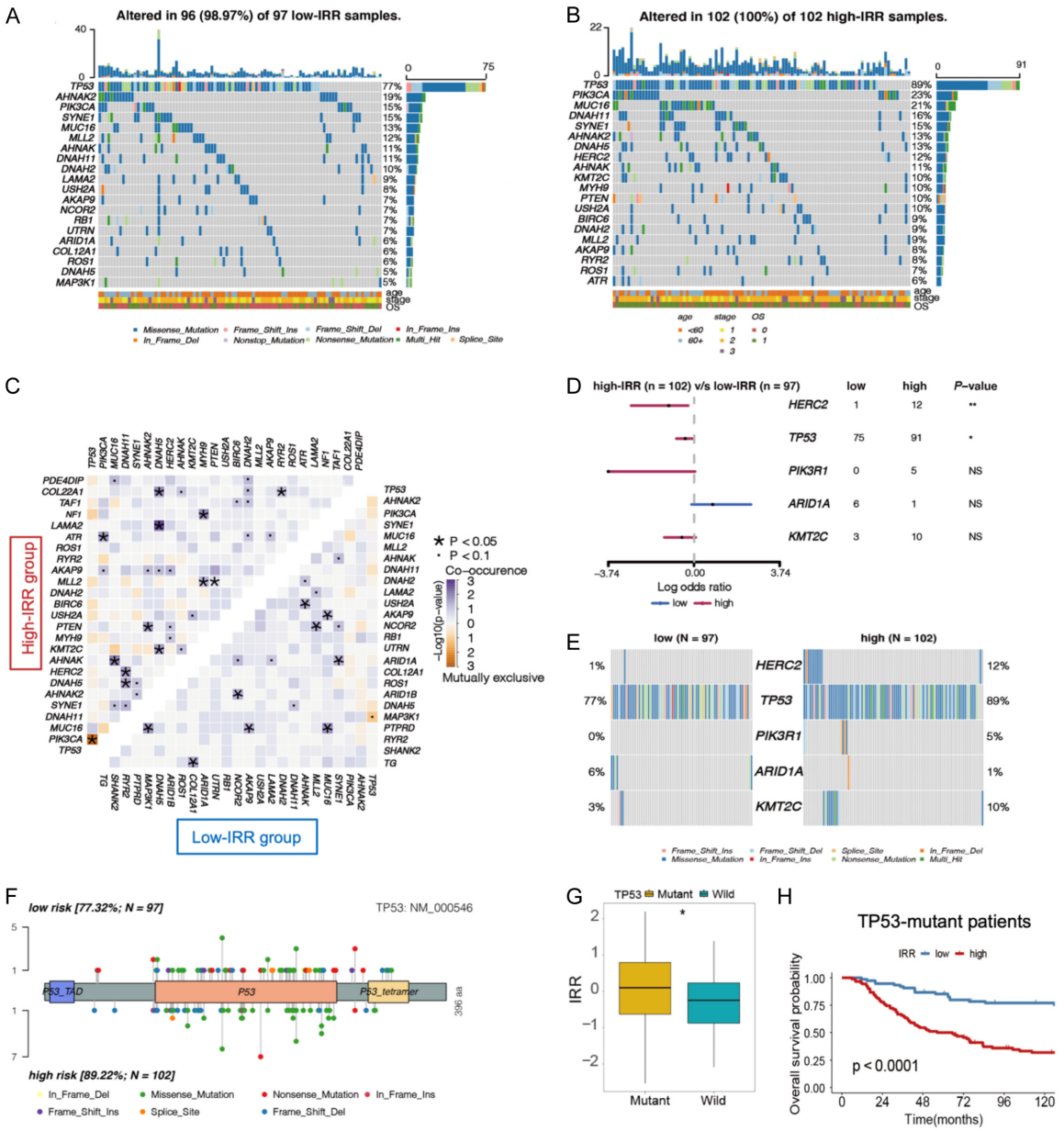


Figure 5. The landscape of genomic alterations in different IRR groups in the METABRIC cohort. (A, B) Waterfall plots illustrating the distribution of the top 20 highly mutated genes in the low-IRR (A) and high-IRR (B) groups. (C) Co-occurrence and coexclusive mutations in the low-IRR and high-IRR groups. (D, E) Top 5 discriminately mutated genes between the high- and low-IRR patients. (F) Lollipop chart showing the mutation sites of the TP53 protein. (G) Comparison of IRR scores in the TP53-mutant and TP53-wild-type groups. (H) Kaplan-Meier analysis of OS in patients with mutant TP53 stratified by IRR score.

the low-IRR group (Figure S10A). Meanwhile, as illustrated in the GSEA results, the low-IRR samples showed a greater response to radiation and less doxorubicin resistance than the high-IRR samples (Figure S10B, S10C). Moreover, we analyzed the distribution of IRR scores in the subgroups with pathological complete response (pCR) and residual disease (RD) of 88 TNBC patients treated with neoadjuvant anthracycline and/or taxane chemotherapy from the GSE106977 dataset. We observed that the IRR score in the RD samples was significantly elevated compared with that in the pCR samples (Figure S10D). Patients with pCR accounted for a larger proportion in the low-IRR group than in the high-IRR group (Figure S10E).

Additionally, we stratified the METABRIC patients with or without adjuvant treatment by the IRR score. Kaplan-Meier analyses demonstrated that a high IRR indicated poorer OS, BCSS, and RFS, regardless of whether the patients received chemotherapy/radiotherapy (Figure S11).

Then, to investigate the predictive performance of the IRR score for immunotherapy in TNBC, we compared the IRR scores of 29 TNBC patients who accepted paclitaxel + pembrolizumab from the I-SPY2 clinical trial (GSE194040). Patients with complete response (CR) to anti-PD-1 therapy showed a tendency for lower IRR score (**Figure 6A**). The proportion of patients with CR was higher in the low-IRR group of the TNBC cohort (**Figure 6B**). The area under the curve (AUC) value of the IRR score for predicting immunotherapy response was 0.684. The sensitivity, specificity and accuracy were 84.2%, 50% and 76.2%, respectively (**Figure 6C**). In addition, in many solid tumors, including TNBC, excluded and ignored immunophenotypes are associated with poor response to immunotherapy. Therefore, we also compared the IRR scores of different immunophenotypes of TNBC and found that the inflamed subtype exhibited a lower IRR score (**Figure 6D**). The proportion of the inflamed subtype was also increased in the low-IRR group compared with the high-IRR group (**Figure 6E**). To further test the potency of the IRR score for the prediction of immunotherapy response, we analyzed the METABRIC dataset using the IPS score and TIDE algorithm. Generally, a higher IPS score and lower TIDE score indicated a greater possibility of benefiting from immunotherapy. The results showed that the low-IRR group exhibited a higher IPS score, suggesting a better response to ICI therapy (**Figure 6F**). As shown in the correlation analysis, the IRR score was negatively correlated with the major histocompatibility complex (MHC), effector cell (EC), and IPS scores but positively correlated with the suppressor cell (SC) score (**Figure 6G**). Additionally, the TIDE results indicated that the IRR score showed a positive correlation with the TIDE and exclusion scores and a negative correlation with the dysfunction score (Figure S12A). Four potential responders to immunotherapy predicted by the TIDE algorithm were all included in the low-IRR group. Compared with the non-responders, the IRR score of responders was significantly decreased (Figure S12B). Next, we estimated the response possibility to anti-CTLA-4 and anti-PD1 immunotherapy in the high- and low-IRR patients with a submap method. After Bonferroni correlation, only patients in the low-IRR group prospectively benefited from anti-PD1 therapy (**Figure 6H**). Finally, we chose three cohorts, GSE35640 (metastatic melanoma patients treated with recombinant MAGE-A3 antigen combined with an immunostimulant), GSE78220 (metastatic melanoma treated with PD1 inhibitor), and IMvigor210 (metastatic urothelial carcinoma treated with PD-L1 inhibitor), to test the predictive efficacy in other cancer types. For metastatic melanoma patients, we observed a significantly higher IRR score in the non-responders versus responders (**Figure 6I**). In addition, responders occupied a larger proportion than non-responders in the low-IRR group, and high-IRR patients tended to have a poorer prognosis than low-IRR patients (**Figure 6J**, **6K**). In the IMvigor210 cohort, there was no difference in responses to anti-PD-L1 therapy between the high-IRR and low-IRR groups, but the Kaplan-Meier plot still demonstrated that a higher IRR score indicated poorer OS (**Figure 6L**). Collectively, these discoveries revealed that the IRR score was a promising indicator for responses to immunotherapy.

In summary, the IRR score is a promising marker for characterizing tumor immune phenotypes, predicting response to ICI treatment and assessing the clinical outcomes for TNBC patients (**Figure 7**).

Discussion

Some studies have constructed risk models according to the biological characteristics of TNBC, such as autophagy, RNA binding protein, N6 methylation and protein-protein interaction networks [30-33]. With the development of basic and translational research, the heterogeneity of the TME of breast cancer has been increasingly understood. Emerging evidence has shown that TNBC has higher immune cell infiltration and correlates with prognosis. In our study, we constructed an IRR signature based on the DEGs between the “inflammatory/active” and “desert/inactive” tumors and validated it in different cohorts. The IRR model is a prognostic factor independent of TNM staging,

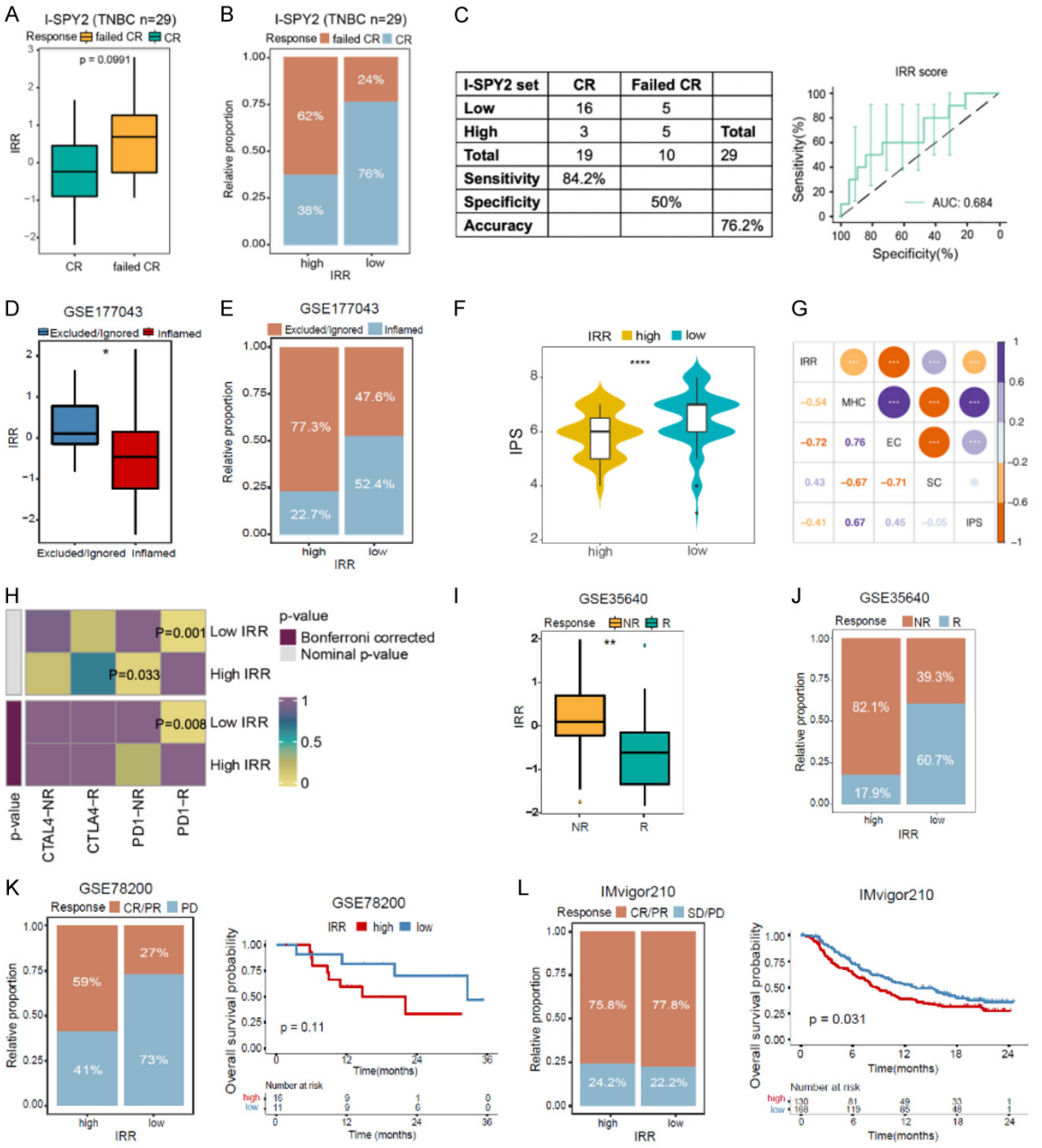
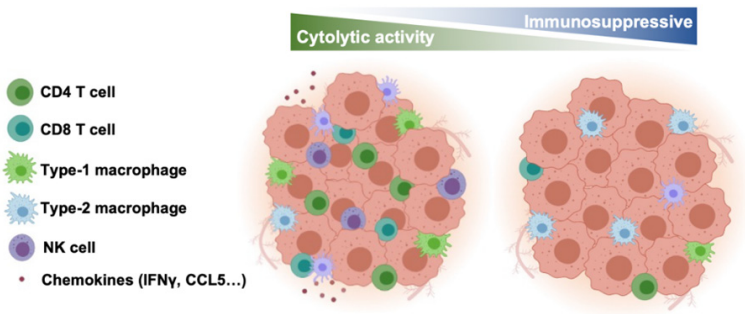


