



OPEN ACCESS

EDITED BY

Nguyen Quoc Khanh Le,
Taipei Medical University, Taiwan

REVIEWED BY

Tao Liu,
Wuhan University, China
Yutian Zou,
Sun Yat-sen University Cancer Center
(SYSUCC), China

*CORRESPONDENCE

Yifeng Mai,
✉ fymaiyifeng@nbu.edu.cn
Kai Hong,
✉ hongkai0629@163.com

[†]These authors have contributed equally to this work

SPECIALTY SECTION

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

RECEIVED 07 September 2022

ACCEPTED 01 December 2022

PUBLISHED 04 January 2023

CITATION

Guo Y, Chen Q, Zhang Y, Cheng X,
Cen K, Dai Y, Mai Y and Hong K (2023),
Prognostic implication and
immunotherapy response prediction of
a ubiquitination-related gene signature
in breast cancer.
Front. Genet. 13:1038207.
doi: 10.3389/fgene.2022.1038207

COPYRIGHT

© 2023 Guo, Chen, Zhang, Cheng, Cen,
Dai, Mai and Hong. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Prognostic implication and immunotherapy response prediction of a ubiquitination-related gene signature in breast cancer

Yangyang Guo^{1,2†}, Qiaoqiao Chen^{3,4†}, Yingjue Zhang^{5†},
Xu Cheng^{6†}, Kenan Cen⁷, Ying Dai⁷, Yifeng Mai^{7*} and
Kai Hong^{1,2*}

¹Department of Thyroid and Breast Surgery, Ningbo First Hospital, Ningbo, China, ²Department of Thyroid and Breast Surgery, Ningbo Hospital of Zhejiang University, Ningbo, China, ³Reproductive Medicine Center, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China, ⁴Key Laboratory of Reproductive Dysfunction Management of Zhejiang Province Assisted Reproduction Unit, Department of Obstetrics and Gynecology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China, ⁵Department of Molecular Pathology, Division of Health Sciences, Graduate School of Medicine, Osaka University, Suita, Japan, ⁶Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou, China, ⁷The Affiliated Hospital of Medical School of Ningbo University, Ningbo, China

Breast cancer (BC) is one of the most common tumor types and has poor outcomes. In this study, a ubiquitination-related prognostic signature was constructed, and its association with immunotherapy response in BC was explored. A list of ubiquitination-related genes was obtained from the molecular signatures database, and a ubiquitination-related gene signature was obtained by least absolute shrinkage and selection operator Cox regression. The genes, *TCN1*, *DIRAS3*, and *IZUMO4*, had significant influence on BC outcomes. Patients were categorized into two clusters—a high-risk group with poor survival and a low-risk group with greater chances of controlling BC progression. Univariate and multivariate Cox regression analyses revealed that the risk signature was an independent prognostic factor for BC. Gene set enrichment analysis suggested that the high-risk group was enriched in cell cycle and DNA replication pathways. The risk score was positively linked to the tumor microenvironment and negatively correlated with the immunotherapy response. The IC50 values for rapamycin were higher in the low-risk group, whereas those for axitinib, AZD6244, erlotinib, GDC0941, GSK650394, GSK269962A, lapatinib, and PD0325901 were higher in the high-risk group. Therefore, the ubiquitination-related signature is considered a promising tool for predicting a BC patient's immunotherapy response.

Abbreviations: AUC, area under the curve; BC, breast cancer; DEG, differentially expressed genes; GSEA, gene set enrichment analysis; OS, overall survival; PCA, principal component analysis; ROC, receiver operating characteristic; TME, tumor microenvironment; t-SNE, t-distributed stochastic neighbor embedding; URG, ubiquitination-related gene.

KEYWORDS

breast cancer, ubiquitination, immunotherapy, LASSO, signature, tumor microenvironment

Introduction

Breast cancer (BC) is the predominant type of cancer in women. The incidence rate of BC has risen steadily over the past decades, affecting nearly 100,000 patients worldwide annually (Britt et al., 2020; Sung et al., 2021). Currently, BC is commonly treated with chemotherapy either alone or in combination with other treatments. However, resistance to the available BC chemotherapeutic drugs is growing. The TNM staging system is widely used for assessing drug responsiveness, but it accounts for only a portion of the survival variation in BC patients (Edge et al., 2019). High-coverage and next-generation sequencing technologies are very efficient and sensitive, making it possible to compare whole-genome sequences of selected patients. Therefore, a panel of robust genetic biomarkers would be beneficial for improving the accuracy in estimating the prognosis of BC patients.

