



OPEN

Comprehensive evaluation of breast cancer immunotherapy and tumor microenvironment characterization based on interleukin genes-related risk model

Yalei Lv¹, Zihe Bai¹, Xiaoyan Wang², Jiayin Liu¹, Yuntao Li³, Xiaolin Zhang⁴ & Yujie Shan^{1✉}

Breast cancer (BRCA) is the most prevalent malignancy and the leading cause of death in women. Interleukin (IL) genes are critical in tumor initiation and control. Nevertheless, the prognosis value of the IL in BRCA remains unclear. We collected data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), and 94 IL genes were identified from GeneCard. Based on the random forest (RF), least absolute shrinkage and selection operator (LASSO) analysis, and multivariate Cox regression analysis, we constructed an IL signature. GSE22219, GSE25065, and GSE21653 were derived as validation sets. The expression differences in the tumor microenvironment (TME), immunotherapy, and chemosensitivity of BRCA between the high- and low-risk groups were evaluated. Overall, 21 IL genes were selected to construct an IL risk model, of which IL18BP, IL17D, and IL23A were the first time identified as prognostic genes in BRCA. IL score could distinguish BRCA patients with inferior outcomes, and AUC of it was 0.70, 0.76, and 0.72 for 1-, 3- and 5- years, respectively, which was also verified in GSE22219, GSE25065, and GSE21653 cohorts. Meanwhile, compared to luminal A and luminal B, HER2-positive and TNBC had significantly higher IL score. Besides, the high-risk group had a significantly higher prevalence of *TP53* and *TTN* but a lower prevalence of *PIK3CA*, as well as higher tumor mutation burden (TMB) and neoantigen level. High- and low-risk groups exhibited notable differences in immunomodulators and tumor infiltrates immune cells (TIICs), and the high-risk group had significantly lower Tumor Immune Dysfunction and Exclusion (TIDE) score. Additionally, the high-risk group has more responders to immune or anti-HER2 combination therapy, whereas the low-risk group has higher sensitivity to docetaxel and paclitaxel. Consequently, we constructed a reliable risk model based on the IL genes, which can provide more information on both the risk stratification and personalizing management strategies for BRCA.

Breast cancer (BRCA) is the most frequently diagnosed malignancy and the fifth greatest cause of cancer-related death in women worldwide. Despite revolutionary advances have been made in the early-detection, there are still approximately 2.3 million new cases diagnosed with it in 2020, accounting for 11.7% of all cancer cases worldwide¹. Meantime, BRCA is a widely known heterogeneous disease that can be classified into different histopathological subtypes according to the expression of hormone receptors (estrogen receptor (ER) and progesterone receptor (PR)) and epidermal growth factor receptor 2 (HER2/Neu)^{2,3}. Then, different treatment methods, such as surgery, radiation therapy, chemotherapy, and endocrine would be adopted according to the patient's histological subtype⁴. However, primary and or developed subclones promoted by therapeutic drugs would further result in drug resistance and tumor recurrence⁵. In addition, primary defined histological subtypes,

¹Department of Medical Oncology, Fourth Hospital of Hebei Medical University, 12 Jiankang Road, Shijiazhuang 050011, China. ²The Fifth Ward of Medical Oncology, Shijiazhuang People's Hospital, Shijiazhuang, China. ³Breast Center, Fourth Hospital of Hebei Medical University, Shijiazhuang, China. ⁴Department of Epidemiology and Statistics, Hebei Medical University, Shijiazhuang, China. ✉email: hdshanyujie@163.com

such as triple-negative breast cancer (TNBC), are also molecularly heterogeneous and can be further categorized into subtypes with varying prognoses⁶. Mounting evidence demonstrates the emerging role of immunotherapy in the BRCA treatment, and emphasizes the association between tumor microenvironment and BRCA metastasis^{7,8}. Lv et al.⁹ generated an aging genes-related risk stratification that can be used for predicting immunotherapy in BRCA. Zhou et al.¹⁰ developed a prognostic model of cellular senescence-related genes, which was applied to characterize the tumor microenvironment infiltration in BRCA. Nevertheless, the existing risk models still have an inadequacy in the prognostic role, thus, it is of great significance to construct a novel signature for improving the prognosis of BRCA.

Interleukins (IL) and associated cytokines regulate the innate and adaptive immunity in normal and tumor tissue, with roles in immunomodulatory, promoting signal transduction, and maintaining tissue homeostasis^{11,12}. Previous studies have indicated that IL servers as a “double agent” in tumor development: on one hand, chronic inflammation has been recognized as a driver of carcinogenesis which produces carcinogenic mediators, and IL further promotes tumor growth, metastasis and progression¹³; on the other hand, IL also governs the innate and adaptive immunity-mediated cancer cell death, starting from the lymphocytes’ proliferation to the termination^{13,14}. Specifically, numerous studies have delineated the role of IL in the development of BRCA, and most of them have long been implicated as the promoters contributed to tumor invasion, migration and therapy resistance^{15–17}. For instance, IL6 not only interacts with STAT3 signaling to drive ER positive BRCA metastasis and lead to resistance to hormone therapy¹⁸, but also alters tumor microenvironment (TME) in TNBC promoting epithelial-mesenchymal transition (EMT) progression, cancer stemness and M2 macrophage polarization¹⁹. On the contrary, some IL family members have also been proven with antitumor activity in BRCA^{20,21}. Over-expression of IL-1 β has been revealed the association with better outcomes in BRCA patients with lymph node metastasis²². Moreover, as IL directly participates in or mediates the regulation of the TME, it could be targeted to increase the sensitivity of immune checkpoint inhibitors²³. Prior research mainly has only established a correlation between a single IL gene and BRCA prognosis^{24–26}; hence, the comprehensive prognostic function of IL family genes in breast cancer, to our best knowledge, remains unknown.

In this research, the underlying role of IL in prognosis, gene mutation, TME, and immunotherapy of BRCA were our primary concerns. Based on the IL genes, we have developed an IL score that could predict the OS of BRCA patients in the TCGA cohort and further validated this risk model in three GEO cohorts. In addition, the correlation of mutation characteristics and risk score were comprehensively applied, and we further evaluated the TME landscape between high- and low-risk groups. Moreover, the response to immunotherapy in BRCA patients with different IL scores was performed to contribute reliable insights into the treatment for BRCA. Taken together, our study might provide robust biological targets for improving prognosis and clinical treatment for BRCA patients, and offer further reliable guidance on the personalized medicine of BRCA.

Materials and methods

Data collection and analysis. The RNA expression data and clinical details of the TCGA-BRCA cohort were downloaded from The Cancer Genome Atlas (TCGA) database (<http://xena.ucsc.edu/>) as the training set, with male BRCA patients excluded. Meanwhile, the GSE22219, GSE25065, and GSE21653 were collected from Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>) and were used as three independent validation sets. Meantime, we searched “interleukin” as a keyword from the GeneCard (<https://www.genecards.org/>) with the category protein-coding and high correlation score, and finally a total of 94 IL genes (Table S1) were identified and utilized for the subsequent analysis.

Construction and validation of the IL score. Firstly, we performed random forest (RF) with R package “randomSurvivalForest” (version 3.6.4)²⁷ to obtain hub IL genes (with variable relative importance ≥ 0.4). Secondly, the least absolute shrinkage and selection operator (LASSO) regression analysis²⁸ was applied to screen out the optimal genes through the “glmnet” R package²⁹, and the best λ value was obtained by tenfold cross-validation used for gene screening. We took the intersection genes of hub genes and optimal genes from RF and LASSO, respectively, and applied multivariate Cox regression analysis to construct a risk model based on them. The risk model formula was exhibited as follows: IL score = $\sum_{j=1}^n Expr_{genej} * Coef_{genej}$, and the BRCA patients were divided into a high-risk group and a low-risk group according to the median IL score. To further evaluate the prognostic ability of the IL score, we used the Kaplan–Meier survival curve to evaluate the overall survival (OS) ability based on the “survival” and “survminer” packages³⁰, and performed receiver operating characteristic (ROC) at 1, 3 and 5 years to calculate the area under the curve (AUC) values using the “timeROC” R package.

Clinical features analysis and nomogram establishment. The Kruskal Wallis test was used to analyze the difference in clinical characteristics between two risk groups, including age, tumor node metastasis (TNM) and pathological stage, and the stratified analysis was then performed in the TCGA cohort. According to the results of univariate and multivariate Cox regression analysis, we generated a nomogram based on IL score and clinicopathological parameters using “rms” R package, and the performance of the nomogram was assessed by calibration curves and ROC curves.

Molecular subtypes of BRCA. BRCA was mainly divided into four molecular subtypes, including luminal A, luminal B, HER2-positive, and triple-negative BRCA (TNBC)³¹. Then, we performed an alluvial diagram to visualize the changes in IL score in four molecular subtypes of BRCA and applied the Kaplan–Meier survival curves to analyze the prediction ability in the IL score of four molecular subtypes with “survival” and “sur-

vminer” packages. In addition, we compared the differences in IL score, tumor mutation burden (TMB), and neoantigen level among four different molecular subtypes of BRCA.

Functional enrichment analysis. To deeply explore the underlying biological activities of the risk model, we firstly screen the genes that were strongly associated with IL score using the Pearson ($|R| > 0.25$, $p < 0.05$), and we finally obtained IL score -associated genes. Kyoto Encyclopedia of Genes and Genomes (KEGG)³² and Gene Ontology (GO)³³ functional annotation was applied to analyze the functional enrichment of IL score-associated genes using the “cluster Profile” R package. The GO mainly performed enrichment analysis in three aspects, including biological process (BP), cellular component (CC), and molecular function (MF). p -value < 0.05 and $q < 0.05$ were set as cutoff values.

Mutation analysis. To evaluate the mutational landscape of BRCA patients between high- and low-risk groups, we analyzed the mutation information of BRCA patients using “maftools” package. Subsequently, we used cBioPortal (https://www.cbioportal.org/mutation_mapper) to draw lollipop plots that could identify the variation distribution of genes and the underlying functional impact of mutations. Finally, the TMB of the high- and low-risk groups was also assessed.

Tumor microenvironment of BRCA. The tumor microenvironment (TME) included non-cancerous host cells in the tumor and non-cellular components, which played an important role in the progression and therapeutic effect of cancer³⁴. To identify the immunological features of the TME in BRCA, we assessed the expression level of immunomodulators between high- and low-risk groups. In addition, we performed CIBERSORT³⁵ to analyze tumor infiltrates immune cells (TIICs) abundance in high- and low-groups.

The immunotherapy landscape. First of all, we evaluated the different expression levels between 24 immune checkpoints and two risk groups. Moreover, we calculated the enrichment scores of immunotherapy prediction pathway between high- and low-risk groups in the TCGA cohort, and the ESTIMATE algorithm was applied to calculate the immune score, stromal score, and tumor purity. Tumor Immune Dysfunction and Exclusion (TIDE) score, T cell dysfunction score, T cell exclusion score, and microsatellite instability (MSI) score were evaluated based on the TIDE website (<http://tide.dfci.harvard.edu/>). In addition, the dataset GSE173839 was applied to evaluate the immunotherapy benefit of the risk model. Finally, we downloaded the immunophenoscore (IPS) of BRCA from The Cancer Immunome Atlas (TCIA) (<https://tcia.at/home>) to predict the BRCA patients’ response to immunotherapy.

The correlation between chemosensitivity and IL score. We collected three treatment targets, including EGFR network, immune inhibited oncogenic pathways, and radiotherapy predicted pathways to assess the correlation between treatment targets and the risk model. In addition, the drugbank database (<https://go.drugbank.com/>) was used to obtain the BRCA-related drug target genes. Furthermore, we used the “ggplot2” R package to compare the proportion of pathological complete response (pCR) among patients in the high- and low-risk groups who received neoadjuvant chemotherapy in the GSE194040 cohort.

Statistical analysis. We used R software and its associated software package (v.4.1.2) to analyze the data. Kaplan–Meier curves and log-rank tests were used to assess differences in OS between groups. Continuous data processing was performed by Student’s t -test and Wilcoxon’s test, and Fisher’s exact test was applied to categorical data. All tests were two-way and $p < 0.05$ was considered statistically significant.

Results

Development of the IL score. First, we used RF to screen the IL genes and found 39 hub IL genes with relative significance values greater than 0.4 (Fig. 1A). In the meantime, LASSO regression analysis further determined 42 optimal IL genes (Fig. 1B) (Table S2). Then, 21 intersection genes were overlapped by these two algorithms, including IL6ST, IL18BP, NFIL3, IRAK4, IL1RL1, IL21, IL17D, IL23A, IL22RA2, IL26, IL1RAPL2, IRAK1, IL17RE, IL27, ILF2, IL1RN, IL31RA, IL20RB, IL1RL2, IL13RA1, and IL17B (Table S3). Based on the 21 IL genes, the multivariate Cox regression analysis was employed, aiming to construct a risk model (Fig. 1C). The IL score was calculated by multiplying the expression of each gene by its respective coefficient and adding the results together (Table S4). The BRCA patients in the training cohort were divided into high- and low-risk groups according to the median IL score. Figure S1 displayed the distribution of risk score and survival status in BRCA patients. Low-risk patients had better OS than those with high-risk (median OS: 9.3 months vs 11.7 months, $p < 0.0001$, Fig. 1D), and the AUC for prediction survival of the IL score was 0.70 in 1 year, 0.76 in 3 years, and 0.72 in 5 years (Fig. 1E).