Figure 6. Predictive value of the IRR score for immunotherapeutic response. A. The box plot displayed that the patients with CR tended to have lower IRR score. B. Bar graph showing the proportions of different responses to paclitaxel + pembrolizumab in the high- and low-IRR groups of TNBC patients. C. Left panel: confusion matrices of immunotherapy responses and results predicted by the IRR score in TNBC patients. Right panel: ROC curve of the IRR score in predicting immunotherapy response. D. The box plot shows IRR scores of different immunophenotypes in the GSE177043 cohort. E. Bar graph showing the proportions of immunophenotypes in the high- and low-IRR groups of the GSE177043 cohort. F. Violin plot showing the distribution of IPS scores in the high- and low-IRR groups of the METABRIC cohort. G. Correlation between the IRR score and MHC, EC, SC, and IPS scores in the METABRIC cohort. H. Submap analysis indicated the response sensitivities to anti-PD-1 therapy and anti-CTLA therapy in the high- and low-IRR patients of the METABRIC cohort. I. The box plot showed that patients with better responses to recombinant MAGE-A3 antigen exhibited lower IRR scores in the GSE35640 cohort. J. Bar graph showing the proportions of different responses to recombinant MAGE-A3 antigen in the high- and low-IRR groups of the GSE35640 cohort. K. Bar graph showing the proportions of different responses to PD-1 inhibitors in the high- and low-IRR groups of the GSE78220 cohort. Kaplan-Meier survival curves of OS in the high- and low-IRR groups. L. Bar graph showing the proportions of different responses to PD-1 inhibitors in the high- and low-IRR groups of the IMvigor210 cohort. Kaplan-Meier survival curves of OS in the high- and low-IRR groups. NR, non-responder; R, responder.



	Low IRR	High IRR
TMB	Low	High
MHC class I/II	Up	Down
cGAS-STING	Up	Down
CTL level	High	Low
M2/M1 ratio	Low	High
Response to ICIs	Good	Poor
Prognosis	Good	Poor

Figure 7. The schema chart displaying the characteristics of high- and low-IRR patients. Compared with the high-IRR patients, low-IRR patients are characterized by intact antigen presentation, higher CTL levels, lower M2/M1 ratios and more stimulatory chemokines. Low-IRR patients also tend to have a better ICI response and favorable prognosis.

which can provide additional prognostic information for TNBC patients. TNBC patients with a high IRR exhibited an inhibitory TIME and highly activated oncogenic pathways, which may result in a worse prognosis and more resistance to chemotherapy, radiotherapy and immunotherapy. We revealed that the suppressive TIME in high-IRR patients might be caused by the hypoxic environment and upregulation of glycolytic activity. In addition, we evaluated the prediction potential of the IRR model for immunotherapy response. Our findings demonstrated that the IRR model could be used to evaluate the prognosis of TNBC patients and guide clinical decision-making, including immunotherapy-related choices.

Genes involved in the IRR model show positive or negative correlations with the immune response. Among the 12 genes, CA9 and C7orf68 are risk factors, whereas the other 10 genes are protective factors. Carbonic anhydrase 9 (CA9) is a transmembrane enzyme that plays a crucial role in pH regulation in the hypoxic TME [34]. Numerous studies have validated the importance of CA9 for promoting tumor progression and metastasis, including in breast cancer [35, 36]. C7orf68, also named hypoxia-inducible lipid droplet associated (HILPDA), can increase hypoxic lipid droplet formation, associate with the glycolysis pathway and enhance tumor immunosuppression status [37, 38]. High expression of these two genes in high-IRR patients reflects that hypoxia is one of the critical causes in the immunosuppressive microenvironment for promoting TNBC progression. HLA-DRB4, one of the protective factors, encodes the beta chain of MHC class II, which plays a central role in antigen presentation between antigen presenting cells (APCs) and CD4 T cells [39]. Intact antigen presentation between APCs and CD4 T cells maintains CTL activation when it is restricted by unsuccessful recognition of CTLs with tumor cells. With the auxiliary function of chemokines secreted by CD4 T cells, CTL activity can be stimulated effectively [40, 41]. In addition, megakaryocyte-associated tyrosine kinase (MATK), encoding a protein containing an amino acid sequence similar to that of CSK tyrosine kinase, is less expressed in human colorectal cancer, neuroblastoma and glioblastoma [42, 43]. Importantly, MATK downregulates HER2-mediated Src kinase activation by interacting with HER2 via the SH2 domain of CHK and acts as an inhibitory growth regulator in human breast cancer [44]. Moreover, MATK can reduce CXCR4 expression through YY1, leading to impairment of CXCR4/CXCL12-mediated breast cancer cell motility and migration [45]. For the other genes included in this model, their underlying mechanisms for regulating the malignant behavior of TNBC as well as their involvement within the TIME remain unclear; thus, further investigation is needed.

In our study, patients with a high IRR presented higher glycolytic activity, less T-cell infiltration, and a higher M2/M1 ratio. Generally, tumor cells tend to induce an inhibitory TIME that benefits immune evasion and tolerance [46, 47]. Multiple studies have confirmed that metabolic reprogramming is an important hallmark of tumors and is closely associated with the TIME.

Tumor cells are prone to exhibit enhanced glycolytic activity under aerobic or hypoxic conditions [48, 49]. High glucose consumption in tumor cells limits the nutrient sources required by effector cells and results in a low pH environment that is harmful to antitumor effectors such as T and NK cells [50, 51]. In addition, glycolysis upregulation influences the differentiation direction of tumor-associated macrophages (TAMs) [52]. For example, the glycolytic metabolite lactate can stabilize HIF-1α-mediated VEGF expression, activate GPR132, and prompt TAMs to differentiate into the protumor M2 type, finally promoting breast cancer cell metastasis [53]. Moreover, we also found that the Hippo pathway was highly upregulated in the high-IRR group compared with the low-IRR group. Recent studies have indicated that some oncogenic pathways play regulatory roles in cancer cell metabolism, such as the Hippo signaling pathway, whose activation and upregulation of its downstream effectors YAP and TAZ promote the glycolytic pathway to increase glucose uptake, lactate production and cell growth [54]. Thus, it can be inferred that the Hippo pathway is critical for the immunosuppressive microenvironment in the high-IRR group. Importantly, Roberta et al. found that inhibition of tumor glycolysis is beneficial to the therapeutic efficacy of CTLA-4 blockade [55]. Gong et al. reported that lactate dehydrogenase inhibitor therapy could sensitize TNBC with high glycolytic activity to ICIs [56]. Therefore, we proposed that ameliorating the hypoxic environment and inhibiting glycolytic activity have the potential to overcome the suppressive TIME and improve the immunotherapeutic response in high-IRR patients.

Generally, a higher TMB is associated with a higher neoantigen load, which is more likely to trigger T-cell infiltration and benefit from immunotherapy. Unexpectedly, in our study, although patients in the high-IRR group had a higher TMB than those in the low-IRR group, they showed lower immunogenicity and CTL infiltration. Similarly, one study indicated that TMB does not correlate with CD8 T-cell infiltration and immunotherapy efficacy in some tumor types, such as breast cancer and prostate cancer [57]. Denis et al. reported that 5% of TMB-low patients can respond well to ICIs, and > 50% of TMB-high patients do not respond [58]. These discoveries demonstrate that the immune response in the human body is a complex process. Multiple other variables in the immune system, such as antigen presentation and TIME characteristics, should also be taken into account in a composite model.

Finally, we assessed the performance of the IRR model for predicting immunotherapeutic response. The IRR score exhibited a favorable performance for predicting responses to anti-PD-1 therapy in 29 TNBC patients, which indicated that the IRR was a promising predictor for immunotherapy. In addition, we used the TIDE algorithm and the submap analysis to estimate the predictive efficiency for response to ICI therapy in different risk groups. The TIDE score is a combined evaluation of two kinds of tumor immune escape mechanisms. One is that cytotoxic T-cell infiltration is high in some tumors, but T cells remain in a dysfunctional state. The other is that T cells are excluded due to immunosuppressive factors in tumor tissue [27]. Interestingly, we observed a negative correlation between the IRR score and dysfunction score but a positive correlation between the IRR score and exclusion score. Therefore, we speculate that tumorigenesis in the low-IRR group is mainly associated with dysfunction of cytotoxic T cells, which indicates a favorable response to ICI treatment. In the high-IRR group, tumorigenesis is mainly induced by the exclusion of T cells and an immunosuppressive microenvironment. More interestingly, in the submap analysis, only patients in the low-IRR group showed promising responses to anti-PD-1 therapy after Bonferroni correlation. These analyses suggested that low-IRR TNBC patients are more sensitive to ICI therapy than patients with high IRR. Our model may be a useful tool to help clinicians conduct immunotherapies.

Although our model is robust in evaluating prognosis and predicting therapy response for TNBC patients, it should be emphasized that because our study is retrospective and relatively small-scale, further validation in larger prospective randomized trials is needed. Thus, future investigations should be performed to validate its predictive efficacy and compare it with other biomarkers.

In conclusion, we established a robust model for immune risk evaluation and immunotherapeutic response prediction. The IRR model

could be an effective tool to identify TNBC patients who may benefit from immunotherapy and facilitate the optimization of therapeutic regimens. Our findings also suggested that ICIs combined with glycolytic inhibitors may be a potential treatment strategy for high-IRR TNBC patients.