Ubiquitination, the addition of a ubiquitin molecule to the substrate, is an important function that involves a series of adaptive mechanisms among strongly invasive, proliferating cancer cells (Morrow et al., 2015). Ubiquitination regulates a large repertoire of cellular functions and requires the presence of an E3 ligase, which transfers ubiquitin (Ub) to substrates. The control system determining the specificity of the process is closely associated with carcinogenesis and tumor progression (Sun, 2006). Moreover, the ubiquitylation function may be directed towards degradation by the proteasome (degradative ubiquitylation) or towards altering the function (regulatory ubiquitylation) (Bergmann, 2010). Ubiquitination could be considered an essential hallmark of cancer; hence, it has been extensively investigated in previous studies using statistical tests and machine learning to integrate transcriptomics and ubiquitination into the underlying molecular mechanisms, with significant benefits (Popovic et al., 2014; Ge et al., 2018). Aberrant ubiquitination could underlie some of the heterogeneity of lung cancer, and drugs that target ubiquitin could offer effective new cancer treatments (Che et al., 2022). Thus, it is important to identify robust tumor-associated ubiquitination biomarkers, which could not only improve BC diagnosis and prognosis but also help develop novel therapeutic strategies.

To investigate the impact of genetic alteration of ubiquitination on BC, we explored the influence of ubiquitination-related genes (URGs) on BC, identified new prognostic groups, and constructed a predictive risk signature based on URG function.

Methods

Data profiles

Raw transcription statistics (fragments per kilobase of exon model per million reads mapped [FPKM] and count data) and corresponding clinical information (age, sex, pathological stage, TNM stage, survival time, and survival position) were downloaded from the Cancer Genome Atlas (TCGA, <https://www.cancer.gov>, BRCA project). The log₂ transformation was used to normalize the TCGA-BRCA cohort. We downloaded two microarray datasets (GSE20685 and GSE25066) (Kao et al., 2011; Baldasici et al., 2022) from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) and used them to validate the cohorts. Batch effects among the TCGA-BRCA and GEO datasets were eliminated by using the “ComBat” method in the *sva* package from R. The BC models of Tang et al., Wu et al., and Hu et al. were used to test and demonstrate the advantages of our ubiquitination-related prognostic signature (Tang et al., 2020; Wu et al., 2020; Hu et al., 2022).

Consensus clustering analysis of ubiquitination-related genes

Seventy-nine ubiquitination-related genes (URGs) were downloaded from the MSigDB database (<http://www.broad.mit.edu/gsea/msigdb/>), including gene symbol, official full name, ensembl ID, and gene type (Supplementary File S1). The R package ConsensusClusterPlus was used to conduct an agreement-unsupervised clustering analysis to sort BCs into distinct clusters according to the expression of URGs (Supplementary File S2) (Wilkerson and Hayes, 2010).

Relationship between molecular subtypes, clinical features, and BC prognosis

To examine the medical value of the two subtypes identified by harmony gathering, we compared the relationships between the different molecular subtypes, clinicopathological characteristics, and prognoses. Patient characteristics included sex, age, and TNM stage. Differences in overall survival (OS) between different subtypes were also evaluated using Kaplan–Meier curves generated by the “survminer” and “survival” R packages (Rich et al., 2010).

Tumor microenvironment of ubiquitination-related clusters

Immune cell infiltration was established using the MCPcounter algorithm to evaluate the tumor microenvironment (TME) of ubiquitination-related clusters (Becht et al., 2016). Differences in immune and stromal cell types were identified using the limma algorithm and presented with ‘violin’ plots (Ritchie et al., 2015). The data are shown using the R packages, limma and MCPcounter.

Construction and authentication of a predictive signature for BC

We normalized raw mRNA expression data from the TCGA-BRCA project and, based on the differentially expressed genes (DEGs) among ubiquitination-related clusters, we identified the downstream genes influenced by URGs using the R limma package ($p < 0.05$ and $|\log_2 FC| > 1$ denoted statistical significance) (Ritchie et al., 2015). Univariate Cox regression analyses were conducted to classify prognostic DEGs ($p < 0.05$). Then, least absolute shrinkage and selection operator (LASSO) regression analysis was performed to classify the prognostic DEGs and build a predictive signature for BC (Tibshirani, 1997).