Validation and comparison of IL score. To further access the robustness of the IL score, we performed three independent cohorts from GEO, including GSE22219, GSE25065, and GSE21653 using the same formula to calculate the risk score (Table S5–S7). Figure S2A–C showed the distribution of risk score and survival status for BRCA patients in three validation sets. As shown in the Fig. 2A–C, high-risk patients in three validation cohorts all had worse OS than low-risk patients. In addition, the AUC values of 1-, 3-, 5 years were 0.66, 0.62 and 0.65 in GSE22219 cohort (Fig. 2D); 0.62, 0.65 and 0.64 in GSE25065 cohort (Fig. 2E); and 0.79, 0.63 and 0.65 in GSE21653 cohort (Fig. 2F). To comprehensively analyze the prognostic ability of IL score, four risk models^{36–39} from previous studies were collected to compare with the prognostic prediction capacity. The result demon-

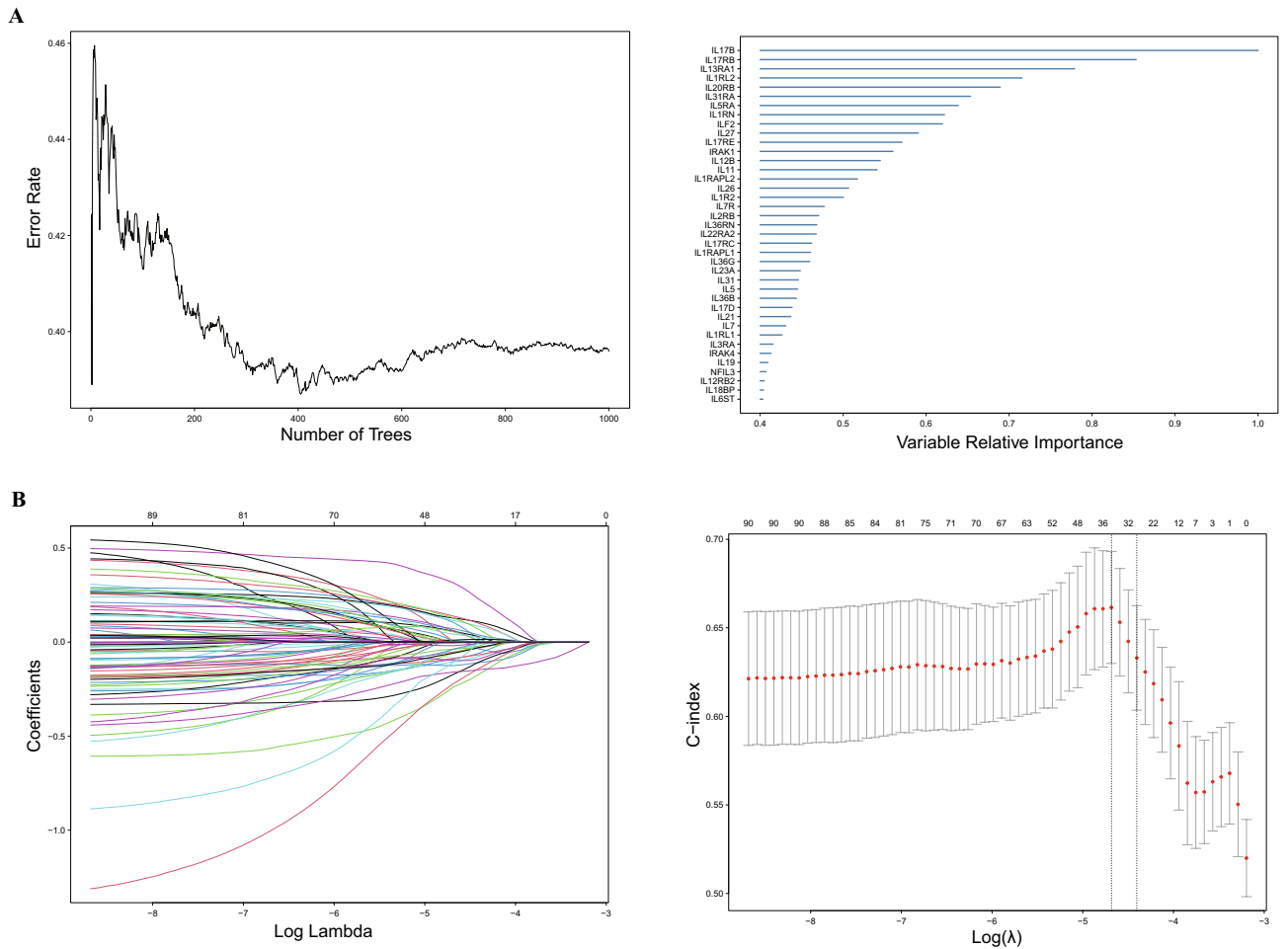


Figure 1. Construction of IL score. **(A)** Screening hub genes by RF with variable relative importance ≥ 0.4 . **(B)** Screening optimal genes by LASSO and calculating the minimum criteria. **(C)** Multivariate Cox regression analysis to obtain intersection genes. **(D)** Kaplan–Meier curve in TCGA cohort. **(E)** Sensitivity and specificity of ROC curve in TCGA cohort. IL, interleukin; RF, random forests; LASSO, Least Absolute Shrinkage and Selection Operator; TCGA, The Cancer Genome Atlas; ROC, receiver operating characteristic; AUC, area under the curve.

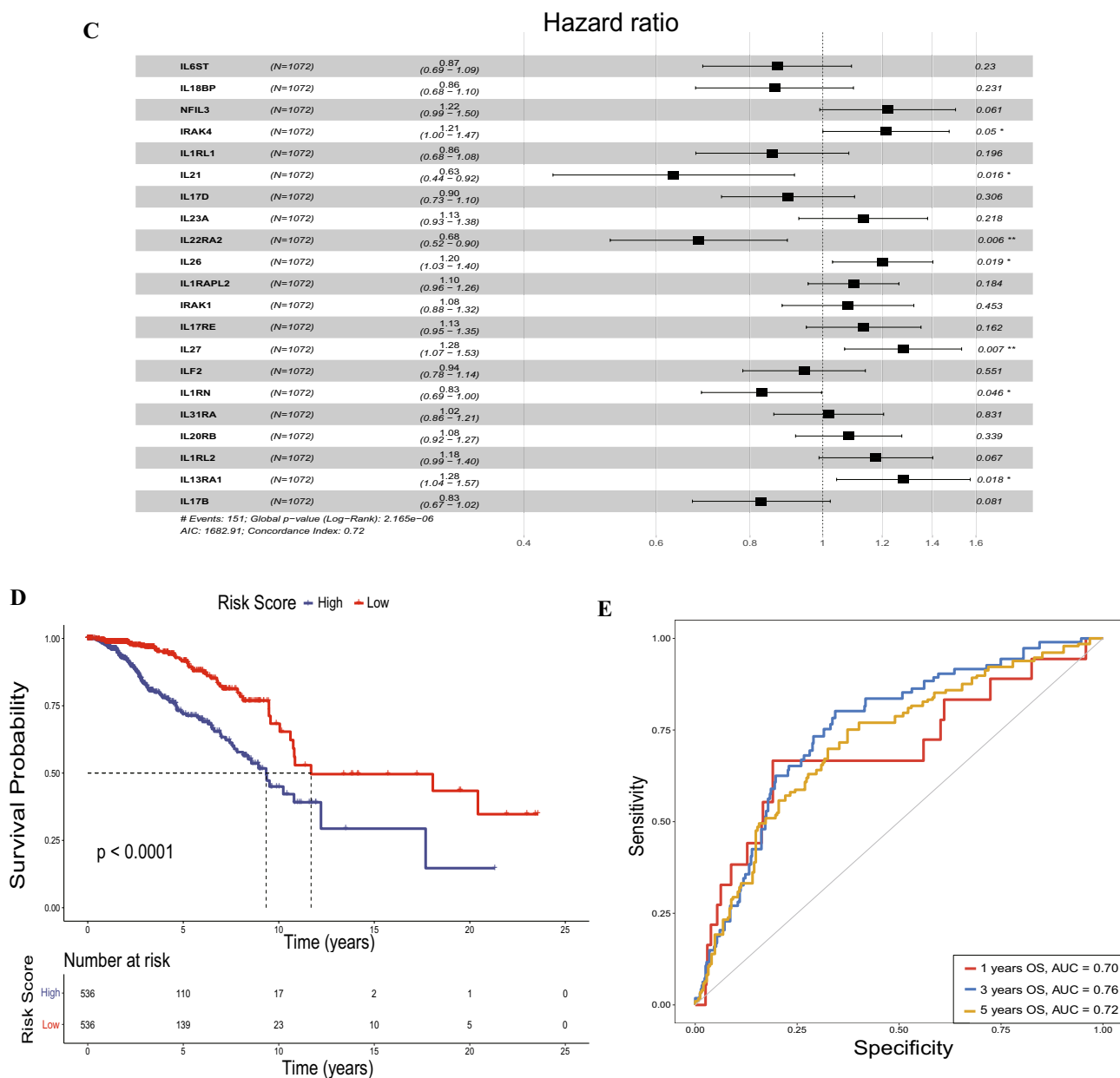


Figure 1. (continued)

stated that the AUC value of the IL score was significantly greater than the other four risk models, showing that our risk model surpassed the others in terms of accuracy (Fig. 2G). In addition, both univariate and multivariate Cox regression analyses demonstrated that the IL score had good prognostic value in the four cohorts (Fig. 2H and I).

The landscape of clinicopathological features in the risk model. Then, we analyzed the difference in IL score across BRCA patients with different clinicopathological features, including age, pathological stage, and TNM stage, and found that patients with M1 or higher stages had a significantly higher risk score (Fig. 3A). In addition, stratification survival analysis was applied according to age (<60 years and ≥ 60 years), pathological stage (stage I–II and stage III–IV), T stage (T1–2 and T3–4), N stage (N0 and N1–3), and M stage (M0 and M1). It was confirmed that the patients with high risk score had significantly shorter OS, regardless of their clinicopathological features (Fig. 3B–K).

Heterogeneity in IL score between different molecular subtypes. The interaction between well known four molecular subtypes (luminal A, luminal B, HER2-positive, and TNBC) and IL score subtypes were visualized using an alluvial diagram (Fig. 4A). Intriguingly, there were more TNBC and HER2-positive patients in the high-risk group, whereas there were more luminal A patients in the low-risk group. Survival analysis revealed that the IL score may distinguish the survival of luminal A patients significantly ($p < 0.0001$). The simi-