Acknowledgements

This work was supported by the 1·3·5 Project for disciplines of excellence (ZYGD18012); the Technological Innovation Project of Chengdu New Industrial Technology Research Institute (2017-CY02-00026-GX); Natural Science Foundation of Sichuan Province (2022NSFSC1385); Post-Doctor Research Project, West China Hospital, Sichuan University (2021HXBH080).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Hong Bu, Department of Pathology, Institute of Clinical Pathology, West China Hospital, Sichuan University, Chengdu 610041, China. E-mail: hongbu@scu.edu.cn; Dr. Luoting Yu, State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu 610041, China. E-mail: yuluot@scu.edu.cn

References

[1] Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, Ruddy K, Tsang J and Cardoso F. Breast cancer. Nat Rev Dis Primers 2019; 5: 66.

[2] Gupta GK, Collier AL, Lee D, Hoefer RA, Zheleva V, van Reesema LLS, Tang-Tan AM, Guye ML, Chang DZ, Winston JS, Samli B, Jansen RJ, Petricoin EF, Goetz MP, Bear HD and Tang AH. Perspectives on triple-negative breast cancer: current treatment strategies, unmet needs, and potential targets for future therapies. Cancers (Basel) 2020; 12: 1-33.

[3] Quinn C and Rakha EA. Eighth edition cancer staging manual of breast cancer by the American Joint Committee on cancer: are the new changes to improve staging or a treatment decision tool? J Clin Pathol 2018; 71: 1028-1029.

[4] Loi S, Michiels S, Adams S, Loibl S, Budczies J, Denkert C and Salgado R. The journey of tumor-infiltrating lymphocytes as a biomarker in breast cancer: clinical utility in an era of checkpoint inhibition. Ann Oncol 2021; 32: 1236-1244.

[5] Zhang J, Tian Q, Zhang M, Wang H, Wu L and Yang J. Immune-related biomarkers in triple-negative breast cancer. Breast Cancer 2021; 28: 792-805.

[6] Xu Y, Su GH, Ma D, Xiao Y, Shao ZM and Jiang YZ. Technological advances in cancer immunity: from immunogenomics to single-cell analysis and artificial intelligence. Signal Transduct Target Ther 2021; 6: 312.

[7] Sui S, An X, Xu C, Li Z, Hua Y, Huang G, Sui S, Long Q, Sui Y, Xiong Y, Ntim M, Guo W, Chen M, Deng W, Xiao X and Li M. An immune cell infiltration-based immune score model predicts prognosis and chemotherapy effects in breast cancer. Theranostics 2020; 10: 11938-11949.

[8] Wang S, Zhang Q, Yu C, Cao Y, Zuo Y and Yang L. Immune cell infiltration-based signature for prognosis and immunogenomic analysis in breast cancer. Brief Bioinform 2021; 22: 2020-2031.

[9] Kuroda H, Jamiyan T, Yamaguchi R, Kakumoto A, Abe A, Harada O and Masunaga A. Tumor-infiltrating B cells and T cells correlate with postoperative prognosis in triple-negative carcinoma of the breast. BMC Cancer 2021; 21: 286.

[10] Guo M and Wang SM. Genome instability-derived genes are novel prognostic biomarkers for triple-negative breast cancer. Front Cell Dev Biol 2021; 9: 1-13.

[11] Sun X, Luo H, Han C, Zhang Y and Yan C. Identification of a hypoxia-related molecular classification and hypoxic tumor microenvironment signature for predicting the prognosis of patients with triple-negative breast cancer. Front Oncol 2021; 11: 1-13.

[12] Huang Z, Pan J, Wang H, Du X, Xu Y, Wang Z and Chen D. Prognostic significance and tumor immune microenvironment heterogenicity of m5C RNA methylation regulators in triple-negative breast cancer. Front Cell Dev Biol 2021; 9: 1-15.

[13] Cortes J, Cescon DW, Rugo HS, Nowecki Z, Im SA, Yusof MM, Gallardo C, Lipatov O, Barrios CH, Holgado E, Iwata H, Masuda N, Otero MT, Gokmen E, Loi S, Guo Z, Zhao J, Aktan G, Karantza V and Schmid P; KEYNOTE-355 Investigators. 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 2020; 396: 1817-1828.

[14] Mittendorf EA, Zhang H, Barrios CH, Saji S, Jung KH, Hegg R, Koehler A, Sohn J, Iwata H, Telli ML, Ferrario C, Punie K, Penault-Llorca F, Patel S, Duc AN, Liste-Hermoso M, Maiya V, Molinero L, Chui SY and Harbeck N. Neoadjuvant atezolizumab in combination with se-

quential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): a randomised, double-blind, phase 3 tria. Lancet 2020; 396: 1090-1100.
[15] Emens LA, Adams S, Barrios CH, Dieras VC, Iwata H, Loi S, Rugo HS, Schneeweiss A, Winer EP, Patel S, Henschel V, Swat A, Kaul M, Molinero L, Chui SY and Schmid P. LBA16 IMpassion130: final OS analysis from the pivotal phase III study of atezolizumab + nab-paclitaxel vs placebo + nab-paclitaxel in previously untreated locally advanced or metastatic triple-negative breast cancer. Ann Oncol 2020; 31: S1148.
[16] Kossai M, Radosevic-Robin N and Penault-Llorca F. Refining patient selection for breast cancer immunotherapy: beyond PD-L1. ESMO Open 2021; 6: 100257.
[17] Zhu Y, Zhu X, Tang C, Guan X and Zhang W. Progress and challenges of immunotherapy in triple-negative breast cancer. Biochim Biophys Acta Rev Cancer 2021; 1876: 188593.
[18] Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB and Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013; 4: 2612.
[19] Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, Jiang P, Shen H, Aster JC, Rodig S, Signoretti S, Liu JS and Liu XS. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol 2016; 17: 174.
[20] Aran D, Hu Z and Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017; 18: 220.
[21] Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M and Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015; 12: 453-457.
[22] Rooney MS, Shukla SA, Wu CJ, Getz G and Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell 2015; 160: 48-61.
[23] Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A, Bruneval P, Fridman WH, Becker C, Pagès F, Speicher MR, Trajanoski Z and Galon JÔ. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 2013; 39: 782-795.
[24] Zhuang W, Sun H, Zhang S, Zhou Y, Weng W, Wu B, Ye T, Huang W, Lin Z, Shi L and Shi K. An immunogenomic signature for molecular classification in hepatocellular carcinoma. Mol Ther Nucleic Acids 2021; 25: 105-115.
[25] Hänzelmann S, Castelo R and Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics 2013; 14: 1-15.
[26] Timms KM, Abkevich V, Hughes E, Neff C, Reid J, Morris B, Kalva S, Potter J, Tran TV, Chen J, Iliev D, Sangale Z, Tikishvili E, Perry M, Zharkikh A, Gutin A and Lanchbury JS. Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. Breast Cancer Res 2014; 16: 475.
[27] Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, Wucherpfennig KW and Liu XS. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med 2018; 24: 1550-1558.
[28] Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H and Trajanoski Z. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep 2017; 18: 248-262.
[29] Lu X, Jiang L, Zhang L, Zhu Y, Hu W, Wang J, Ruan X, Xu Z, Meng X, Gao J, Su X and Yan F. Immune signature-based subtypes of cervical squamous cell carcinoma tightly associated with human papillomavirus type 16 expression, molecular features, and clinical outcome. Neoplasia 2019; 21: 591-601.
[30] Li X, Chen J, Yu Q, Huang H, Liu Z, Wang C, He Y, Zhang X, Li W, Li C, Zhao J and Long W. A signature of autophagy-related long non-coding RNA to predict the prognosis of breast cancer. Front Genet 2021; 12: 1-11.
[31] Liu Y, Sun H, Li X, Liu Q, Zhao Y, Li L, Xu B, Hou Y and Jin W. Identification of a three-RNA Binding Proteins (RBPs) signature predicting prognosis for breast cancer. Front Oncol 2021; 11: 1-14.
[32] Wang S, Zou X, Chen Y, Cho WC and Zhou X. Effect of N6-methyladenosine regulators on progression and prognosis of triple-negative breast cancer. Front Genet 2021; 11: 580036.
[33] Li M, Guo Y, Feng YM and Zhang N. Identification of triple-negative breast cancer genes and a novel high-risk breast cancer prediction model development based on ppi data and support vector machines. Front Genet 2019; 10: 180.
[34] Chafe SC, Vizeacoumar FS, Venkateswaran G, Nemirovsky O, Awrey S, Brown WS, McDonald PC, Carta F, Metcalfe A, Karasinska JM, Huang L, Muthuswamy SK, Schaeffer DF, Renouf DJ, Supuran CT, Vizeacoumar FJ and Dedhar S. Genome-wide synthetic lethal screen unveils

novel CAIX-NFS1/xCT axis as a targetable vulnerability in hypoxic solid tumors. Sci Adv 2021; 7: eabj0364.
[35] Güttler A, Eiselt Y, Funtan A, Thiel A, Petrenko M, Keßler J, Thondorf I, Paschke R, Vordermark D and Bache M. Betulin sulfonamides as carbonic anhydrase inhibitors and anticancer agents in breast cancer cells. Int J Mol Sci 2021; 22: 8808.
[36] Sarnella A, D'avino G, Hill BS, Alterio V, Winum JY, Supuran CT, De Simone G and Zannetti A. A novel inhibitor of carbonic anhydrases prevents hypoxia-induced TNBC cell plasticity. Int J Mol Sci 2020; 21: 1-18.
[37] VandeKopple MJ, Wu J, Auer EN, Giaccia AJ, Denko NC and Papandreou I. HILPDA regulates lipid metabolism, lipid droplet abundance, and response to microenvironmental stress in solid tumors. Mol Cancer Res 2019; 17: 2089-2101.
[38] Liu C, Zhou X, Zeng H, Wu D and Liu L. HILPDA Is a prognostic biomarker and correlates with macrophage infiltration in pan-cancer. Front Oncol 2021; 11: 597860.
[39] Zanetti M. Tapping CD4 T cells for cancer immunotherapy: the choice of personalized genomics. J Immunol 2015; 194: 2049-2056.
[40] Dosset M, Castro A, Carter H and Zanetti M. Telomerase and CD4 T cell immunity in cancer. Cancers (Basel) 2020; 12: 1-19.
[41] Marincola FM, Jaffee EM, Hicklin DJ and Ferrone S. Escape of human solid tumors from T-cell recognition: molecular mechanisms and functional significance. Adv Immunol 2000; 74: 181-273.
[42] Kim SO, Avraham S, Jiang S, Zagozdzon R, Fu Y and Avraham HK. Differential expression of Csk homologous kinase (CHK) in normal brain and brain tumors. Cancer 2004; 101: 1018-1027.
[43] Chüeh AC, Advani G, Foroutan M, Smith J, Ng N, Nandurkar H, Lio DS, Zhu HJ, Chong YP, Verkade H, Fujita DJ, Bjorge J, Basheer F, Lim JP, Luk I, Dhillon A, Sakthianandeswaren A, Mouradov D, Sieber O, Hollande F, Mariadason JM and Cheng HC. CSK-homologous kinase (CHK/MATK) is a potential colorectal cancer tumour suppressor gene epigenetically silenced by promoter methylation. Oncogene 2021; 40: 3015-3029.
[44] Kim S, Zagozdzon R, Meisler A, Baleja JD, Fu Y, Avraham S and Avraham H. Csk homologous kinase (CHK) and ErbB-2 interactions are directly coupled with CHK negative growth regulatory function in breast cancer. J Biol Chem 2002; 277: 36465-36470.
[45] Lee BC, Lee TH, Zagozdzon R, Avraham S, Usheva A and Avraham HK. Carboxyl-terminal Src kinase homologous kinase negatively regulates the chemokine receptor CXCR4 through YY1 and impairs CXCR4/CXCL12 (SDF-1alpha)-mediated breast cancer cell migration. Cancer Res 2005; 65: 2840-2845.
[46] Mortezaee K. Immune escape: a critical hallmark in solid tumors. Life Sci 2020; 258: 118110.
[47] Tang S, Ning Q, Yang L, Mo Z and Tang S. Mechanisms of immune escape in the cancer immune cycle. Int Immunopharmacol 2020; 86: 106700.
[48] Martínez-Reyes I and Chandel NS. Cancer metabolism: looking forward. Nat Rev Cancer 2021; 21: 669-680.
[49] Yan L, Tan Y, Chen G, Fan J and Zhang J. Harnessing metabolic reprogramming to improve cancer immunotherapy. Int J Mol Sci 2021; 22: 10268.
[50] Huber V, Camisaschi C, Berzi A, Ferro S, Lugini L, Triulzi T, Tuccitto A, Tagliabue E, Castelli C and Rivoltini L. Cancer acidity: an ultimate frontier of tumor immune escape and a novel target of immunomodulation. Semin Cancer Biol 2017; 43: 74-89.
[51] Ngwa VM, Edwards DN, Philip M and Chen J. Microenvironmental metabolism regulates anti-tumor immunity. Cancer Res 2019; 79: 4003.
[52] Colegio OR, Chu NQ, Szabo AL, Chu T, Rhebergen AM, Jairam V, Cyrus N, Brokowski CE, Eisenbarth SC, Phillips GM, Cline GW, Phillips AJ and Medzhitov R. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. Nature 2014; 513: 559-563.
[53] Chen P, Zuo H, Xiong H, Kolar MJ, Chu Q, Saghatelian A, Siegwart DJ and Wan Y. Gpr132 sensing of lactate mediates tumor-macrophage interplay to promote breast cancer metastasis. Proc Natl Acad Sci U S A 2017; 114: 580-585.
[54] Ardestani A, Lupse B and Maedler K. Hippo signaling: key emerging pathway in cellular and whole-body metabolism. Trends Endocrinol Metab 2018; 29: 492-509.
[55] Zappasodi R, Serganova I, Cohen IJ, Maeda M, Shindo M, Senbabaoglu Y, Watson MLJ, Leftin A, Maniyar R, Verma S, Lubin M, Ko M, Mane MM, Zhong H, Liu C, Ghosh A, Abu-Akeel M, Ackerstaff E, Koutcher JA, Ho PC, Delgoffe GM, Blasberg R, Wolchok JD and Merghoub T. CTLA-4 blockade drives loss of T reg stability in glycolysis-low tumours. Nature 2021; 591: 652-658.
[56] Gong Y, Ji P, Yang YS, Xie S, Yu TJ, Xiao Y, Jin ML, Ma D, Guo LW, Pei YC, Chai WJ, Li DQ, Bai F, Bertucci F, Hu X, Jiang YZ and Shao ZM. Metabolic-pathway-based subtyping of triple-negative breast cancer reveals potential therapeutic targets. Cell Metab 2021; 33: 51-64, e9.