$$\text{prognosis index (PI)} = \sum_{i=1}^n \text{Coef}(i) * \text{Expr}(i).$$

BC patients from TCGA-BRCA were divided into low- and high-risk subgroups based on the median TCGA-BRCA risk score. The whole TCGA-BRCA dataset was set as a training subset, then half the TCGA-BRCA patients were arbitrarily selected as the examination subset. The R survival package was used to compare the survival between the two groups using the Kaplan–Meier plotter in the training and test subsections. In addition, receiver operating characteristic (ROC) curves with area under the curve (AUC) standards for 1-, 3-, and 5-year survival were used to evaluate the prognostic ability of the ubiquitination-related signature for BC (Hajian-Tilaki, 2013). The survival status of BC patients in the TCGA-BRCA dataset is presented. Principal component analysis (PCA) and *t*-distributed stochastic neighbor embedding (*t*-SNE) analysis were used to evaluate the separation of low- and high-risk BC (Ringnér, 2008; Belkina et al., 2019). To further address the prognostic attributes of the signature, univariate and multivariate Cox regression analyses were performed to determine the independent risk factors for BC, with covariates such as the risk score, age, and TNM stages.

Establishment and verification of the nomogram

To improve the prognostic value of the signature, ubiquitination-related genes and new characteristics were used

to build a nomogram with the “rms” and “regplot” R packages (Iasonos et al., 2008). The nomogram was used to predict the 1-, 3-, and 5-year survival rates of patients, and calibration curves were used to assess the accuracy of the nomogram (Austin et al., 2020). ROC and DCA analyses were performed to evaluate the stability of the prognostic nomogram (Van Calster et al., 2018).

DEG identification and functional annotation

To explore the potential cellular functions and enriched pathways of genes downstream of URGs, functional enrichment analyses, including GO and KEGG, were performed on the DEGs using the “clusterprofiler” R package (Yu et al., 2012). Gene set enrichment analysis (GSEA) was conducted for differentially enriched pathways in low- and high-risk BC (Yu et al., 2012).

Subgroup analysis of low- and high-risk BC

To determine the correlation between the ubiquitination-related signature and the clinical characteristics, risk differences between low- and high-risk BC were assessed in subgroups by age, sex, clinical stage, tumor grade, and TNM stage. In addition, the survival differences of low- and high-risk BC in different subgroups were analyzed using Kaplan–Meier survival analysis. The results are presented with significance set at $p < 0.05$.

TME and immunotherapy response analysis of low- and high-risk BC

The correlation between ubiquitination-related signatures and immunoregulatory genes, immune checkpoint genes, and various types of immune cells was evaluated using Pearson’s coefficient. A heatmap of immune cell infiltration in high- and low-risk BC was generated using Timer, Xcell, Quantiseq, MCPcounter, EPIC, and Cibersort procedures (Becht et al., 2016; Aran et al., 2017; Li et al., 2017; Chen et al., 2018; Plattner et al., 2020; Racle and Gfeller, 2020). Expression of the immunotherapy targets CD47 and CTLA4 was determined, and TIDE and MSI scores were calculated to predict the immunotherapy response. Potential anti-PD-1 treatment efficiency was evaluated using the IMvigor 210 cohort (Balar et al., 2017).

Drug sensitivity evaluation of low- and high-risk BC

The R package “pRRophetic” was used to predict drug sensitivity in high- and low-risk BC patients (Geeleher et al., 2014). The data were extracted from the GDSC database (<https://>

www.cancerrxgene.org/), and the half-maximal inhibitory concentration (IC50) was used as the index.

Cells and cell culture

Normal breast epithelial cells (MCF-10A) and cancer cells (BT-549) were acquired from the American Type Culture Collection (Manassas, VA, United States). The cells were incubated with RPMI-1640 medium (Gibco, Carlsbad, CA, United States), containing 10% fetal bovine serum (Gibco, Carlsbad, CA, United States) and 1% penicillin-streptomycin (Gibco, Carlsbad, CA, United States). All cells were cultured at 37°C with 5% CO₂.

RT-qPCR

Total RNA was extracted from MCF-10 A and BT-549 cells using TRIzol reagent (Invitrogen), and reverse-transcribed into cDNA using PrimeScript[®] RT master mix (Perfect Real Time, Takara Bio, Shiga, Japan). RT-qPCR was performed using the qPCR master mix (Promega, Madison, WI, United States). Reactions were performed in triplicate, and Ct values were normalized to the endogenous housekeeping gene, *β-actin*, using the 2^{-ΔΔCT} method.