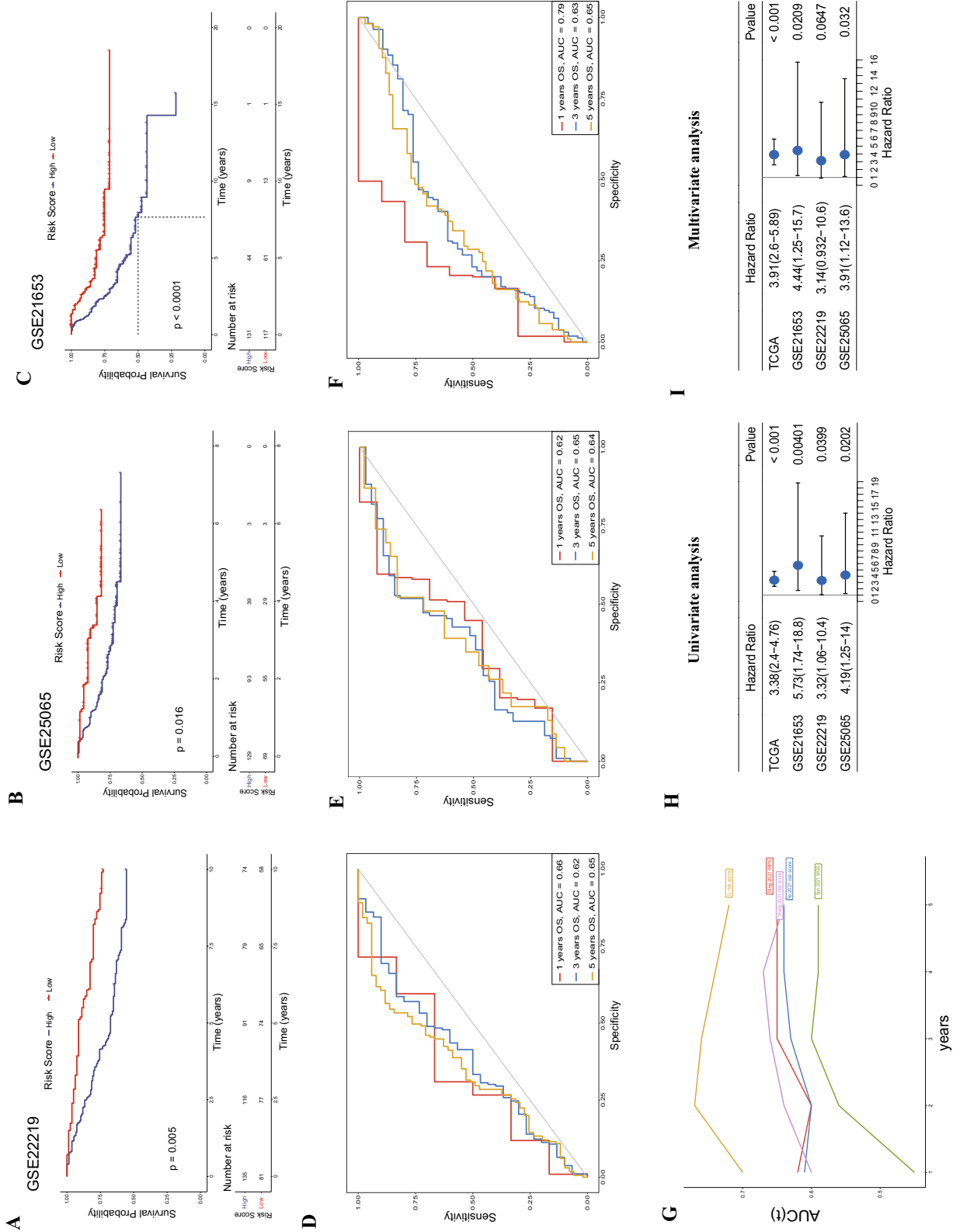


Figure 2. Validation of IL score. Kaplan–Meier curves in validation cohorts of GSE22219 (A), GSE25065 (B), and GSE21653 (C). Sensitivity and specificity of ROC curves in validation cohorts of GSE22219 (D), GSE25065 (E), and GSE21653 (F). (G) Comparison of ROC curves for five risk models. Forest plots of IL score in TCGA, GSE21653, GSE22219, and GSE25065 by univariate Cox regression analysis (H) and multivariate Cox regression analysis (I).

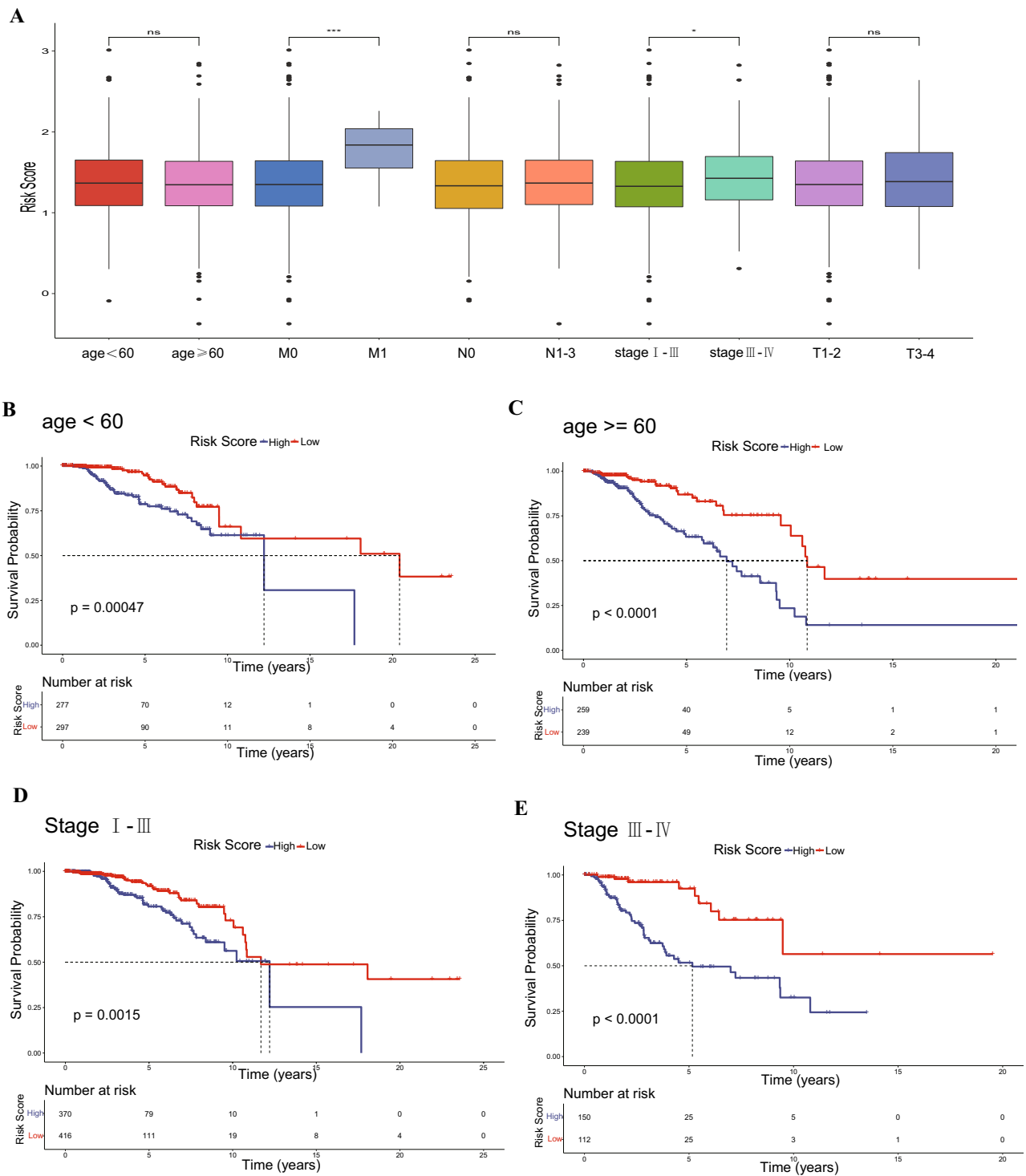


Figure 3. The association between clinical characteristics and IL score. **(A)** Difference between clinical features and IL score. **(B)** Survival analysis of clinical features, including age < 60 years **(B)**, age ≥ 60 years **(C)**, stage I-II **(D)**, stage III-IV **(E)**, T1-2 **(F)**, T3-4 **(G)**, N0 **(H)**, N1-3 **(I)**, M0 **(J)** and M1 **(K)**.

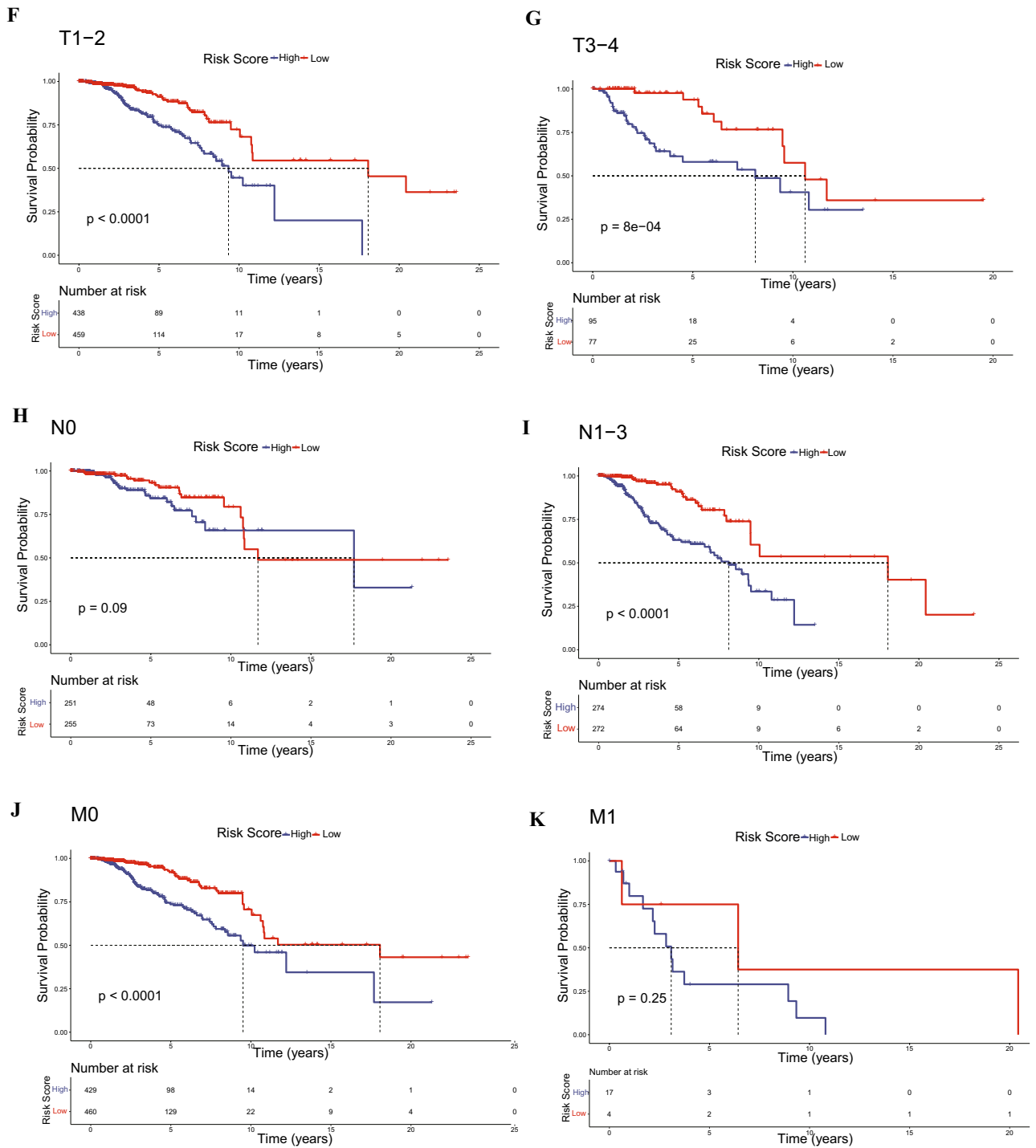


Figure 3. (continued)

lar trend was observed in the other three subtypes, although there was no statistically significant difference due to limited sample size (Fig. 4B). HER2-positive and TNBC had the highest IL score compared to the other molecular subtypes, which was consistent with aforementioned result in Fig. 4A ($p < 2.2e-16$, Fig. 4C). Moreover, TMB values were both significantly elevated in the high-risk group of luminal A and luminal B patients (all $p < 0.001$, Fig. 4D). Subsequently, neoantigen levels were significantly higher in the high-risk group of luminal A patients ($p < 0.05$, Fig. 4E).

Nomogram construction. Univariate and multivariate Cox regression analysis revealed that the IL score was an independent prognostic indicator (HR = 3.91, 95% CI 2.60–5.89, $p < 0.001$, Fig. 5A and B). Next, we constructed a nomogram according to IL score, T stage, N stage, pathological stage, and age (Fig. 5C), which showed a good agreement between actual and predicted survival at 1-, 3- and 5-year (Fig. 5D). The AUC for prediction