[57] McGrail DJ, Pilié PG, Rashid NU, Voorwerk L, Slagter M, Kok M, Jonasch E, Khasraw M, Heimberger AB, Lim B, Ueno NT, Litton JK, Ferrarotto R, Chang JT, Moulder SL and Lin SY. High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types. Ann Oncol 2021; 32: 661-672.

[58] Jardim DL, Goodman A, de Melo Gagliato D and Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. Cancer Cell 2021; 39: 154-173.

Table S1. 29 Immune signatures and their contained genes

Cell types/Pathways	Genes
ADCS	CD83, LAMP3, CCL1
APC CO inhibition	C10orf54, CD274, LGALS9, PDCD1LG2, PVRL3
APC CO stimulation	CD40, CD58, CD70, ICOSLG, SLAMF1, TNFSF14, TNFSF15, TNFSF18, TNFSF4, TNFSF8, TNFSF9
B cells	BACH2, BANK1, BLK, BTLA, CD79A, CD79B, FCRL1, FCRL3, HVCN1, RALGPS2
CCR	CCL16, TPO, TGFBR2, CXCL2, CCL14, TGFBR3, IL11RA, CCL11, IL4I1, IL33, CXCL12, CXCL10, BMPER, BMP8A, CXCL11, IL21R, IL17B, TNFRSF9, ILF2, CX3CR1, CCR8, TNFSF12, CSF3, TNFSF4, BMP3, CX3CL1, BMP5, CXCR2, TNFRSF10D, BMP2, CXCL14, CCL28, CXCL3, BMP6, CCL21, CXCL9, CCL23, IL6, TNFRSF18, IL17RD, IL17D, IL27, CCL7, IL1R1, CXCR4, CXCR2P1, TGFB1I1, IFNGR1, IL9R, IL1RAPL1, IL11, CSF1, IL20RA, IL25, TNFRSF4, IL18, ILF3, CCL20, TNFRSF12A, IL6ST, CXCL13, IL12B, TNFRSF8, IL6R, BMPR2, IFNE, IL1RAPL2, IL3RA, BMP4, CCL24, TNFSF13B, CCR4, IL2RA, IL32, TNFRSF10C, IL22RA1, BMPR1A, CXCR5, CXCR3, IFNA8, IL17REL, IFNB1, IFNAR1, TNFRSF1B, CCL17, IFNL1, IL16, IL1RL1, ILK, CCL25, ILDR2, CXCR1, IL36RN, IL34, TGFB1, IFNG, IL19, ILKAP, BMP2K, CCR10, ILDR1, EPO, CCR7, IL17C, IL23A, CCR5, IL7, EPOR, CCL13, IL2RG, IL31RA, TNFAIP6, IFNL2, BMP1, IL12RB1, TNFAIP8, IL4R, TNFRSF6B, TNFAIP8L1, TNFRSF10B, IFNL3, CCL5, CXCL6, CXCL1, CCR3, TNFSF11, CSF1R, IL21, IL1RAP, IL12RB2, CCL1, IL17RA, CCR1, IL1RN, TNFRSF11B, TNFRSF14, IL13, IL2RB, BMP8B, CCL2, IL24, IL18RAP, TGFBI, TNFSF10, TNFRSF11A, CXCL5, IL5RA, TNFSF9, IL1RL2, TNFRSF13C, IL36G, IL15RA, TNFRSF21, CXCL8, IL22RA2, TNFAIP8L2, IL18R1, IFNLR1, CXCR6, CCL3L3, TNFRSF1A, IL17RE, IFNGR2, IL17RC, TNFAIP8L3, ILVBL, TGFBRAP1, CCL4L1, CSF2RA, CCRN4L, CCL26, TNFAIP1, CCRL2, IFNA10, TNFRSF17, IFNA13, IL20, IL18BP, CCL3L1, TNFSF12, TNFSF13, IL5, IL23R, IL26, TNF, TGFA, CSF2, IL1F10, CXCL17, TNFSF13, IFNA4, IL37, IL12A, IL7R, IFNA1, IL1A, IL4, IL2, CCL22, CSF3R, IL10, IFNK
CD8+ T cells	CD8A
Check-point	IDO1, LAG3, CTLA4, TNFRSF9, ICOS, CD80, PDCD1LG2, TIGIT, CD70, TNFSF9, ICOSLG, KIR3DL1, CD86, PDCD1, LAIR1, TNFRSF8, TNFSF15, TNFRSF14, IDO2, CD276, CD40, TNFRSF4, TNFSF14, HHLA2, CD244, CD274, HAVCR2, CD27, BTLA, LGALS9, TMIGD2, CD28, CD48, TNFRSF25, CD40LG, ADORA2A, VTCN1, CD160, CD44, TNFSF18, TNFRSF18, BTNL2, C10orf54, CD200R1, TNFSF4, CD200, NRP1
Cytolytic activity	PRF1, GZMA
DCS	CCL17, CCL22, CD209, CCL13
HLA	HLA-E, HLA-DPB2, HLA-C, HLA-J, HLA-DQB1, HLA-DQB2, HLA-DQA2, HLA-DQA1, HLA-A, HLA-DMA, HLA-DOB, HLA-DRB1, HLA-H, HLA-B, HLA-DRB5, HLA-DOA, HLA-DPB1, HLA-DRA, HLA-DRB6, HLA-L, HLA-F, HLA-G, HLA-DMB, HLA-DPA1
IDCS	CD1A, CD1E
Inflammation-promoting	CCL5, CD19, CD8B, CXCL10, CXCL13, CXCL9, GNLY, GZMB, IFNG, IL12A, IL12B, IRF1, PRF1, STAT1, TBX21
Macrophages	C11orf45, CD68, CLEC5A, CYBB, FUCA1, GPNMB, HS3ST2, LGMN, MMP9, TM4SF19
Mast cells	CMA1, MS4A2, TPSAB1
MHC class I	B2M, HLA-A, TAP1
Neutrophils	EVI2B, HSD17B11, KDM6B, MEGF9, MNDA, NLRP12, PADI4, SELL, TRANK1, V-NN3
NK cells	KLRC1, KLRF1
Parainflammation	CXCL10, PLAT, CCND1, LGMN, PLAUR, AIM2, MMP7, ICAM1, MX2, CXCL9, ANXA1, TLR2, PLA2G2D, ITGA2, MX1, HMOX1, CD276, TIRAP, IL33, PTGES, TNFRSF12A, SCARB1, CD14, BLNK, IFIT3, RETNLB, IFIT2, ISG15, OAS2, REL, OAS3, CD44, PPARG, BST2, OAS1, NOX1, PLA2G2A, IFIT1, IFITM3, IL1RN
PDCS	CLEC4C, CXCR3, GZMB, IL3RA, IRF7, IRF8, LILRA4, PHEX, PLD4, PTCRA
T cell co-inhibition	BTLA, C10orf54, CD160, CD244, CD274, CTLA4, HAVCR2, LAG3, LAIR1, TIGIT
T cell co-stimulation	CD2, CD226, CD27, CD28, CD40LG, ICOS, SLAMF1, TNFRSF18, TNFRSF25, TNFRSF4, TNFRSF8, TNFRSF9, TNFSF14
T helper cells	CD4
TFH	PDCD1, CXCL13, CXCR5
TH1 cells	IFNG, TBX21, CTLA4, STAT4, CD38, IL12RB2, LTA, CSF2
TH2 cells	PMCH, LAIR2, SMAD2, CXCR6, GATA3, IL26
TIL	ITM2C, CD38, THEMIS2, GLYR1, ICOS, F5, TIGIT, KLRD1, IRF4, PRKCQ, FCRL5, SIRPG, LPXN, IL2RG, CCL5, LCK, TRAF3IP3, CD86, MAL, LILRB1, DOK2, CD6, PAG1, LAX1, PLEK, PIK3CD, SLAMF1, XCL1, GPR171, XCL2, TBX21, CD2, CD53, KLHL6, SLAMF6, CD40, SIT1, TNFRSF4, CD79A, CD247, LCP2, CD3D, CD27, SH2D1A, FYB, ARHGAP30, ACAP1, CST7, CD3G, IL2RB, CD3E, FCRL3, CORO1A, ITK, TCL1A, CYBB, CSF2RB, IKZF1, NCF4, DOCK2, CCR2, PTPRC, PLAC8, NCKAP1L, IL7R, CD28, STAT4, CD8A, LY9, CD48, HCST, PTPRCAP, SASH3, ARHGAP25, LAT, TRAT1, IL10RA, PAX5, CCR7, DOCK11, PARVG, SPNS1, CD52, HCLS1, ARHGAP9, GIMAP6, PRKCB, MS4A1, GPR18, TBC1D10C, GVINP1, P2RY8, EVI2B, VAMP5, KLRK1, SELL, MPEG1, MS4A6A, ARHGAP15, MFNG, GZMK, SELPLG, TARP, GIMAP7, FAM65B, INPP5D, ITGA4, MZB1, GPSM3, STK10, CLEC2D, IL16, NLRC3, GIMAP5, GIMAP4, IFFO1, CFH, PVRIG, CFHR1
TREG	IL12RB2, TMPRSS6, CTSC, LAPTM4B, TFRC, RNF145, NETO2, ADAT2, CHST2, CTLA4, NFE2L3, LIMA1, IL1R2, ICOS, HSDL2, HTATIP2, FKBP1A, TIGIT, CCR8, LTA, SLC35F2, IL21R, AHCYL1, SOCS2, ETV7, BCL2L1, RRAGB, ACSL4, CHRNA6, BATF, LAX1, ADPRH, TNFRSF4, ANKRD10, CD274, CASP1, LY75, NPTN, SSTR3, GRSF1, CSF2RB, TMEM184C, NDFIP2, ZBTB38, ERI1, TRAF3, NAB1, HS3ST3B1, LAYN, JAK1, VDR, LEPROT, GCNT1, PTPRJ, IKZF2, CSF1, ENTPD1, TNFRSF18, METTL7A, KSR1, SSH1, CADM1, IL1R1, ACP5, CHST7, THADA, CD177, NFAT5, ZNF282, MAGEH1
TYPE I IFN RESPONSE	DDX4, IFIT1, IFIT2, IFIT3, IRF7, ISG20, MX1, MX2, RSAD2, TNFSF10
TYPE II IFN RESPONSE	GPR146, SELP, AHR