The primer sequences were as follows: 5'-GATTACCGC GTCGTGGTAGTC-3' (forward) and 5'-TCAATGGTCGGC AGGTACTCA-3' (reverse) for *DIRAS3*; 5'-GGAGCGACG ACACGATGAAG-3' (forward) and 5'-CAGCTCGTTGGG GAAATACCC-3' (reverse) for *IZUMO4*; 5'-CCCCTAGTG GGGCTCTTACT-3' (forward) and 5'-CAGAGGTTTTAG GCGGATGTAG-3' (reverse) for *TCN1*; and 5'-TGACGTGGA CATCCGCAAAG-3' (forward) and 5'-CTGGAAGGTGGA CAGCGAGG-3' (reverse) for *β-actin*.

Statistical analysis

We used the R software (version 4.0.3) to perform all calculations and produce the statistical data. A $p < 0.05$ was considered statistically significant.

Results

Data

The expression ranks of 79 URGs, obtained from the Molecular Signatures Database (MSigDB) in the GSEA

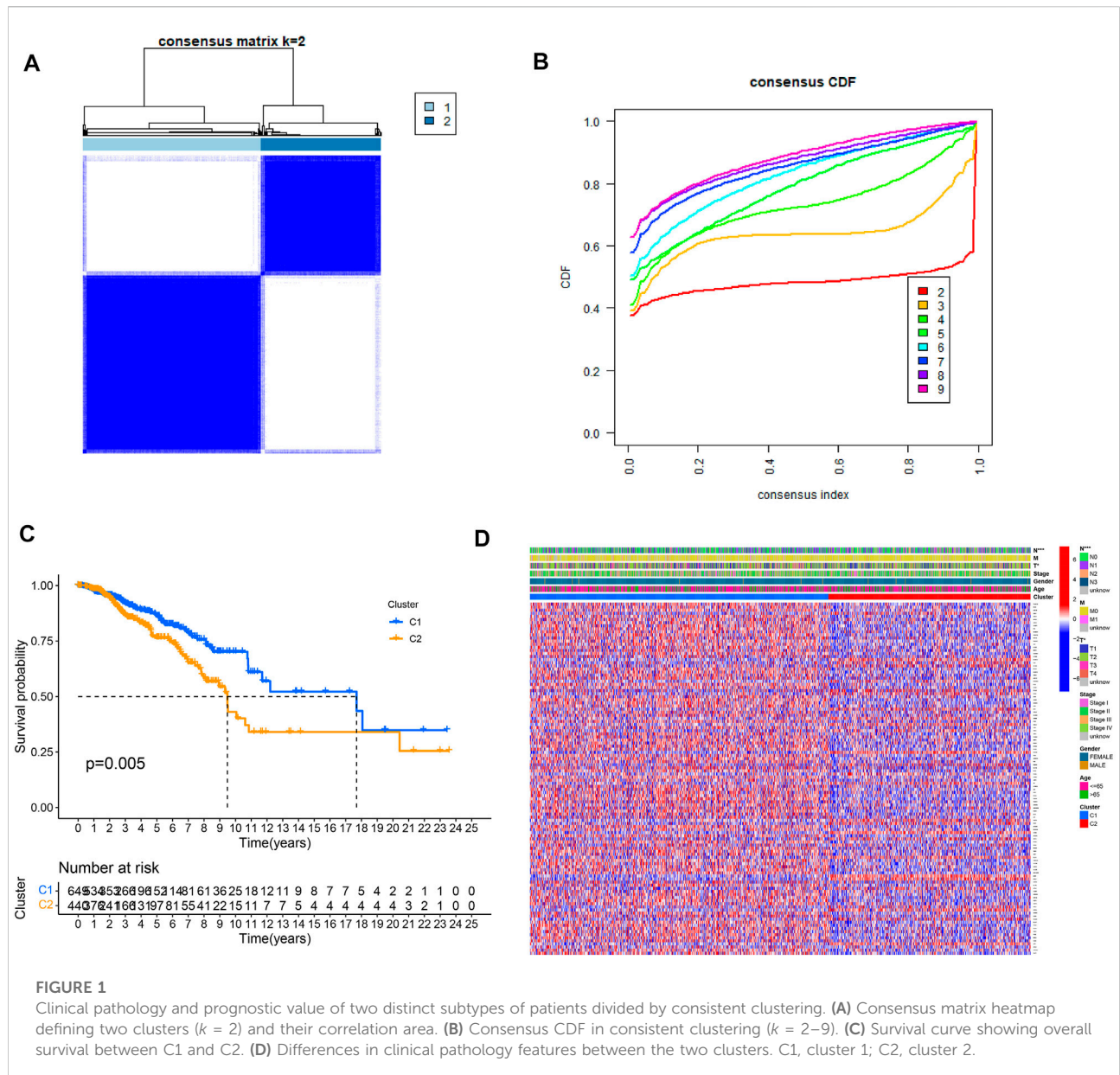
database, were measured in BC and normal breast samples from TCGA database using the Wilcoxon signed-rank test (Supplementary File S1).