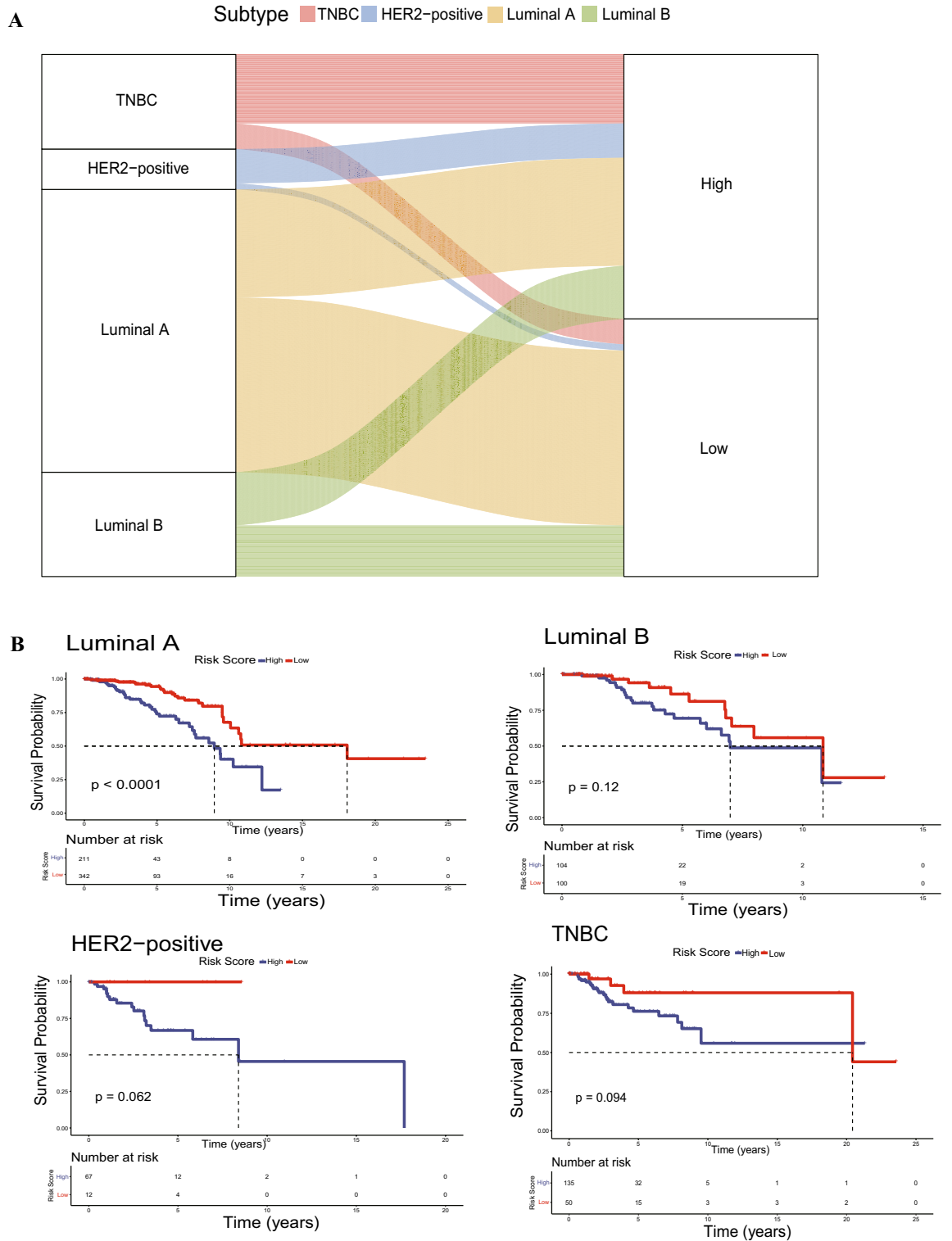


Figure 4. Molecular subtypes in BRCA. (A) Alluvial diagram of four molecular subtypes and IL score. (B) Kaplan–Meier curves of four molecular subtypes. (C) Differences in IL scores among the four molecular subtypes. (D) TMB in the four molecular subtypes. (E) Neoantigen levels of four molecular subtypes. BRCA, breast cancer; TMB, tumor mutation burden.

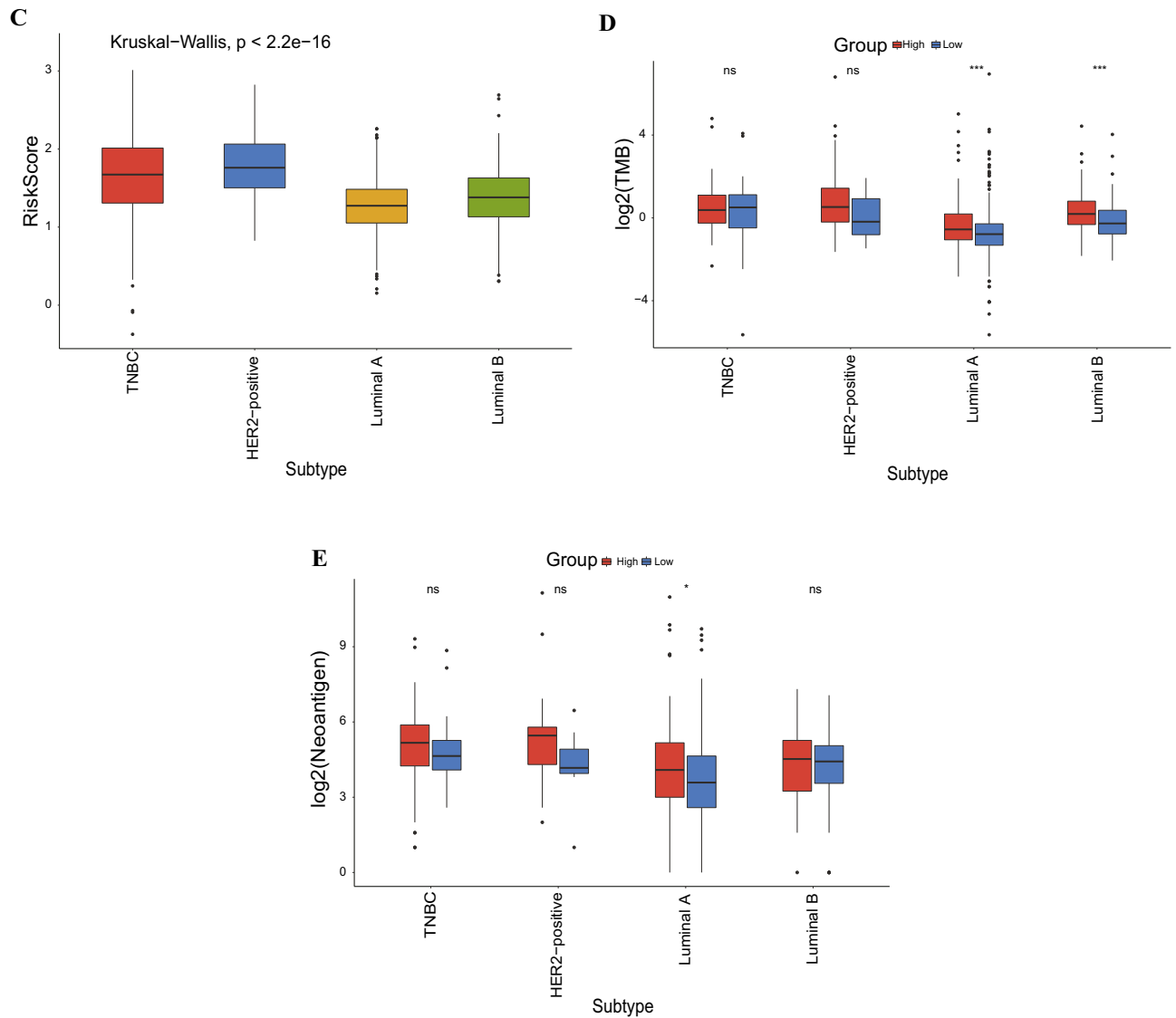


Figure 4. (continued)

survival at 1-, 3- and 5- year was further improved to 0.80, 0.82 and 0.78, respectively, outperforming the risk score and other clinicopathological factors (Fig. 5E–G).

Functional enrichment analysis. To investigate the potential biological activities of the risk model, we performed the GO and KEGG enrichment analysis based on the IL score-associated genes ($R\text{-value} > 0.25$). The results revealed that the IL score-associated genes were primarily enriched in DNA replication, hexose metabolic process, and monosaccharide metabolic process. From the aspect of GO CC, we found these genes mainly clustered in the DNA replication preinitiation complex. And cadherin binding, histone binding, and catalytic activity acting on a tRNA were major concentrated in GO MF (Fig. 6A). Through KEGG analysis we found that the IL score-associated genes were mainly involved in the cell cycle, carbon metabolism, glycolysis, and gluconeogenesis (Fig. 6B).

Genomic landscape. Somatic mutation profiles of BRCA patients from the TCGA cohort were visualized by waterfall plots, which showed that *TP53*, *PIK3CA*, and *TTN* were the most frequently mutated genes in both high-risk and low-risk groups (Fig. 7A and B). However, the prevalence of the top 3 genes, including *TP53*, *TTN*, and *PIK3CA* were significantly differed between the two groups according to Fisher’s test (all $p < 0.05$, Fig. 7C). In addition, as shown in lollipop plots, *PIK3CA* mutations were mainly distributed in specific functional domains, including PI3K-p85B, PI3K-C2, PI3Ka, and P13-P14-kinase (Fig. 7D); the distribution of *TP53* mutations was primarily in specific functional domain P53 (Fig. 7E). In addition, TMB was significantly higher in the high-risk group ($p = 5.1e-15$; Fig. 7F), which was accompanied by a rising level of neoantigen ($p = 2.4e-06$; Fig. 7G).

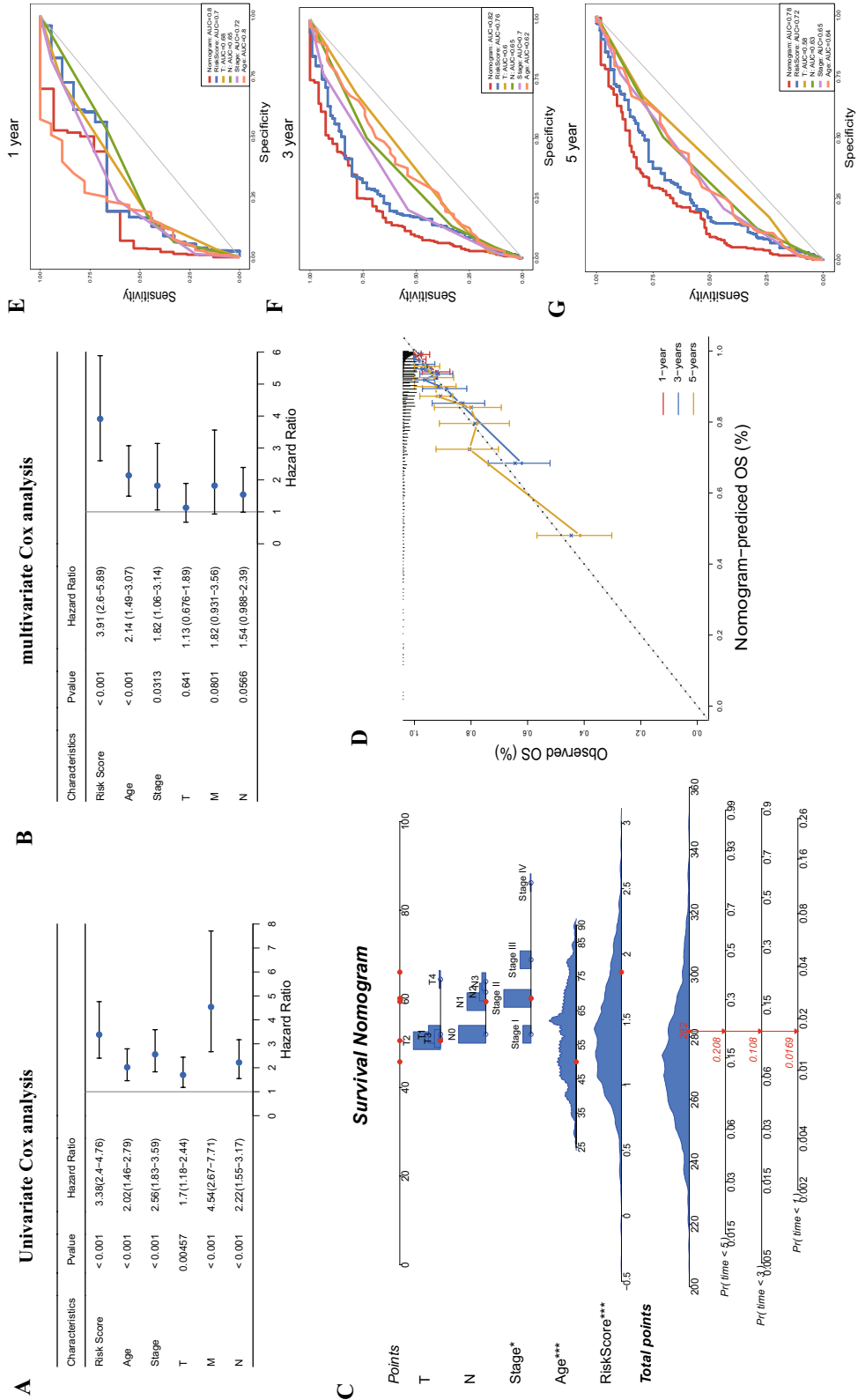
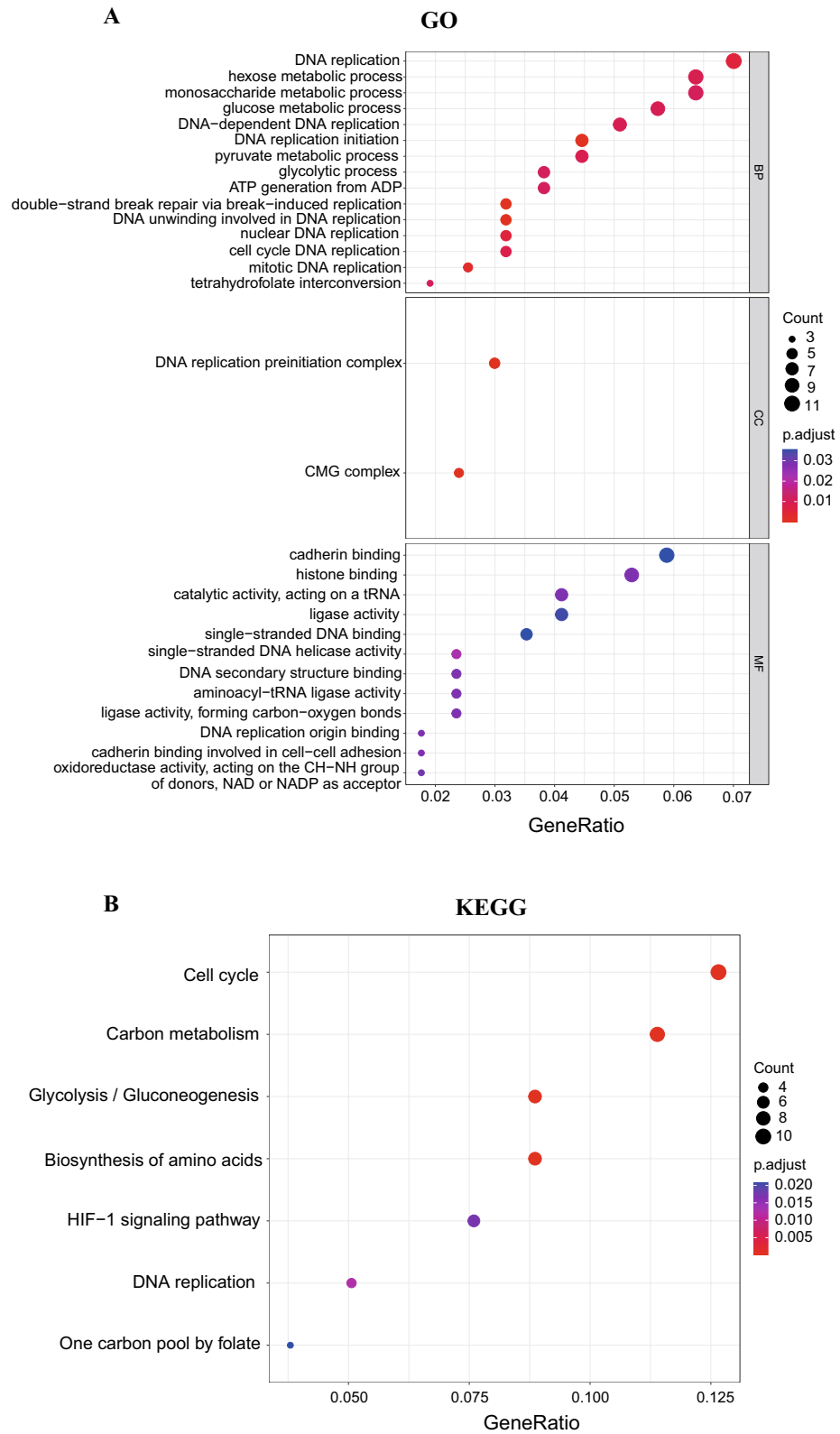