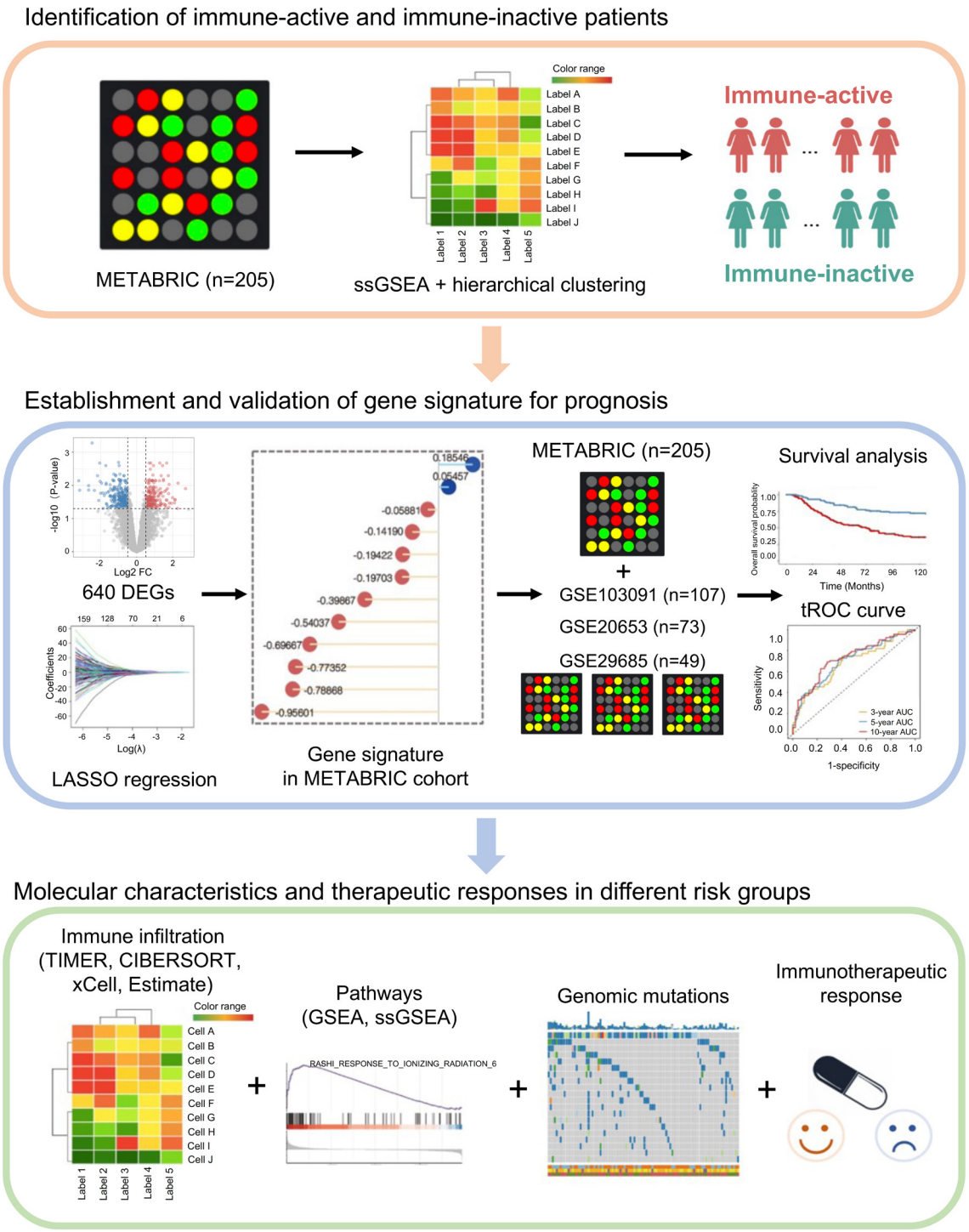


Figure S1. Workflow chart of the study. A. The METABRIC cohort was divided into immune-active and immune-inactive subgroups by ssGSEA and hierarchical clustering based on the transcriptome data of patients; B. A prognostic gene signature was constructed by integrated methods and validated in three external cohorts; C. Differences in microenvironment components, molecular characteristics and response to therapy between the low- and high-IRR groups were evaluated.

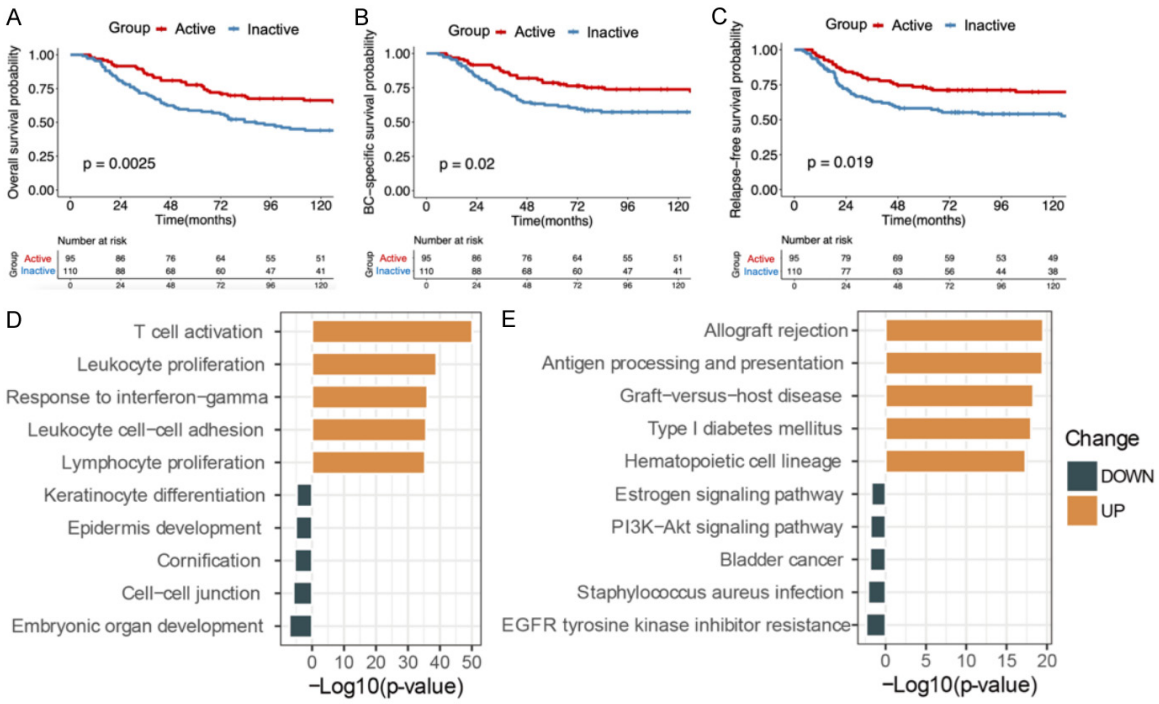


Figure S2. Survival differences in the immune-active and immune-inactive groups. (A-C) Comparison of OS (A), BCSS (B) and RFS (C) between the immune-active and immune-inactive groups by Kaplan-Meier analysis; (D, E) GO and KEGG enrichment analysis of 640 DEGs. The top 5 upregulated and downregulated functions ranked by *p* value are displayed.

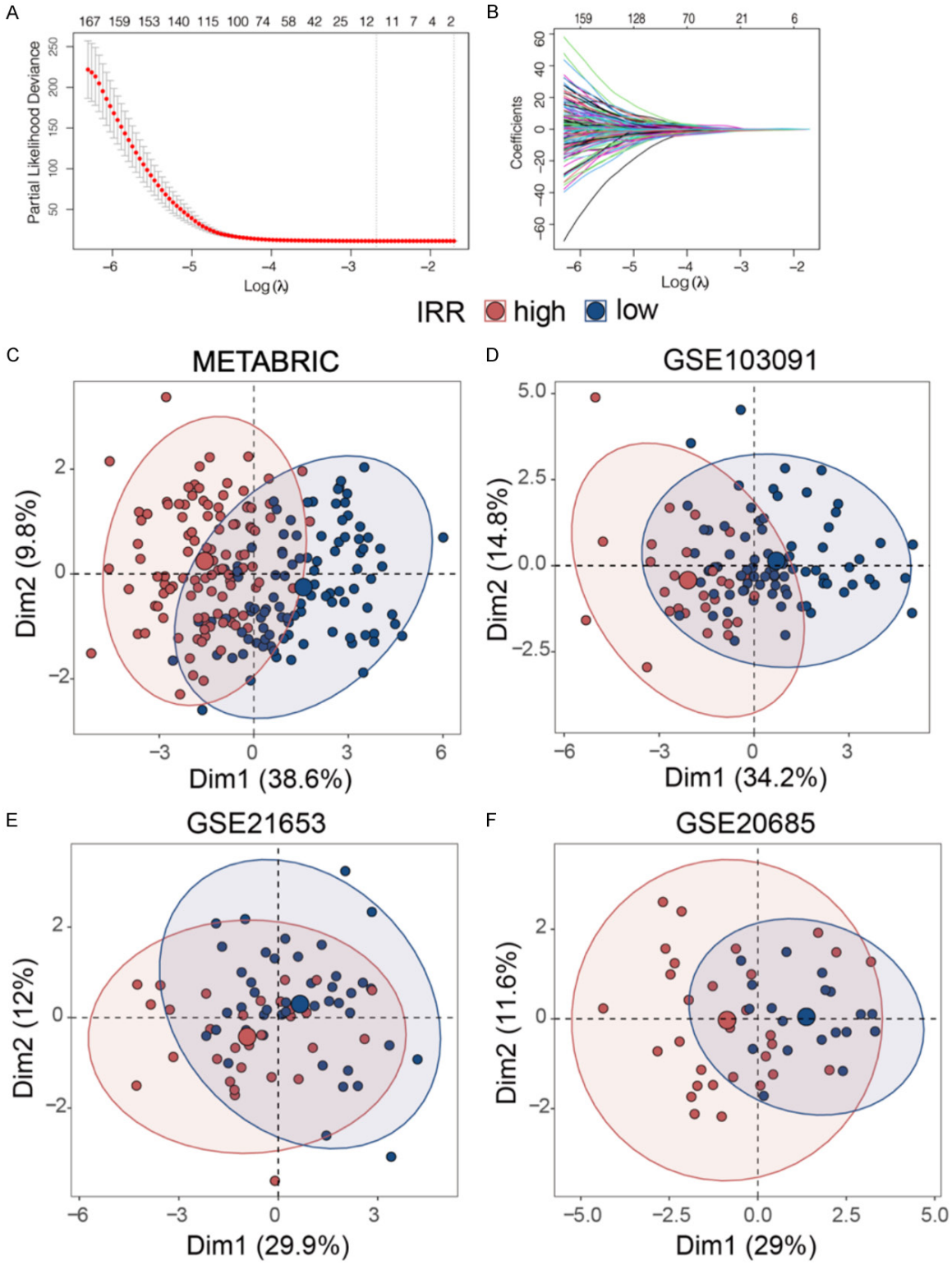


Figure S3. Construction for immune-related risk gene signature. (A) LASSO Cox regression analysis for selecting robust prognostic genes. The black solid vertical lines represent the partial likelihood deviance ± standard error of each lambda. The left vertical dotted line was drawn at the minimum criteria, and the right line was drawn at the 1-SE criteria; (B) LASSO coefficient profiles of the 243 prognosis-related genes. (C-F) Principal component analysis of the high- and low-IRR groups in the METABRIC (C), GSE103091 (D), GSE21653 (E) and GSE20685 (F) datasets.