Identification of ubiquitination-related clusters in BC

To understand how URGs are involved in tumorigenesis, we used an agreement clustering algorithm to classify BC based on the mRNA expression of 79 URGs (Supplementary File S2). In our dataset, $k = 2$ was considered the best choice for organizing the cohort into clusters 1 and 2 (Figure 1A). The consensus CDF of consistent clustering ($k = 2-9$) is shown in Figure 1B. The Kaplan–Meier curve revealed longer survival in cluster 1 BC than in cluster 2 BC ($p = 0.005$; Figure 1C). In addition, the clinicopathological topographies of the different BC subtypes exhibited important differences in URG expression and clinicopathological characteristics (Figure 1D). As shown in Figure 1D, cluster 1 was preferentially related to a lower N ($p < 0.001$) and T stage ($p < 0.05$) than cluster 2. The MCPcounter study showed that immune cell infiltration, including T-cell neutrophils, monocytic lineage cells, myeloid dendritic cells, fibroblasts, and endothelial cells, was significantly enriched in cluster 1 compared to that in cluster 2 (Figure 2).

Identification of gene clusters using DEGs

To determine the biological function of a ubiquitination pattern, we identified ubiquitination subtype-connected DEGs using the R package, limma. We examined the prognostic value of DEGs through BC progression using univariate Cox regression analysis to determine the potential relationship between the gene expression levels of the patient and two clusters in terms of survival. The results confirmed that 17 DEGs were associated with survival ($p < 0.01$) (Figure 3A). Thus, *PRAME*, *RAD54B*, *PXDNL*, *ACTL8*, and *JPH1* were considered high-risk genes ($HR > 1$), whereas *NEK10*, *TPRG1*, *PLD4*, *IGFALS*, *TCN1*, *CALML3*, *SPEF1*, *DIRAS3*, *SLC7A4*, *IZUMO4*, *VSIG2*, and *NPAS1* were considered protective genes ($HR < 1$). We then used LASSO Cox regression to classify the best risk score model for predicting survival in patients with BC (Figures 3B, C). The median risk score of patients with BC was identified as the threshold. As the training subset, all the patients with BC in TCGA-BRCA were divided into high- and low-risk groups. There was a significant difference in OS between the high- and low-risk groups according to the Kaplan–Meier curve ($p < 0.001$) (Figure 3D). Similarly, for the test set, the randomly selected patients (half of the total) with BC in TCGA-BRCA were divided into high- and low-risk groups according to the median risk score. Consistent



with the OS in the training set, a high survival probability was observed in the low-risk group in the test set (Figure 3E).