Figure 5. Construction of nomogram. (A) Univariate Cox regression analysis. (B) Multivariate Cox regression analysis. (C) Nomogram for predicting the probability of 1-, 3-, and 5-years OS in the TCGA cohort. (D) Calibration curves for predicting the fitness of the nomogram in 1-, 3-, and 5-years. ROC curves of a nomogram for 1-, 3-, and 5-year (E) in the TCGA cohort. OS: overall survival.



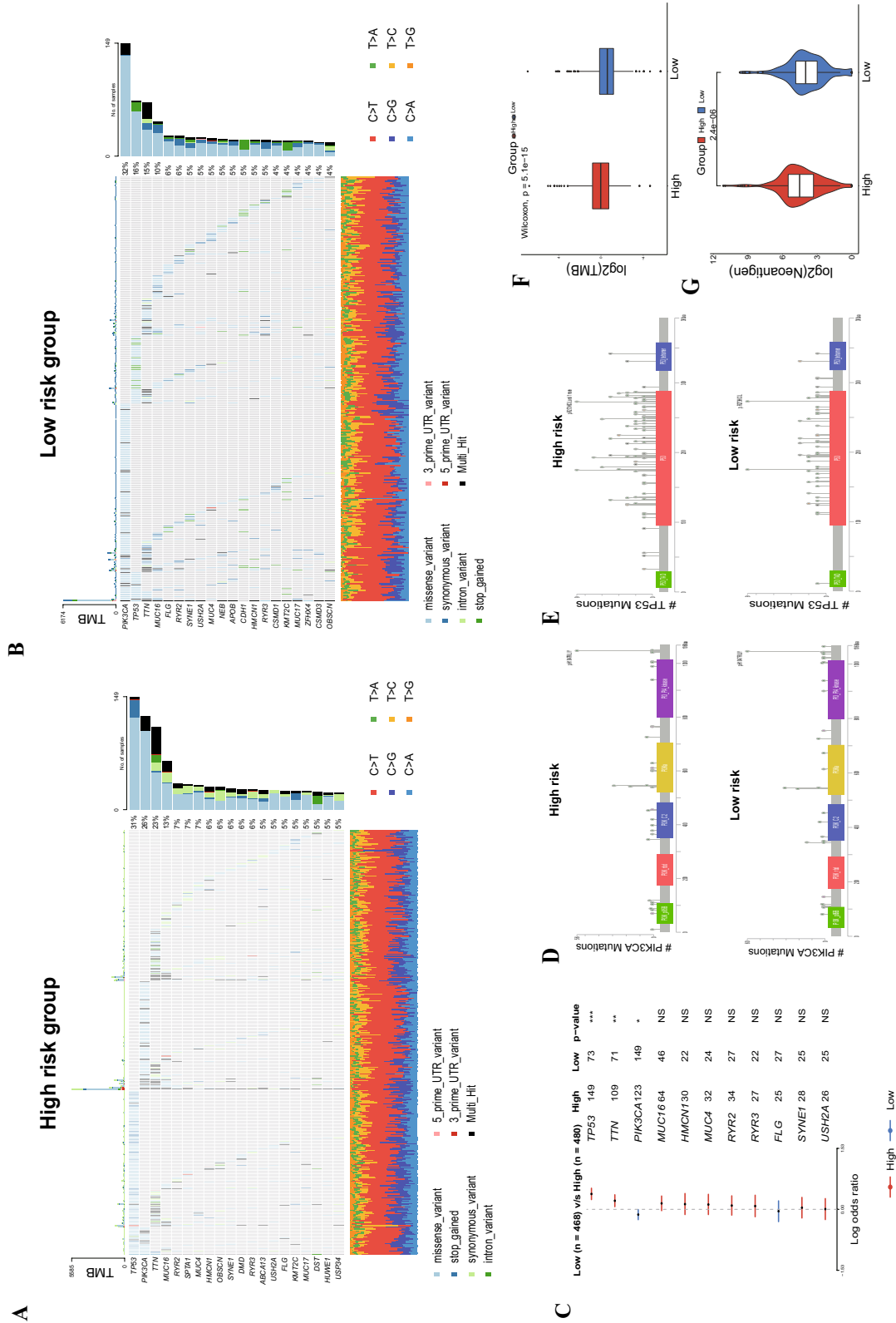


Figure 7. Genetic mutations analysis. Waterfall plots of somatic mutation features in the high-risk group (A), and low-risk group (B). (C) The forest plot exhibits the 11 genes with the highest differences between the high- and low-risk groups. (D) Lollipop plots showing *PIK3CA* gene mutations in the high- and low-risk groups. (E) Lollipop plots showing *TP53* gene mutations in the high- and low-risk groups. (F) The comparison of neoantigen levels between high- and low-risk groups. (G) The comparison of neoantigen levels between high- and low-risk groups.

Association between IL score and tumor microenvironment. We explored the expression level of immunomodulators, and the distribution of these 122 immunomodulators in high- and low-risk groups was shown in Fig. 8A. Interestingly, though high risk group had notable elevated TMB value, more significantly overexpression of immunomodulators were presented in the low-risk group, including CCL1, CCL19, CX3CR1. Subsequently, we identified the differences in TIICs between the high- and low-risk groups, high-risk group had significantly higher abundances of macrophage M0, macrophage M2, neutrophil and NK cell resting, as well as lower abundances of B cell naïve, mast cell activated, monocyte, myeloid dendritic cell resting and T cell CD4+ memory resting than those of low-risk group (all $p < 0.05$, Fig. 8B).

Correlation of the IL score with immunotherapy. First of all, we analyzed the expression profile of 24 immune checkpoints between the high- and low-risk groups. The expression of PD-L1, ADORA2A, BTLA, CD200, CD200R1, KIR3DL1, TIM-3, and VISTA was more prevalent in the low-risk group, whereas CD276 and PVR were up-regulated in the high-risk group (Fig. 9A). We determined, using the ESTIMATE algorithm, that the ESTIMATE score, stromal score and immune score of the low-risk group were significantly higher than those of the high-risk group (Fig. 9B). In addition, the TIDE score ($p = 6.5e-06$) and T cell dysfunction score ($p = 9e-06$) were markedly lower in the high-risk group, indicating that the high-risk group may have a higher possibility to respond to immunotherapy (Fig. 9C). The survival analysis further showed that BRCA patients with high TIDE and lower risk score had the best OS ($p < 0.0001$, Fig. 9D). We subsequently applied GSE173839 to further explore the predictive ability of IL score for immunotherapy response in BRCA. Responders to durvalumab and olaparib therapy had higher risk score than non-responder in the GSE173839 ($p = 0.0045$, Fig. 9E). Additionally, anti-CTLA4 and anti-PD-1 therapy, either alone or in combination, were more beneficial for patients in the low-risk group ($p = 0.0044$; $p = 0.0032$; $p = 0.041$; $p = 0.0053$; Fig. 9F–I). In consequence, we believed that the risk model had the potential to be employed to identify the immunotherapy response of BRCA patients.

Chemotherapeutic sensitivity and treatment efficiency of IL score. We explored the relationship between IL score and therapeutic signatures, such as EGFR network, immune inhibited oncogenic pathways, and radiotherapy predicted pathways, the results showed that the high-risk group showed a significant positive correlation in the radiotherapy predicted pathway (Fig. 10A). Furthermore, the target genes from chemotherapy drugs indicated that the low-risk group was positively associated with response to docetaxel and paclitaxel (Fig. 10B). According to the GSE194040 cohort, the pCR proportion of neoadjuvant chemotherapy (paclitaxel + pertuzumab + trastuzumab) in the high-risk group was significantly higher than that in the low-risk group, but it did not reach statistical significance ($p = 0.074$, Fig. 10C), and twelve other neoadjuvant chemotherapy treatments were presented in Fig. S3. Moreover, neoadjuvant chemotherapy (paclitaxel + pertuzumab + trastuzumab) had a higher risk score in pCR than non-pCR ($p = 0.0086$, Fig. 10D), risk scores for the other eleven neoadjuvant chemotherapy treatments were shown in Fig. S4. These results illustrated the potential therapeutic implications of the risk model for chemosensitivity in BRCA.

Discussion

In this study, multiple machine learning algorithms, such as RF, LASSO, and multivariate Cox regression analysis, were used to extensively examine the role of IL genes and ultimately create a 21 IL gene-related prognostic model in BRCA. Not only did high-risk and low-risk BRCA individuals exhibit varied prognoses, but also distinct clinicopathologic characteristics, TME landscape, immunotherapy, and chemotherapy. Our research demonstrated the predictive relevance, biological significance, and therapeutic potential of IL genes in BRCA.

As formentioned, IL plays an important role in human disease and is closely related to the occurrence and development of cancer⁴⁰, and multiple IL genes have been confirmed their function in the proliferation and growth of BRCA cells^{41,42}. Considering the critical role of IL genes in cancer, it is necessary to comprehensively investigate their prognostic value in BRCA. Notably, in current study, we found 21 genes with significant prognostic predictive functions, and IL18BP, IL17D, and IL23A were identified for the first time as being significantly correlated with the prognosis of BRCA. At the same time, IRAK4, IL21, IL22RA2, IL26, IL27, IL1RN and IL13RA1 were significantly related to the OS of BRCA. A previous research discovered that variants in IRAK4 are highly associated with the prognosis of BRCA, which increase nearly fivefold risk of developing BRCA⁴³. As a member of the IL-10 family of cytokines, IL26 participates in inflammatory signaling and is overexpressed in TNBC⁴⁴; nevertheless, a previous study found that the introduction of recombinant IL21 could enhance the anticancer impact of trastuzumab in the treatment of metastatic HER2-positive patients and that IL21 expression is essential for CD8 T cells to optimal anti-HER2 antibody effectiveness⁴⁵. Furthermore, it was reported that the mutant alleles of IL1RN are correlated with shorter OS in BRCA patients, mainly by altering IL1 receptor binding and resulting in the production of IL-1, which produces a proinflammatory status and enhances the aggressiveness of BRCA tumor⁴⁶. In addition, several studies also have confirmed that IL13RA1 is strongly related to the survival of BRCA^{37,47}. All of these data corroborated our conclusions and demonstrated the prognostic function of IL genes in BRCA. Additionally, the IL gene play a key role on the biological functions, recent work from Liubomirski et al.⁴⁸ found that the interaction of Notch with IL-6 and the transcription factor STAT3 enhances pro-tumor functions in BRCA. There is evidence that inactivating mutations of P53 can guide the loss of methylation of IL-6, leading to epigenetic reprogramming and driving the development of BRCA⁴⁹. Accordingly, we constructed a 21 IL gene-related risk model in BRCA, which is proved to be robust and reliable with the AUC for OS of the IL score was 0.70 in 1 year, 0.76 in 3 years, and 0.72 in 5 years in the TCGA dataset. Besides, based on the nomogram, the AUCs for the OS predictions for 1, 3, and 5 years were 0.80, 0.82 and 0.78, respectively, demonstrating the practical application value of this risk model in clinical management for BRCA patients.