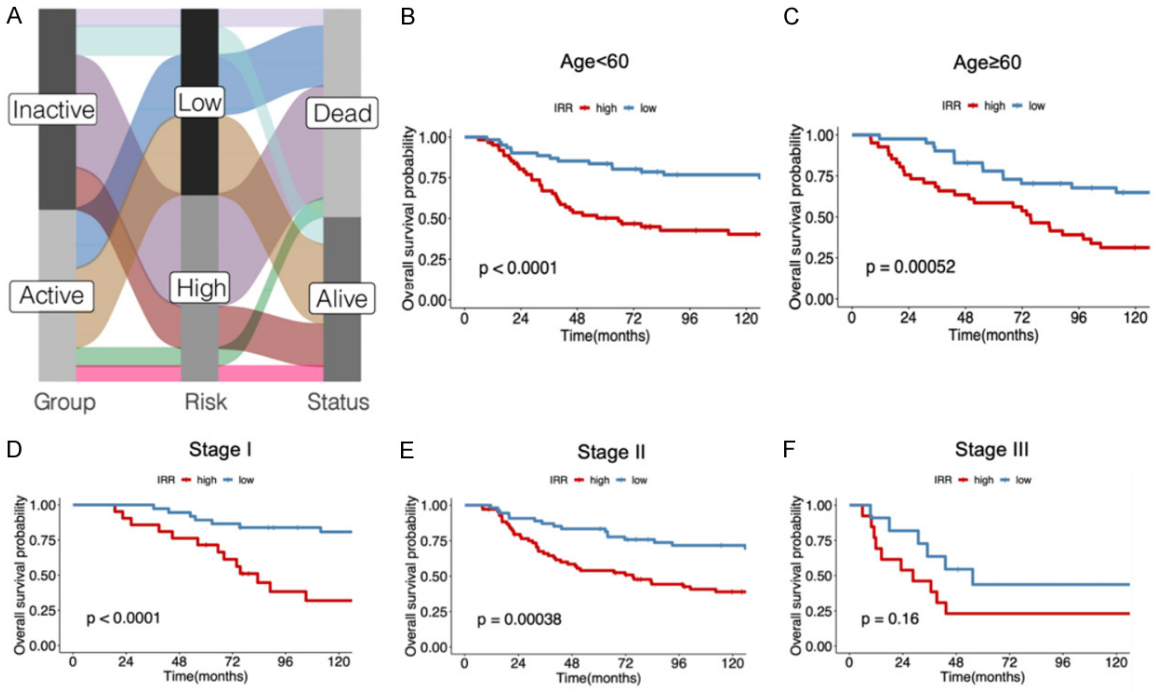


Figure S4. Prognostic value in different clinical subgroups of the METABRIC dataset. A. Sankey diagram illustrating the flows from immune infiltration to IRR group and vital status in the METABRIC cohort; B-F. Kaplan-Meier survival curves of OS in age and stage subgroups stratified by IRR score.

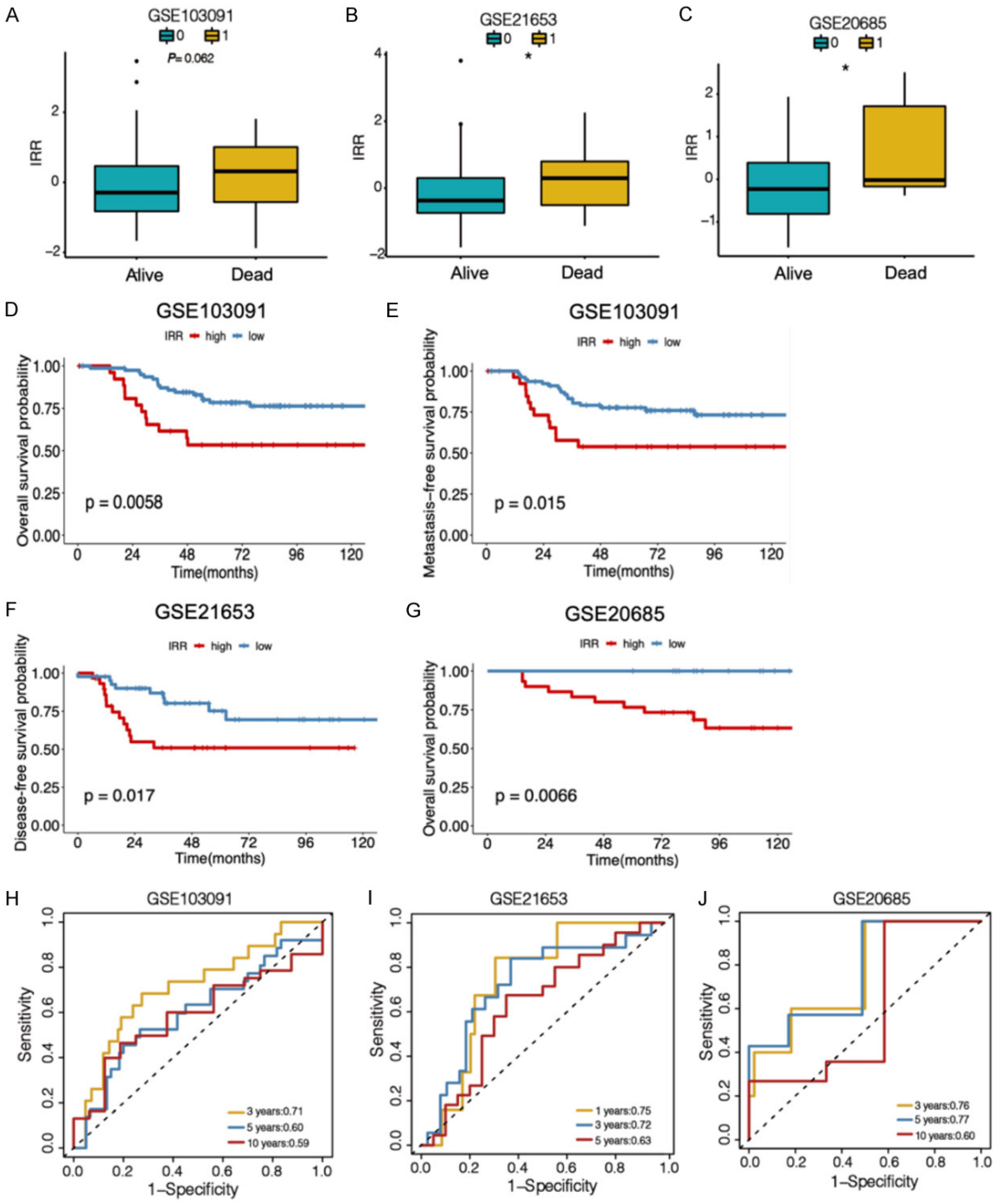


Figure S5. Validation of IRR model performance on risk stratification in three GEO datasets. A-C. Distribution of IRR scores for patients with different clinical outcomes in three GEO cohorts; D-G. Kaplan-Meier survival curves for the high- and low-IRR groups in three GEO cohorts; H-J. ROC curves of IRR scores with regard to 3-, 5- and 10-year survival in three GEO cohorts.

Table S2. Univariate and multivariate Cox regression analyses of clinical variables in the METABRIC and GEO datasets

METABRIC-OS	Univariate		Multivariate	
	HR	*p*-value	HR	*p*-value
Age (vs. < 60)				
≥60	1.48 (1.02-2.14)	0.037	1.47 (1.02-2.14)	0.041
Grade (vs. Grade 1/2)				
Grade 3	1.24 (0.71-2.18)	0.45		
Stage (vs. stage I)				
Stage II	1.41 (0.90-2.21)	0.134	1.23 (0.78-1.94)	0.379
Stage III	2.79 (1.53-5.09)	0.001	2.47 (1.35-4.53)	0.003
IRR (vs. low)				
high	2.85 (1.92-4.23)	< 0.0001	2.74 (1.84-4.08)	< 0.0001
Chemotherapy (vs. NO)				
YES	1.09 (0.75-1.58)	0.663		
Radiotherapy (vs. NO)				
YES	0.86 (0.56-1.30)	0.467		
GSE103091-OS				
Age (vs. < 60)				
≥60	3.19 (1.50-6.76)	0.003	3.23 (1.52-6.85)	0.002
IRR (vs. low)				
high	2.72 (1.30-5.69)	0.008	2.76 (1.32-5.79)	0.007
GSE21653-DFS				
Age (vs. < 60)				
≥60	1.86 (0.80-4.31)	0.148		
Grade (vs. Grade 1/2)				
Grade 3	0.50 (0.19-1.37)	0.178		
T (vs. T1)				
T2	0.53 (0.16-1.73)	0.295		
T3	1.18 (0.36-3.84)	0.784		
N (vs. N0)				
N1	2.43 (1.05-5.64)	0.039	3.08 (1.30-7.27)	0.01
IRR (vs. low)				
high	2.72 (1.16-6.39)	0.022	3.38 (1.41-8.07)	0.006
GSE20685-OS				
Age (vs. < 60)				
≥60	1.01 (0.95-1.08)	0.738		
Stage (vs. stage I)				
Stage II	2.34 (0.29-19)	0.427	2.27 (0.28-18.5)	0.445
Stage III/IV	17.19 (1.55-190.72)	0.021	22.04 (1.90-255.84)	0.013
IRR (vs. low)				
high	2.05 (1.15-3.64)	0.015	2.30 (1.22-4.34)	0.01

Abbreviation: HR, Hazard ratio; IRR, immune-related risk; OS, overall survival; DFS, disease-free survival.