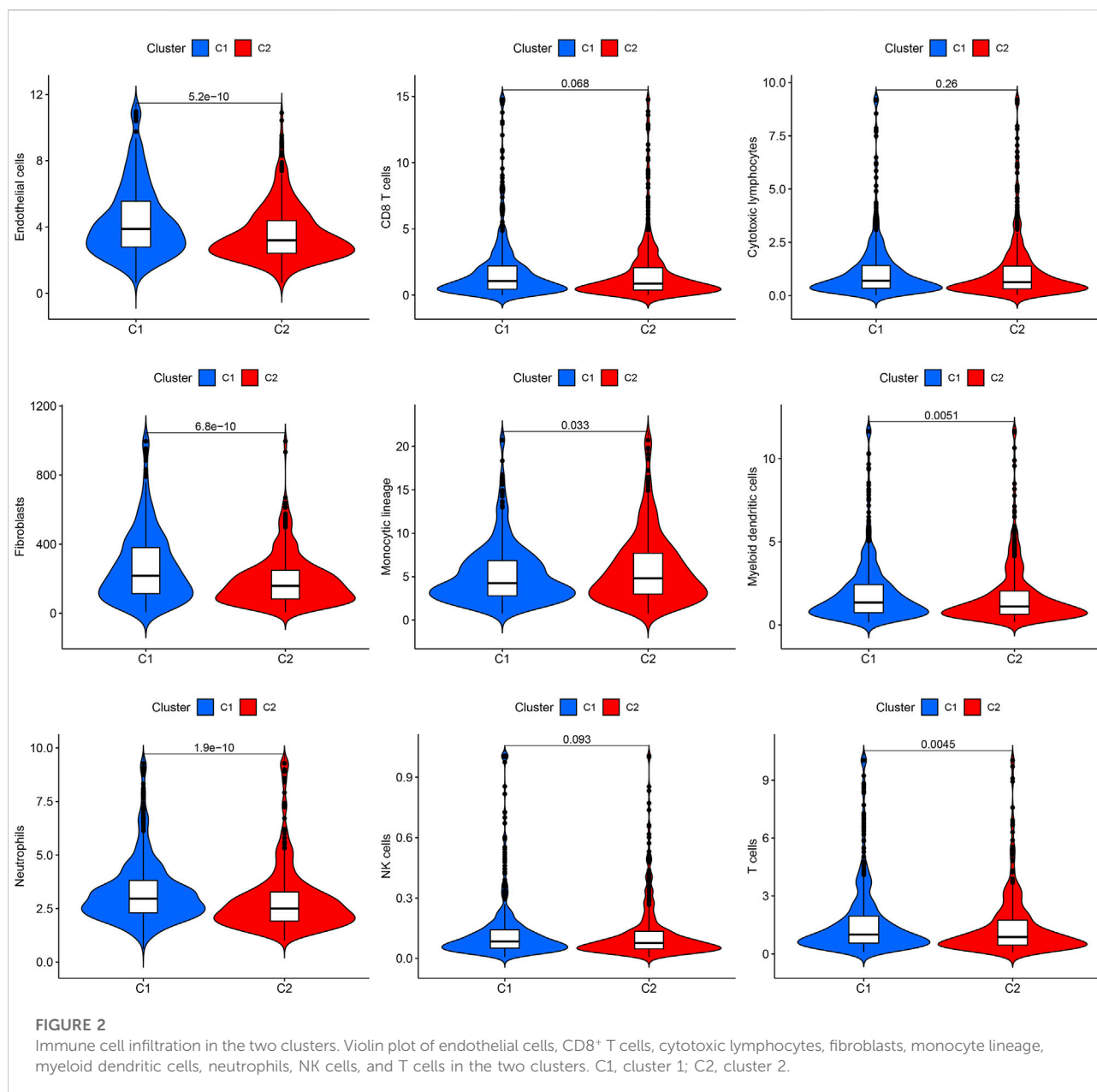
ROC curve analysis of the BC prognosis risk scoring model

To determine whether the BC signature was predictive, ROC curves were generated. ROC curve analysis revealed a significant prognostic effect for BC based on the risk model group, the AUC values for 1-, 3-, and 5-year survival being 0.632, 0.628, and 0.614 in the training set and 0.643, 0.668, and 0.645 in the test set, respectively (Figures 4A, B). Our survival and risk status maps

showed that the high-risk groups had more deaths than the low-risk groups (Figures 4C,D). In addition, an important difference was observed between the two clusters in terms of ubiquitination-related transcription profiles through PCA and *t*-SNE analysis (Figures 4E,F).