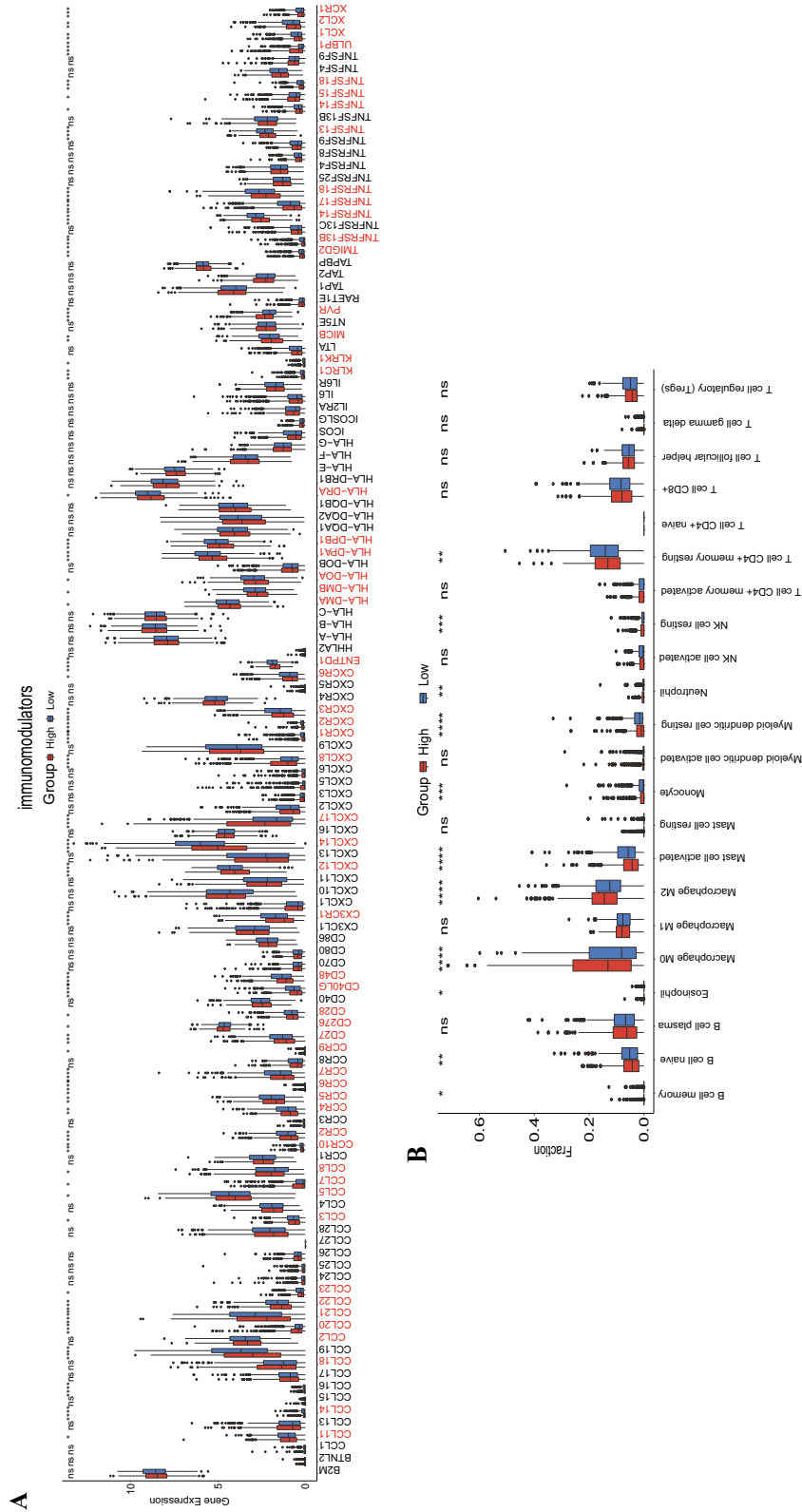


Figure 8. The role of IL score in TME. (A) Expression profile of immunomodulators in high- and low-risk groups. (B) The proportions of T1Cs between high- and low-risk groups. TME, tumor microenvironment; T1Cs, tumor infiltrates immune cells.

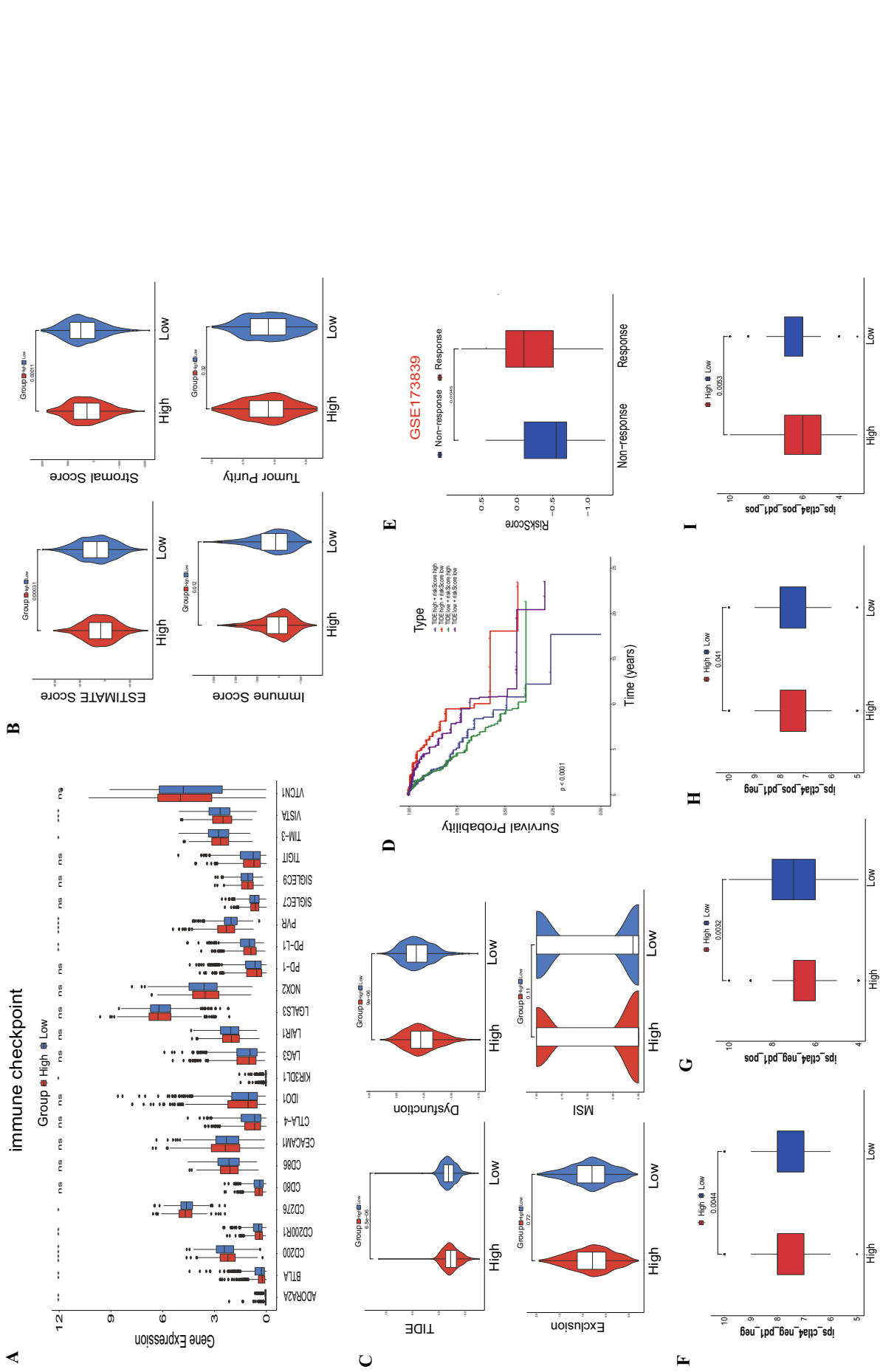


Figure 9. Immunotherapy landscape in BRCA. (A) The correlation between the immune checkpoints and IL score. (B) Relationship of IL score and ESTIMATE score, stromal score, immune score, and tumor purity. (C) Association of IL score and TIDE, T cell dysfunction, T cell exclusion, and MSI. (D) Survival analysis with different combinations of TIDE and IL scores. (E) The risk scores of non-response and response in the GSE173839 cohort. (F–I) IPS difference of BRCA with different statuses of CTLA-4 or PD-1. TIDE, Tumor Immune Dysfunction and Exclusion; MSI, microsatellite instability; IPS, immunophenoscore.

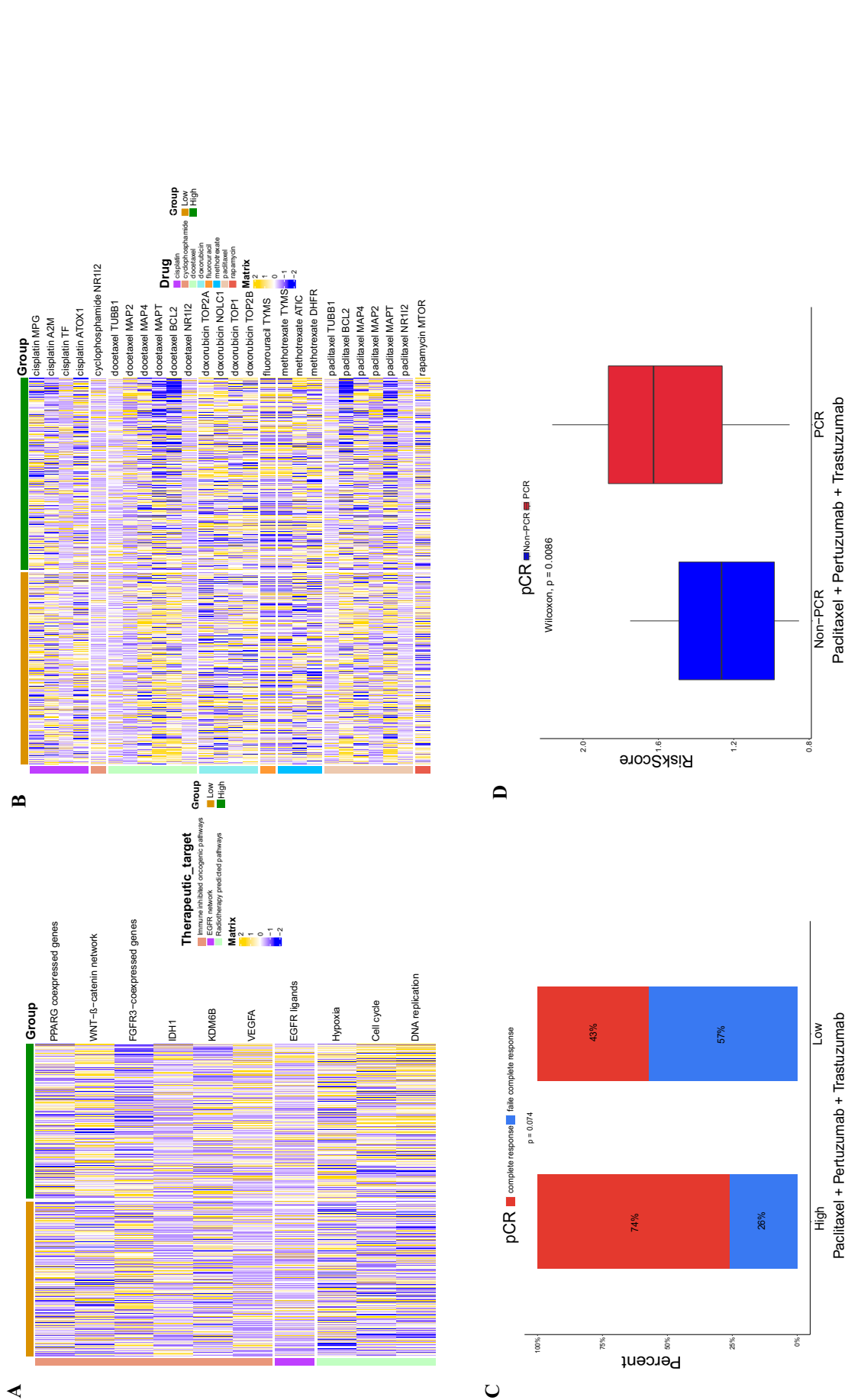


Figure 10. Sensitivity analysis of chemotherapy. **(A)** Expression profiles of three treatment characteristics between high- and low-risk groups. **(B)** Correlation between chemotherapeutic drug target genes and IL score in BRCA. **(C)** The pCR ratio of neoadjuvant chemotherapy (paclitaxel + pertuzumab + trastuzumab) in the high- and low-risk groups. **(D)** Comparison of risk scores for neoadjuvant chemotherapy (paclitaxel + pertuzumab + trastuzumab) in non-pCR and pCR. pCR, pathological complete response.