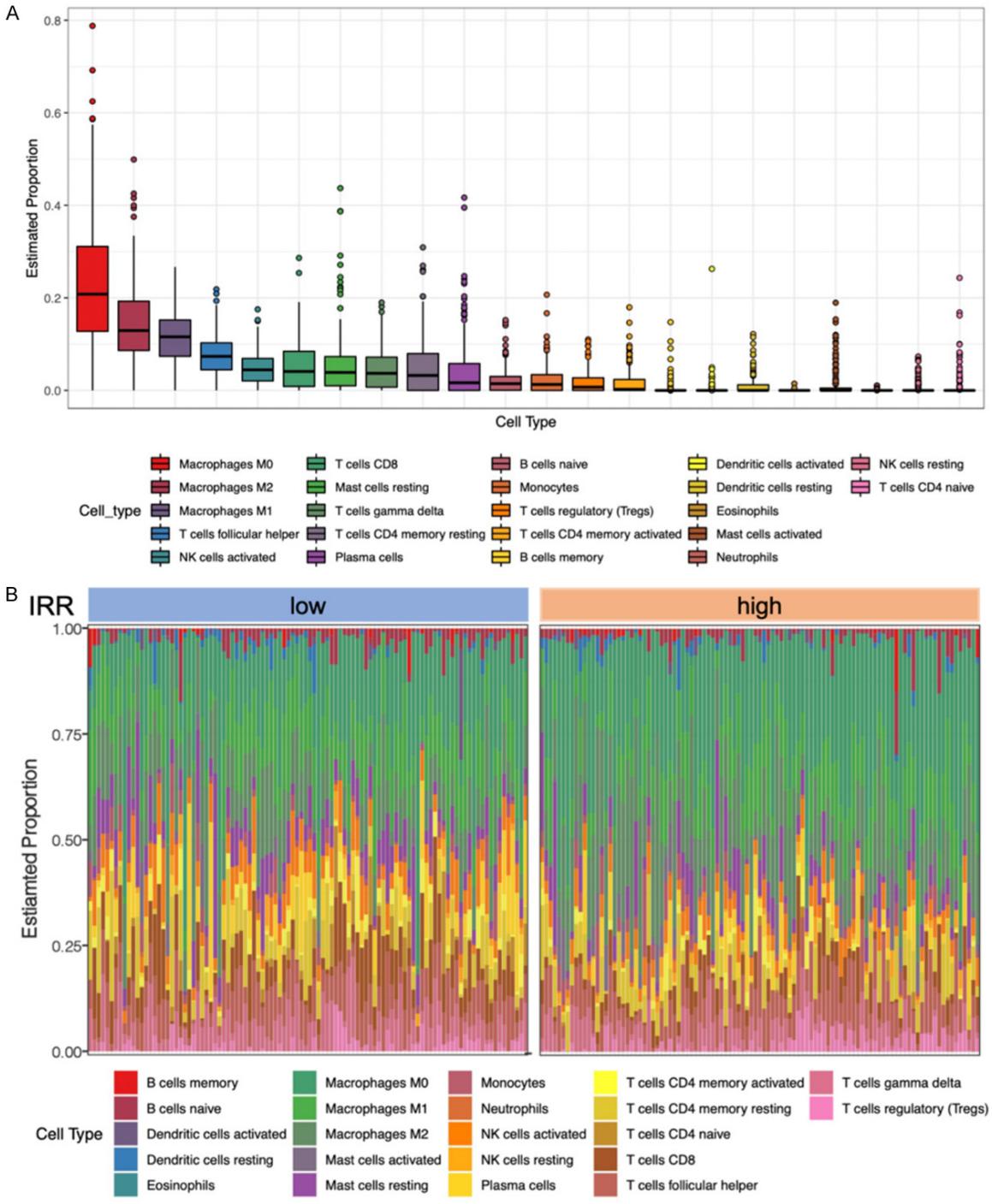


Figure S6. Distribution of 22 immune cell types in the METABRIC cohort. A. Boxplot showing the distribution tendencies of 22 immune cell types in the METABRIC cohort; B. Stacked bar plots illustrating the infiltration patterns of 22 immune cell types in the high- and low-IRR groups.

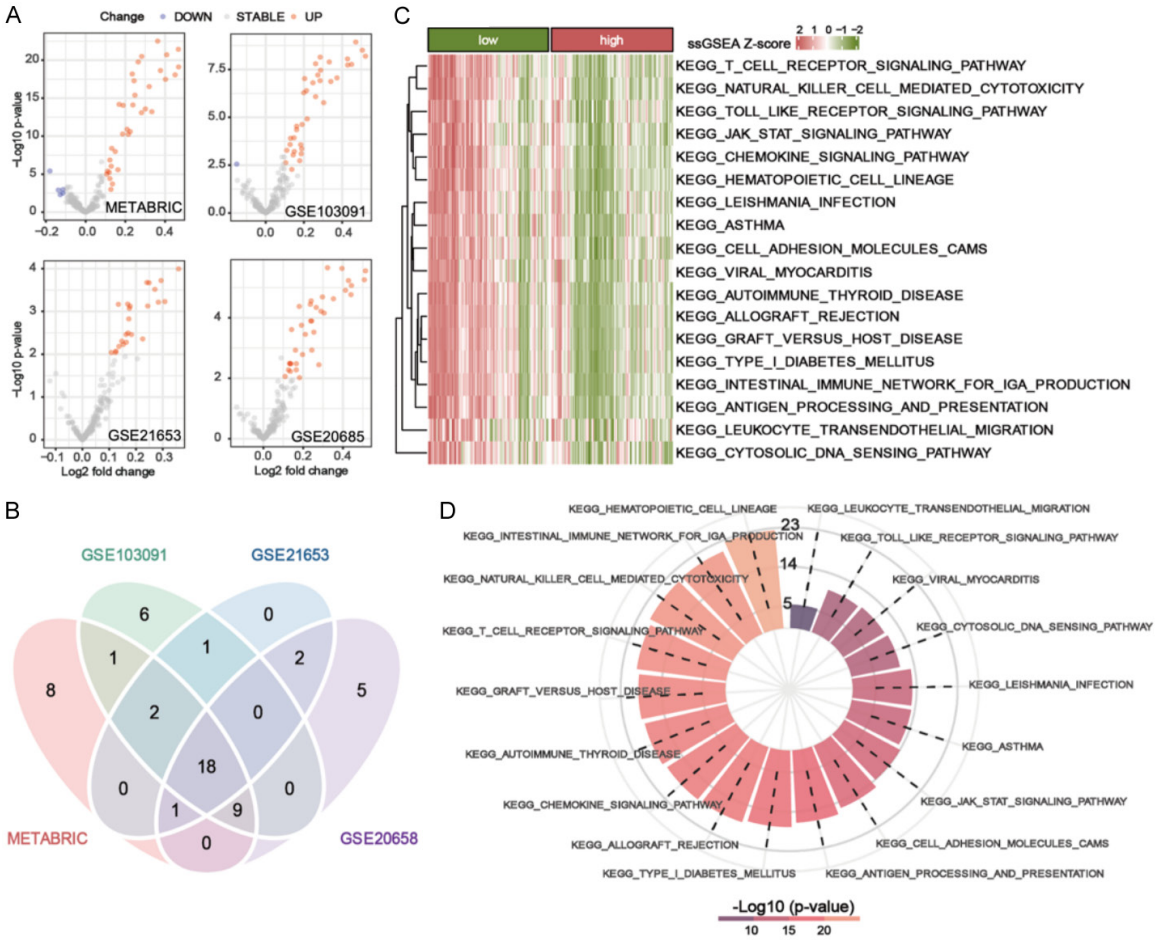


Figure S7. GSVA based on KEGG gene sets in the METABRIC and three GEO cohorts. A. Volcano plots showing enriched KEGG pathways identified by GSVA in the high- and low-IRR groups (|Log2FC| > 0.1; *p* value < 0.01); B. Venn plot showing 18 commonly upregulated KEGG pathways in the high-IRR group of four cohorts; C. Heatmap showing the enrichment scores of 18 enriched KEGG pathways in the METABRIC cohort; D. Statistic significance of 18 enriched KEGG pathways in the high-IRR group versus the low-IRR group of the METABRIC cohort.

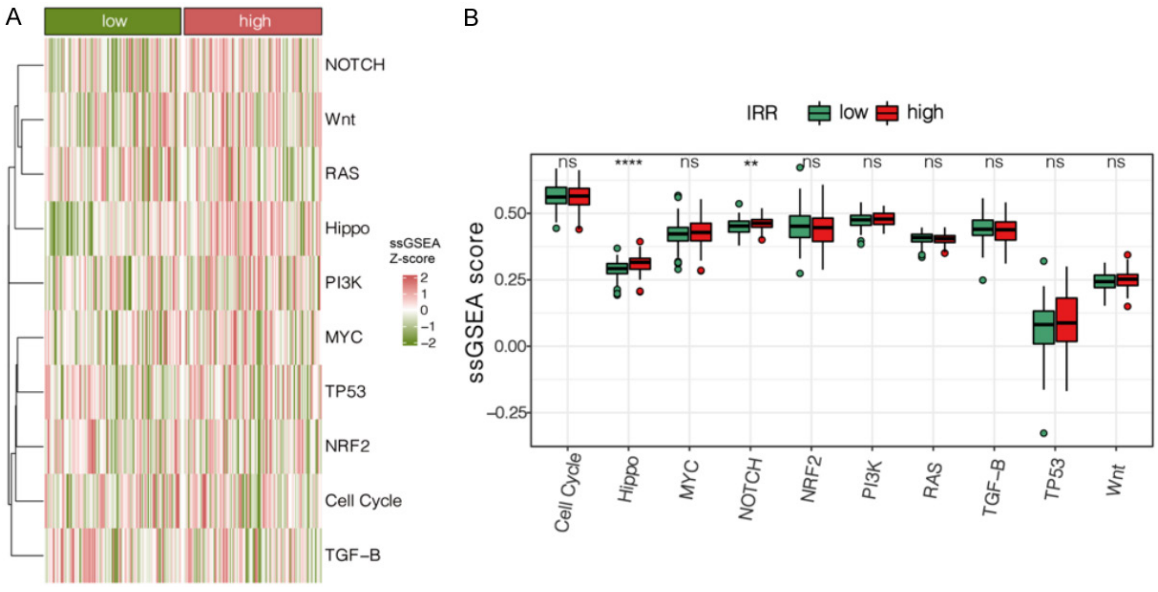


Figure S8. A. Heatmap displaying the normalized ssGSEA scores of 10 oncogenic pathways in the high- and low-IRR groups of the METABRIC cohort; B. Boxplot showing the enriched scores of 10 oncogenic pathways in the high- and low-IRR groups of the METABRIC cohort.

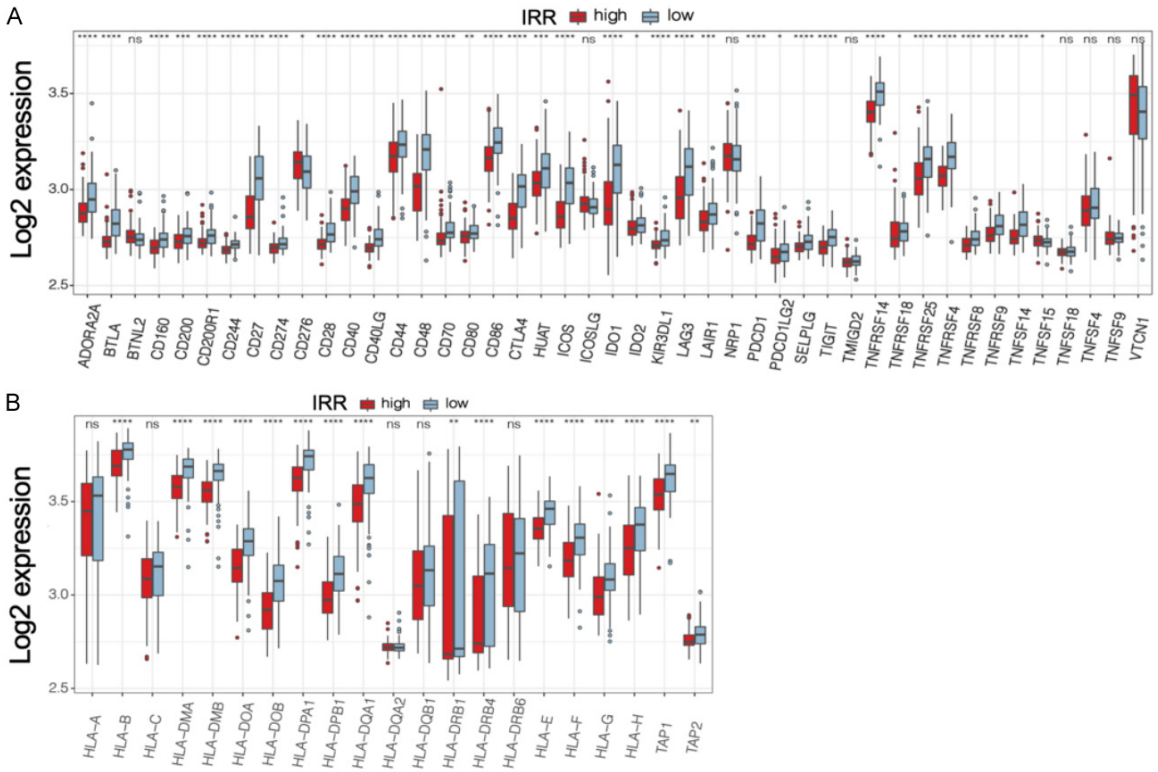


Figure S9. (A, B) Comparison of the expression levels of immune checkpoints (A) and HLA family genes (B) in the high- and low-IRR groups of the METABRIC cohort.