Independent validation of the signature's stability and advantages

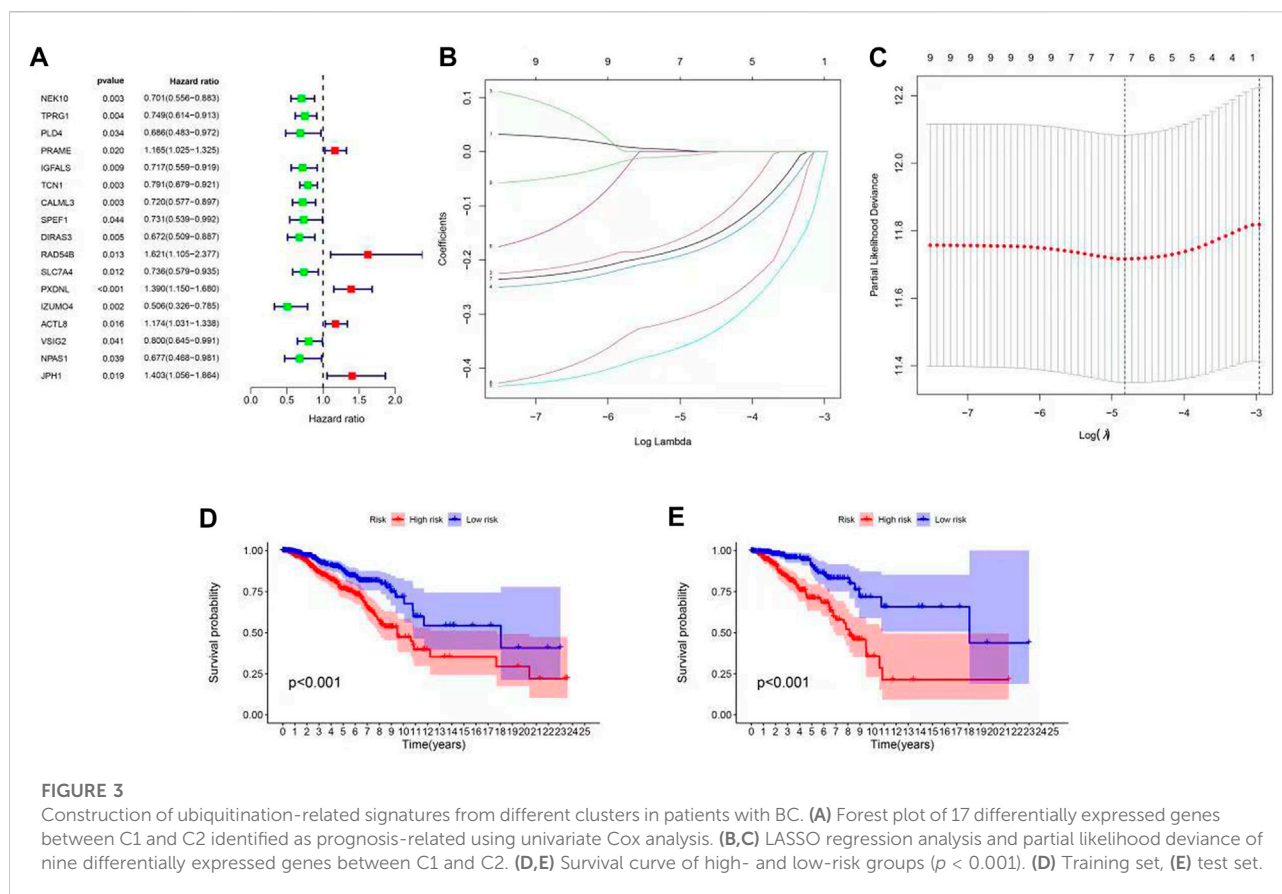
To further corroborate the advantages and stability of our ubiquitination-related signature, three breast cancer models were selected for comparison. The models of Tang et al., Wu



et al., and Hu et al. showed a lower C-index than our signature (0.509, 0.598, and 0.645-0.646) (Supplementary Figure S1). Consequently, we determined the K-M survival and performed ROC analyses on the GSE20685 and GSE25066 datasets. The results confirmed that the low- and high-risk groups had a significantly different prognosis in GSE20685 and GSE25066 ($p = 0.006$; $p = 0.037$). The two ROC analyses showed acceptable stability for 1-, 3-, and 5-year OS (0.687, 0.647, and 0.636; 0.641, 0.656, and 0.632) (Supplementary Figure S1).

Structure and validation of the analytical nomogram

The risk scores were identified as independent analytical indicators using multivariate and univariate Cox regression analyses (Figures 5A–D). Based on several clinical parameters and risk scores, separate numerical probabilities for OS were generated using a prognostic nomogram (Figure 5E). Through calibration curves, the nomogram predicted 1-, 3-, and 5-year OS well compared to the ideal model (Figure 5F). ROC curves were



used to evaluate the prognostic accuracy of the nomogram and the other signatures. A high prognostic value was indicated by an area under the ROC curve of 0.888 (Figure 5G). Compared to risk, age, and stage, the nomogram produced a greater net benefit in forecasting OS according to DCA (Figure 5H). Furthermore, the nomogram demonstrated a better net benefit than age and tumor stage, indicating its reliability and predictability.