Moreover, the GO and KEGG enrichment analysis found that the DNA replication, glucose metabolic process, and glycolysis pathways were the most enriched in the pathways related to IL score. Studies have discovered that cancer is correlated with errors that occur during DNA replication⁵⁰, and the inhibition of DNA replication is shown to be beneficial to improve the prognosis of BRCA⁵¹. Recent research has demonstrated the crucial role played by glucose metabolic process in the growth of TNBC⁵². Besides, glycolysis exhibits elevated activity in BRCA, mainly through the effect of mTOR hyperactivation⁵³.

Furthermore, we systematically studied the genomic differences between high- and low-risk groups, and *TP53* was the most frequently mutated gene in the high-risk group. It has been reported that nearly one third BRCA patients have *TP53* mutations, which reduces the transcriptional activity of p53, resulting in the development of BRCA⁵⁴. Funda et al.⁵⁵ have proved that BRCA patients with *TP53* alterations had significantly inferior recurrence-free survival, progression-free survival, and overall survival. Accordingly, *TP53* mutations may contribute to the poor prognosis in the high-risk group. On the contrary, the mutation frequency of *PIK3CA*, which is a prevalent gene in BRCA⁵⁶, was higher in the low-risk group, and is also an effective prognostic marker of BRCA⁵⁷. The increasing prevalence of *PIK3CA* may be explained by higher proportion of hormone receptor positive patients in the low risk group. As alpelisib, a α -selective *PIK3CA* inhibitor, has been approved by FDA for treating BRCA patients with this kind of mutants, low risk group had potentially higher clinical benefits to the targeted therapy⁵⁸. Moreover, TMB is associated with a higher probability of response to immune checkpoint inhibitors (ICIs)⁵⁹, and we found that high-risk group had a significantly higher TMB. As Targeted Agent and Profiling Utilization Registry (TAPUR) and GeparNuevo Study both revealed that ICI monotherapy showed satisfied antitumor activity in BRCA patients with high TMB^{60,61}, it suggested that high risk patients may have higher response rate to ICIs.

Notably, more TNBC and HER2-positive BRCA patients were in the high-risk group, which was consistent with the fact that TNBC and HER2-positive BRCA patients exhibit highly invasiveness and high proliferation characteristics and have a worse prognosis than other types of BRCA^{4,62}. This difference in subtype distribution was consistent with the finding in the TMB that TNBC had the highest TMB compared to other BRCA subtypes⁶³. On the other hand, TME is a cellular community composed of tumor cells, endothelial cells, and stroma, which plays a key role in tumor control and progression, and is highly associated with malignant cell immune evasion, chemotherapy resistance, and tumor cell proliferation^{64,65}. The expression of immunomodulators reflects the immunological characteristics of TME, MHC, receptors, chemokines, and immune stimulators are the main immunomodulators⁶⁶. In the current study, we comprehensively analyzed the expression levels of immunomodulators between two groups. Interestingly, we discovered that most immunomodulators were highly expressed in the low-risk group, which contradicted our finding that the high-risk group would respond better to immunotherapy. Several findings demonstrated that the high expression of *CCL11* and *CCL19* predicts better OS of BRCA^{67,68}, and Li et al.⁶⁹ indicated that the high expression of *CCL21* is highly associated with a lower distant recurrence rate of BRCA, through increasing the infiltration of CD8+T cells, thus, we confirmed that these studies may explain the up-regulation of immunomodulators in the low-risk group. Consequently, we hypothesized that the high-risk group would respond better to immunotherapy. In addition, we studied the abundances of the TIICs using CIBERSORT to evaluate the immune heterogeneity between high- and low-groups. Compared with the low-risk group, the increased tumor infiltrating abundances of macrophage M0, macrophage M2, neutrophil, and NK cell resting were discovered in the high-risk group. It has been recently highlighted that TIICs are a novel treatment target for immunotherapy of cancer⁷⁰, in which macrophage M0 is a key factor in regulating immune responses⁷¹. Previous studies proved that the polarization of macrophage M2 is highly related to BRCA and that reducing macrophage M2 can help to suppress the process of BRCA^{72,73}. These findings indicated that patients in the high-risk group have a worse prognosis, which may be related to the high abundance of macrophage M2. Collectively, the IL score revealed the vital role of TME in BRCA and might offer a novel strategy for the treatment of BRCA patients.

At present, immunotherapy is an approach to cancer treatment by activating the anti-tumor immune response, in which immune checkpoints are a group of inhibitory immune receptors that exert an immunosuppressive effect on the cell surface⁷⁴, which can be beneficial to a subset of cancer patients in immunotherapy⁷⁵. In this study, we found higher expression levels of immune checkpoints in the low-risk group, including PD-L1, TIM-3, VISTA, ADORA2A, BTLA, CD200, CD200R1, and KIR3DL1. Nevertheless, TIDE is shown to have better prognostic ability than other indicators⁷⁶, we found a significantly lower TIDE score and T cell dysfunction score in the high-risk group, indicating a better response to immunotherapy in the high-risk group. In particular, PD-L1 inhibitor durvalumab and PARP inhibitor olaparib can be used for the treatment of HER2-negative BRCA⁷⁷, our research found that the responders to durvalumab and olaparib therapy had higher risk score than non-responders in GSE173839 dataset, further suggesting that the BRCA patients in high-risk group response better to immunotherapy. Together, these results demonstrated that the IL score can be employed as a potential biomarker for BRCA immunotherapy. Docetaxel and paclitaxel are known to be commonly used chemotherapeutic drugs for BRCA^{78,79}, and sensitivity analysis of chemotherapeutic drugs showed that the low-risk group had a significantly higher response to docetaxel and paclitaxel. Besides, neoadjuvant chemotherapy is a novel treatment method for BRCA, which aims to reduce the tumor size and improve OS in BRCA patients with locally advanced cancer⁸⁰, and the combination neoadjuvant chemotherapy (paclitaxel + pertuzumab + trastuzumab) had a higher risk score in pCR. Consequently, we speculated that there is a potential interactive correlation between IL score and chemotherapeutic drug sensitivity of BRCA.

In summary, a 21-gene signature based on IL genes was developed and validated to have robust performance in predicting the survival outcomes of BRCA patients. In addition, the risk model was proved to provide novel insights into clinical characteristics, TME landscape, immunotherapy response, and chemotherapeutic drug sensitivity between high- and low-risk groups. Taken together, we believed that our findings could offer a powerful prognostic biomarker for individualized prediction of clinical decision-making and provide a theoretical basis for further studies in patients with BRCA.

Data availability

The datasets generated and analyzed during the current study are available in The Cancer Genome Atlas (TCGA) repository (<http://xena.ucsc.edu/>), and GSE22219, GSE25065, and GSE21653 from Gene Expression Omnibus (GEO) repository (<https://www.ncbi.nlm.nih.gov/geo/>). Access to both repositories is not required accession number. The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Received: 17 September 2022; Accepted: 24 November 2022

Published online: 28 November 2022

References

- Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249. <https://doi.org/10.3322/caac.21660> (2021).
- Bonfiglio, R. & Di Pietro, M. L. The impact of oral contraceptive use on breast cancer risk: State of the art and future perspectives in the era of 4P medicine. *Semin. Cancer Biol.* **72**, 11–18. <https://doi.org/10.1016/j.semcancer.2020.10.008> (2021).
- Dittmer, J. Breast cancer stem cells: Features, key drivers and treatment options. *Semin. Cancer Biol.* **53**, 59–74. <https://doi.org/10.1016/j.semcancer.2018.07.007> (2018).
- Harbeck, N. *et al.* Breast cancer. *Nat. Rev. Dis. Primers* **5**, 66. <https://doi.org/10.1038/s41572-019-0111-2> (2019).
- Wang, S. *et al.* Clinical significance and immunogenomic landscape analyses of the immune cell signature based prognostic model for patients with breast cancer. *Brief Bioinform.* **22**, 4. <https://doi.org/10.1093/bib/bbaa311> (2021).
- Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Investig.* <https://doi.org/10.1172/jci45014> (2011).
- Emens, L. A. Breast cancer immunotherapy: Facts and hopes. *Clin. Cancer Res.* **24**, 511–520. <https://doi.org/10.1158/1078-0432.Ccr-16-3001> (2018).
- Xie, H. Y., Shao, Z. M. & Li, D. Q. Tumor microenvironment: Driving forces and potential therapeutic targets for breast cancer metastasis. *Chin. J. Cancer* **36**, 36. <https://doi.org/10.1186/s40880-017-0202-y> (2017).
- Ly, W. *et al.* Identification of an aging-related gene signature in predicting prognosis and indicating tumor immune microenvironment in breast cancer. *Front. Oncol.* **11**, 796555. <https://doi.org/10.3389/fonc.2021.796555> (2021).
- Zhou, Y. *et al.* Identification of senescence-related subtypes, establishment of a prognosis model, and characterization of a tumor microenvironment infiltration in breast cancer. *Front. Immunol.* **13**, 921182. <https://doi.org/10.3389/fimmu.2022.921182> (2022).
- Gonzalez-Aparicio, M. & Alfaro, C. Implication of interleukin family in cancer pathogenesis and treatment. *Cancers (Basel)* **13**, 1016. <https://doi.org/10.3390/cancers13051016> (2021).
- Akdis, M. *et al.* Interleukins (from IL-1 to IL-38), interferons, transforming growth factor β , and TNF- α : Receptors, functions, and roles in diseases. *J. Allergy Clin. Immunol.* **138**, 984–1010. <https://doi.org/10.1016/j.jaci.2016.06.033> (2016).
- Briukhovetska, D. *et al.* Interleukins in cancer: From biology to therapy. *Nat. Rev. Cancer* **21**, 481–499. <https://doi.org/10.1038/s41568-021-00363-z> (2021).
- Waldmann, T. A. Cytokines in cancer immunotherapy. *Cold Spring Harb. Perspect. Biol.* <https://doi.org/10.1101/cshperspect.a028472> (2017).
- Coffelt, S. B. *et al.* IL-17-producing $\gamma\delta$ T cells and neutrophils conspire to promote breast cancer metastasis. *Nature* <https://doi.org/10.1038/nature14282> (2015).
- Kuan, E. L. & Ziegler, S. F. A tumor-myeloid cell axis, mediated via the cytokines IL-1 α and TSLP, promotes the progression of breast cancer. *Nat. Immunol.* <https://doi.org/10.1038/s41590-018-0066-6> (2018).
- Setrerrahmane, S. & Xu, H. Tumor-related interleukins: Old validated targets for new anti-cancer drug development. *Mol. Cancer* **16**, 153. <https://doi.org/10.1186/s12943-017-0721-9> (2017).
- Siersbaek, R. *et al.* IL6/STAT3 signaling hijacks estrogen receptor α enhancers to drive breast cancer metastasis. *Cancer Cell* **38**, 412–423.e419. <https://doi.org/10.1016/j.ccell.2020.06.007> (2020).
- Weng, Y. *et al.* MCT-1/miR-34a/IL-6/IL-6R signaling axis promotes EMT progression, cancer stemness and M2 macrophage polarization in triple-negative breast cancer. *Mol. Cancer* **18**, 42. <https://doi.org/10.1186/s12943-019-0988-0> (2019).
- Moossavi, M., Parsamanesh, N., Bahrami, A., Atkin, S. L. & Sahebkar, A. Role of the NLRP3 inflammasome in cancer. *Mol. Cancer* <https://doi.org/10.1186/s12943-018-0900-3> (2018).
- Qiu, N. *et al.* Tumor-associated macrophage and tumor-cell dually transfecting polyplexes for efficient interleukin-12 cancer gene therapy. *Adv. Mater.* **33**, e2006189. <https://doi.org/10.1002/adma.202006189> (2021).
- Castano, Z. *et al.* IL-1 β inflammatory response driven by primary breast cancer prevents metastasis-initiating cell colonization. *Nat. Cell Biol.* **20**, 1084–1097. <https://doi.org/10.1038/s41556-018-0173-5> (2018).
- Kaplanov, I. *et al.* Blocking IL-1 β reverses the immunosuppression in mouse breast cancer and synergizes with anti-PD-1 for tumor abrogation [Immunology and Inflammation]. *Proc. Natl. Acad. Sci. USA* <https://doi.org/10.1073/pnas.1812266115> (2019).
- Zhao, Z. *et al.* Increased interleukin-35 expression in tumor-infiltrating lymphocytes correlates with poor prognosis in patients with breast cancer. *Cytokine* **89**, 76–81. <https://doi.org/10.1016/j.cyto.2016.09.012> (2017).
- You, Y. *et al.* IL-21 gene polymorphism is associated with the prognosis of breast cancer in Chinese populations. *Breast Cancer Res. Treat.* **137**, 893–901. <https://doi.org/10.1007/s10549-012-2401-1> (2013).
- Hanavadi, S., Martin, T. A., Watkins, G., Mansel, R. E. & Jiang, W. G. Expression of interleukin 11 and its receptor and their prognostic value in human breast cancer. *Ann. Surg. Oncol.* **13**, 802–808. <https://doi.org/10.1245/aso.2006.05.028> (2006).
- Taylor, J. M. Random survival forests. *J. Thorac. Oncol.* **6**, 1974–1975. <https://doi.org/10.1097/JTO.0b013e318233d835> (2011).
- Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.: Ser. B (Methodol.)* **58**, 267–288 (1996).
- Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).
- Rich, J. T. *et al.* A practical guide to understanding Kaplan-Meier curves. *Otolaryngol. Head Neck Surg.* **143**, 331–336. <https://doi.org/10.1016/j.otohns.2010.05.007> (2010).
- Loibl, S., Poortmans, P., Morrow, M., Denkert, C. & Curigliano, G. Breast cancer. *Lancet* **397**, 1750–1769. [https://doi.org/10.1016/S0140-6736\(20\)32381-3](https://doi.org/10.1016/S0140-6736(20)32381-3) (2021).
- Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. <https://doi.org/10.1093/nar/28.1.27> (2000).
- Harris, M. A. *et al.* The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**, D258–261. <https://doi.org/10.1093/nar/gkh036> (2004).
- Xiao, Y. & Yu, D. Tumor microenvironment as a therapeutic target in cancer. *Pharmacol. Ther.* **221**, 107753. <https://doi.org/10.1016/j.pharmthera.2020.107753> (2021).
- Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457. <https://doi.org/10.1038/nmeth.3337> (2015).