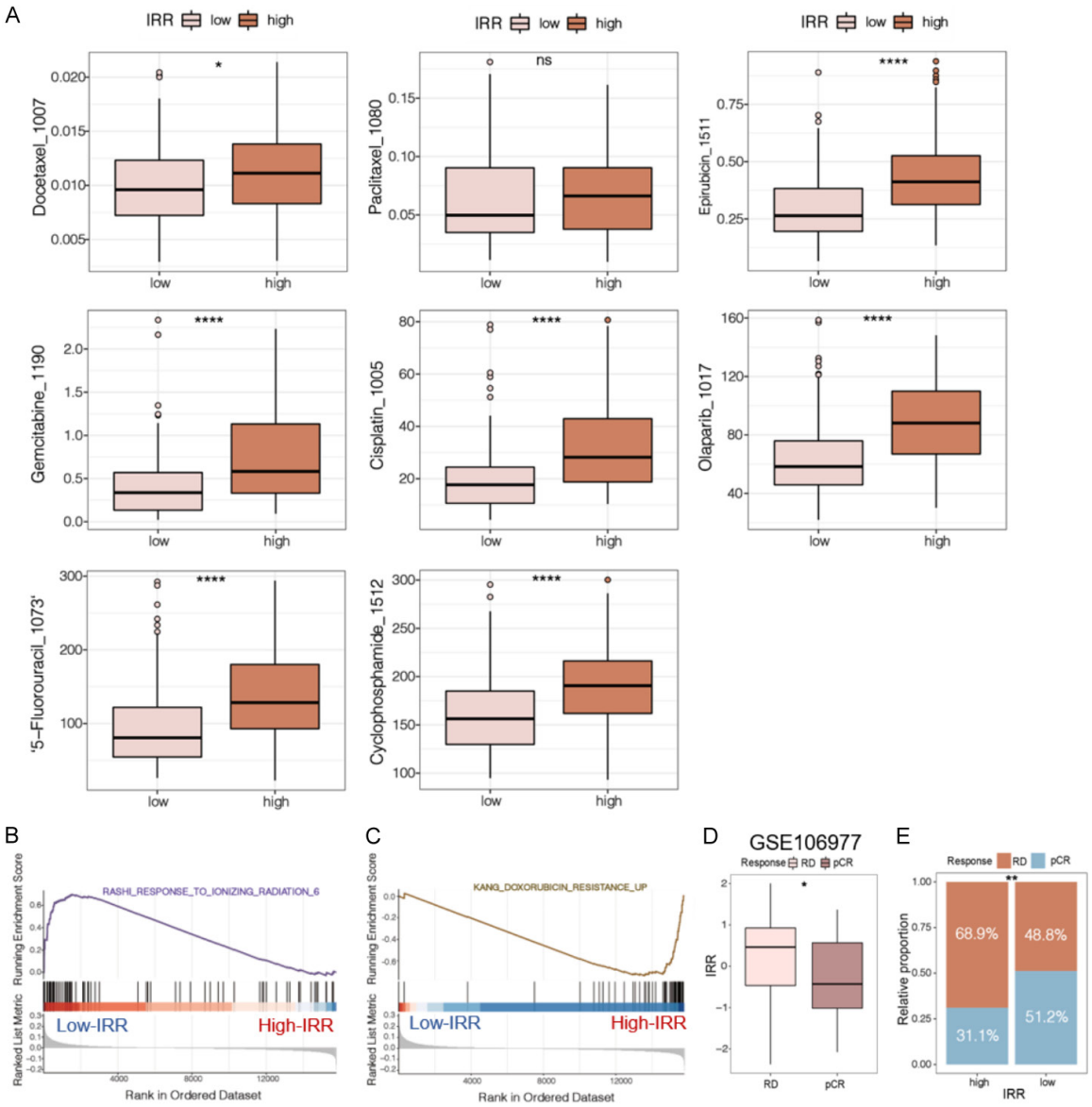


Figure S10. High-IRR patients showed more resistance to chemotherapy. A. Estimated chemotherapy responses of the high- and low-IRR patients for eight common chemotherapy drugs in TNBC treatment. B. GSEA showed that response to the ionizing radiation pathway was enriched in the low-IRR group. C. GSEA showed that doxorubicin resistance pathway was enriched in the high-IRR group. D. Boxplot illustrating the distribution of IRR scores for different neoadjuvant chemotherapy responses of 88 patients treated with neoadjuvant anthracyclines and/or taxanes in the GSE106977 cohort. E. Bar graph showing the proportions of different responses to neoadjuvant anthracyclines and/or taxanes treatment in the high- and low-IRR groups of the GSE106977 cohort.

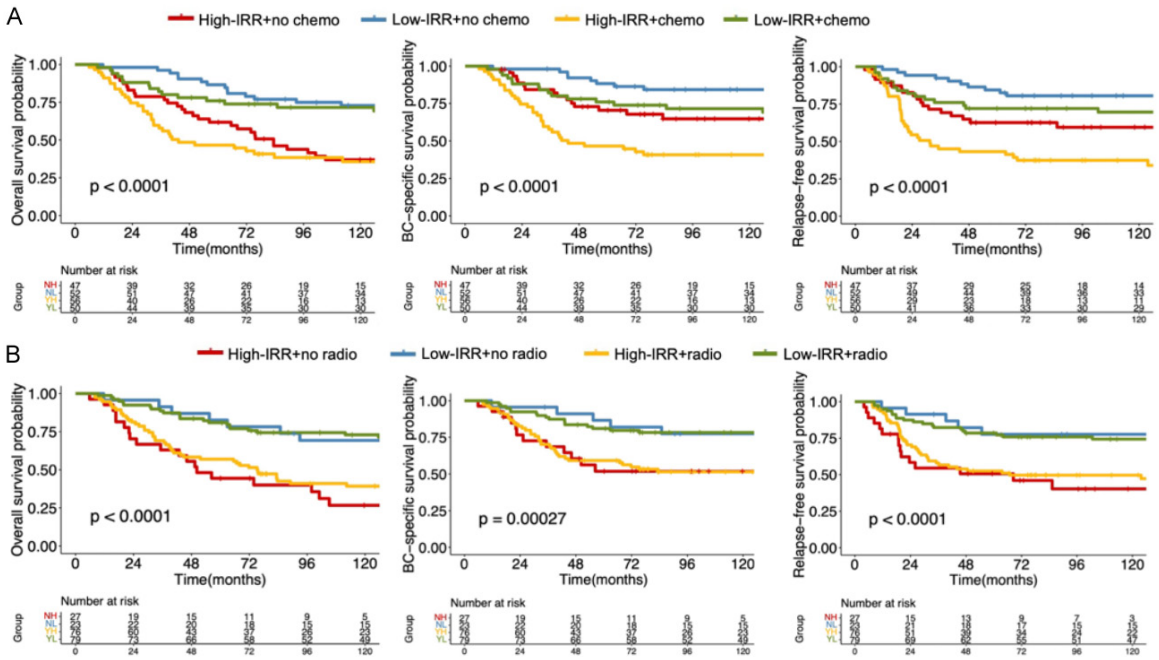


Figure S11. Survival analyses for patients stratified by combining the IRR score and adjuvant chemotherapy or radiotherapy in the METABRIC cohort.

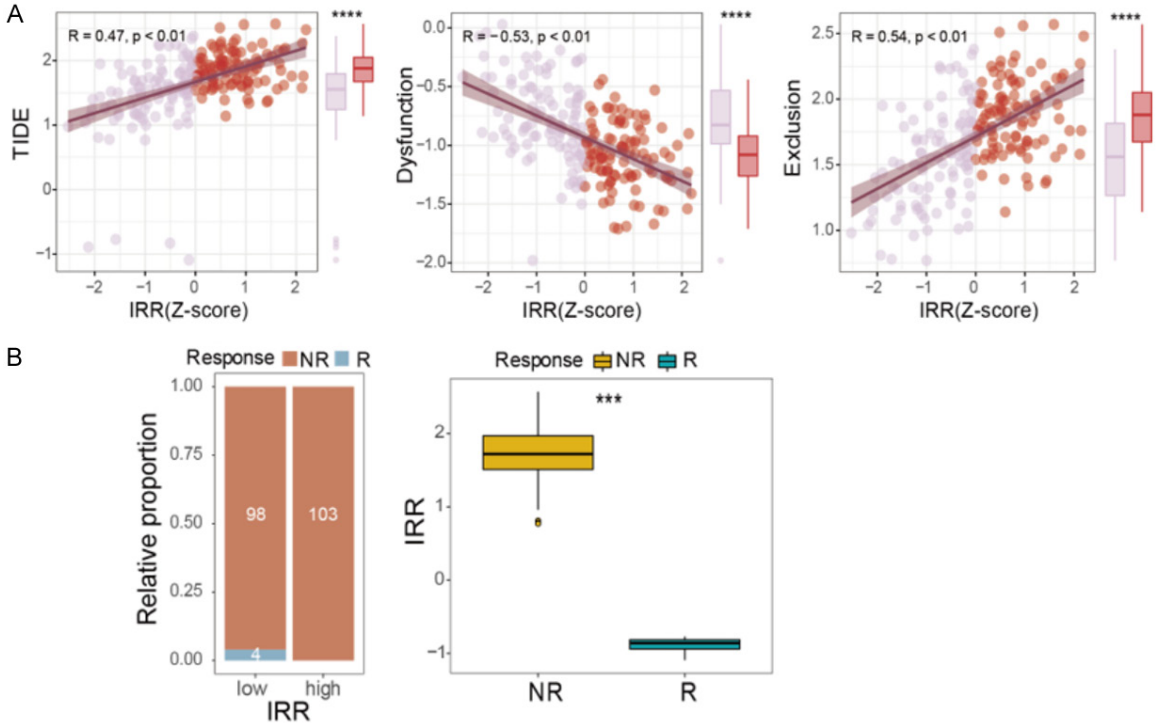


Figure S12. Prediction of immunotherapy response based on the TIDE algorithm. A. Correlation analyses for TIDE, Dysfunction, and Exclusion scores with the IRR score. B. Bar graph showing the numbers of different clinical responses to immunotherapy estimated by the TIDE algorithm in the high- and low-IRR groups for the METABRIC cohort. Boxplot illustrating the distribution of IRR scores of patients with different immunotherapeutic responses estimated by the TIDE algorithm in the METABRIC cohort. NR, non-responder; R, responder.