36. Ding, S. *et al.* Identification of a novel immune-related prognostic signature associated with tumor microenvironment for breast cancer. *Int. Immunopharmacol.* **100**, 108122. <https://doi.org/10.1016/j.intimp.2021.108122> (2021).
37. He, M., Hu, C., Deng, J., Ji, H. & Tian, W. Identification of a novel glycolysis-related signature to predict the prognosis of patients with breast cancer. *World J. Surg. Oncol.* **19**, 294. <https://doi.org/10.1186/s12957-021-02409-w> (2021).
38. Sun, X. *et al.* A novel metabolic gene signature-based nomogram to predict overall survival in breast cancer. *Ann. Transl. Med.* **9**, 367. <https://doi.org/10.21037/atm-20-4813> (2021).
39. Wang, D. *et al.* Identification of the prognostic value of ferroptosis-related gene signature in breast cancer patients. *BMC Cancer* **21**, 645. <https://doi.org/10.1186/s12885-021-08341-2> (2021).
40. Holder, P. G. *et al.* Engineering interferons and interleukins for cancer immunotherapy. *Adv. Drug Deliv. Rev.* **182**, 114112. <https://doi.org/10.1016/j.addr.2022.114112> (2022).
41. Huang, S. C. *et al.* TGF- β 1 secreted by Tregs in lymph nodes promotes breast cancer malignancy via up-regulation of IL-17RB. *EMBO Mol. Med.* **9**, 1660–1680. <https://doi.org/10.15252/emmm.201606914> (2017).
42. Alraouji, N. N. & Aboessekhra, A. Tocilizumab inhibits IL-8 and the proangiogenic potential of triple negative breast cancer cells. *Mol. Carcinog.* **60**, 51–59. <https://doi.org/10.1002/mc.23270> (2021).
43. Yeyeodu, S. T. *et al.* IRAK4 and TLR3 sequence variants may alter breast cancer risk among African-American WOMEN. *Front. Immunol.* **4**, 338. <https://doi.org/10.3389/fimmu.2013.00338> (2013).
44. Trotter, T. N. *et al.* IL26, a Noncanonical Mediator Of DNA inflammatory stimulation, promotes tnbc engraftment and progression in association with neutrophils. *Cancer Res.* **80**, 3088–3100. <https://doi.org/10.1158/0008-5472.Can-18-3825> (2020).
45. Mittal, D. *et al.* Improved treatment of breast cancer with anti-HER2 therapy requires interleukin-21 signaling in CD8+ T cells. *Cancer Res.* **76**, 264–274. <https://doi.org/10.1158/0008-5472.Can-15-1567> (2016).
46. Grimm, C. *et al.* The prognostic value of four interleukin-1 gene polymorphisms in Caucasian women with breast cancer: A multicenter study. *BMC Cancer* **9**, 78. <https://doi.org/10.1186/1471-2407-9-78> (2009).
47. Jiang, F., Wu, C., Wang, M., Wei, K. & Wang, J. Identification of novel cell glycolysis related gene signature predicting survival in patients with breast cancer. *Sci. Rep.* **11**, 3986. <https://doi.org/10.1038/s41598-021-83628-9> (2021).
48. Liubomirski, Y. & Ben-Baruch, A. Notch-inflammation networks in regulation of breast cancer progression. *Cells* **9**, 201. <https://doi.org/10.3390/cells9071576> (2020).
49. D'Anello, L. *et al.* Epigenetic control of the basal-like gene expression profile via Interleukin-6 in breast cancer cells. *Mol. Cancer* **9**, 300. <https://doi.org/10.1186/1476-4598-9-300> (2010).
50. Tomasetti, C., Li, L. & Vogelstein, B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* **355**, 1330–1334. <https://doi.org/10.1126/science.aaf9011> (2017).
51. Zhang, H. *et al.* Endocytic protein intersectin1-S shuttles into nucleus to suppress the DNA replication in breast cancer. *Cell Death Dis.* **12**, 922. <https://doi.org/10.1038/s41419-021-04218-1> (2021).
52. He, J. *et al.* MIG-6 is essential for promoting glucose metabolic reprogramming and tumor growth in triple-negative breast cancer. *EMBO Rep.* **22**, e50781. <https://doi.org/10.15252/embr.202050781> (2021).
53. Holloway, R. W. & Marignani, P. A. Targeting mTOR and glycolysis in HER2-positive breast cancer. *Cancers (Basel)* **13**, 2922. <https://doi.org/10.3390/cancers13122922> (2021).
54. Shahbandi, A., Nguyen, H. D. & Jackson, J. G. TP53 mutations and outcomes in breast cancer: Reading beyond the headlines. *Trends Cancer* **6**, 98–110. <https://doi.org/10.1016/j.trecan.2020.01.007> (2020).
55. Meric-Bernstam, F. *et al.* Survival outcomes by TP53 mutation status in metastatic breast cancer. *JCO Precis. Oncol.* <https://doi.org/10.1200/pon.17.00245> (2018).
56. Verret, B., Cortes, J., Bachelot, T., Andre, F. & Arnedos, M. Efficacy of PI3K inhibitors in advanced breast cancer. *Ann. Oncol.* **30**, x12–x20. <https://doi.org/10.1093/annonc/mdz381> (2019).
57. Jiang, Y. Z., Yu, K. D., Bao, J., Peng, W. T. & Shao, Z. M. Favorable prognostic impact in loss of TP53 and PIK3CA mutations after neoadjuvant chemotherapy in breast cancer. *Cancer Res.* **74**, 3399–3407. <https://doi.org/10.1158/0008-5472.Can-14-0092> (2014).
58. Verret, B., Cortes, J., Bachelot, T., Andre, F. & Arnedos, M. Efficacy of PI3K inhibitors in advanced breast cancer. *Ann. Oncol.* **30**(Suppl 10), x12–x20. <https://doi.org/10.1093/annonc/mdz381> (2019).
59. Fusco, M. J., West, H. J. & Walko, C. M. Tumor mutation burden and cancer treatment. *JAMA Oncol.* **7**, 316. <https://doi.org/10.1001/jamaoncol.2020.6371> (2021).
60. Alva, A. S. *et al.* Pembrolizumab in patients with metastatic breast cancer with high tumor mutational burden: Results from the targeted agent and profiling utilization registry (TAPUR) study. *J. Clin. Oncol.* <https://doi.org/10.1200/jco.20.02923> (2021).
61. Karn, T. *et al.* Tumor mutational burden and immune infiltration as independent predictors of response to neoadjuvant immune checkpoint inhibition in early TNBC in GeparNuevo. *Ann. Oncol.* <https://doi.org/10.1016/j.annonc.2020.05.015> (2020).
62. Hua, Z., White, J. & Zhou, J. Cancer stem cells in TNBC. *Semin. Cancer Biol.* **82**, 26–34. <https://doi.org/10.1016/j.semcancer.2021.06.015> (2022).
63. Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399. <https://doi.org/10.1038/nature10933> (2012).
64. Downs-Canner, S. M., Meier, J., Vincent, B. G. & Serody, J. S. B cell function in the tumor microenvironment. *Annu. Rev. Immunol.* **40**, 169–193. <https://doi.org/10.1146/annurev-immunol-101220-015603> (2022).
65. Cheng, H. S., Lee, J. X. T., Wahli, W. & Tan, N. S. Exploiting vulnerabilities of cancer by targeting nuclear receptors of stromal cells in tumor microenvironment. *Mol. Cancer* **18**, 51. <https://doi.org/10.1186/s12943-019-0971-9> (2019).
66. Hu, J. *et al.* Siglec15 shapes a non-inflamed tumor microenvironment and predicts the molecular subtype in bladder cancer. *Theranostics* **11**, 3089–3108. <https://doi.org/10.7150/thno.53649> (2021).
67. Wang, Y. *et al.* Identification of differentially expressed hub genes associated with immune cell recruitment in claudin-low breast cancer. *Front. Oncol.* **12**, 848206. <https://doi.org/10.3389/fonc.2022.848206> (2022).
68. Wang, J. *et al.* CCL19 has potential to be a potential prognostic biomarker and a modulator of tumor immune microenvironment (TIME) of breast cancer: A comprehensive analysis based on TCGA database. *Aging (Albany NY)* **14**, 4158–4175. <https://doi.org/10.18632/aging.204081> (2022).
69. Li, K. *et al.* CD8(+) T cell immunity blocks the metastasis of carcinogen-exposed breast cancer. *Sci. Adv.* **7**, 26. <https://doi.org/10.1126/sciadv.abd8936> (2021).
70. Kitamura, T., Qian, B. Z. & Pollard, J. W. Immune cell promotion of metastasis. *Nat. Rev. Immunol.* **15**, 73–86. <https://doi.org/10.1038/nri3789> (2015).
71. Nakkala, J. R. *et al.* Macrophage membrane-functionalized nanofibrous mats and their immunomodulatory effects on macrophage polarization. *Acta Biomater.* **141**, 24–38. <https://doi.org/10.1016/j.actbio.2021.12.026> (2022).
72. Esser, A. K. *et al.* Nanotherapy delivery of c-myc inhibitor targets protumor macrophages and preserves antitumor macrophages in breast cancer. *Theranostics* **10**, 7510–7526. <https://doi.org/10.7150/thno.44523> (2020).
73. Deng, Y. *et al.* Sphingomyelin synthase 2 facilitates M2-like macrophage polarization and tumor progression in a mouse model of triple-negative breast cancer. *Acta Pharmacol. Sin.* **42**, 149–159. <https://doi.org/10.1038/s41401-020-0419-1> (2021).
74. He, X. & Xu, C. Immune checkpoint signaling and cancer immunotherapy. *Cell Res.* **30**, 660–669. <https://doi.org/10.1038/s41422-020-0343-4> (2020).
75. Sharma, P. & Allison, J. P. Dissecting the mechanisms of immune checkpoint therapy. *Nat. Rev. Immunol.* **20**, 75–76. <https://doi.org/10.1038/s41577-020-0275-8> (2020).

76. Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **24**, 1550–1558. <https://doi.org/10.1038/s41591-018-0136-1> (2018).
77. Pusztai, L. *et al.* Durvalumab with olaparib and paclitaxel for high-risk HER2-negative stage II/III breast cancer: Results from the adaptively randomized I-SPY2 trial. *Cancer Cell* **39**, 989–998.e985. <https://doi.org/10.1016/j.ccell.2021.05.009> (2021).
78. Ramagopalan, S. V., Pisoni, R., Rathore, L. S., Ray, J. & Sammon, C. Association of pertuzumab, trastuzumab, and docetaxel combination therapy with overall survival in patients with metastatic breast cancer. *JAMA Netw. Open* **4**, e2027764. <https://doi.org/10.1001/jamanetworkopen.2020.27764> (2021).
79. Delaloge, S. *et al.* Paclitaxel plus bevacizumab or paclitaxel as first-line treatment for HER2-negative metastatic breast cancer in a multicenter national observational study. *Ann. Oncol.* **27**, 1725–1732. <https://doi.org/10.1093/annonc/mdw260> (2016).
80. Vaidya, J. S. *et al.* Rethinking neoadjuvant chemotherapy for breast cancer. *BMJ* **360**, j5913. <https://doi.org/10.1136/bmj.j5913> (2018).

Acknowledgements

We are grateful to all participants in the Department of Medical Oncology, Fourth Hospital of Hebei Medical University; The Fifth Ward of Medical Oncology, Shijiazhuang people's hospital; Breast center, Fourth Hospital of Hebei Medical University; and Department of Epidemiology and statistics, Hebei Medical University.

Author contributions

Y.L. designed and drafted the work; Z.B. and X.W. acquired the data; J.L. and Y.L. analyzed the data; X.Z. and Y.S. substantively revised the work. All authors read and approved the final manuscript.

Funding

This study was funded by Health Commission of Hebei Province, China; Grant ID: ZL20140060.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25059-8>.

Correspondence and requests for materials should be addressed to Y.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022