Clinical utility and functional enrichment analysis of the predictive signature

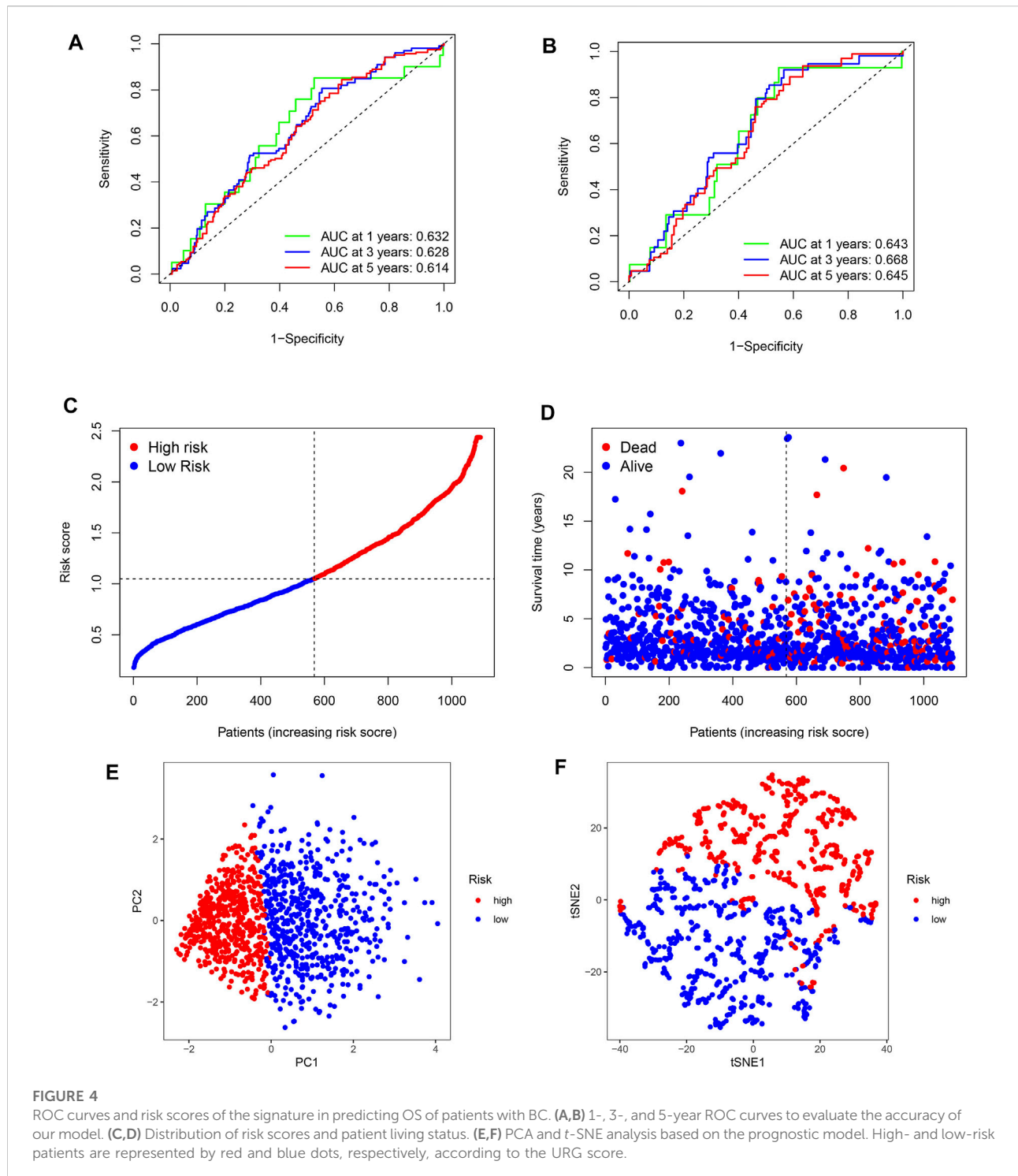
We also examined the association between ubiquitination-related predictive factors and medical variables. A heatmap was constructed to display the expression levels of the three signature URGs in the high- and low-risk groups (Figure 6A). The low-risk group had high expression levels of *TCN1*, *DIRAS3*, and *IZUMO4*. Genes associated with ubiquitination-related prognostic signatures were significantly enriched in biological processes related to homeostasis (Figure 6B). KEGG analysis revealed pathways related to immune function and cancer (Figure 6C), suggesting that ubiquitination plays a pivotal role in the immune regulation of the TME. To further explore this mechanism, GSEA plots showed the presence of metabolic, cell

cycle, cell communication, and DNA replication pathways in the high- and low-risk groups (Figures 6D, E).

The function of the important risk scores in disease progression was assessed by evaluating the relationship between the risk score in the predictive signature and clinicopathological features. High-risk scores were associated with advanced M1 stages, T stages, and high-grade tumors, suggesting a strong correlation between risk scores and low prognosis in BC. High-risk scores were seen in the most advanced clinicopathological stages, T4, stage IV, and M1 (Figure 7A). Next, we performed survival examinations stratified by age, TNM stage, and clinical stage to better estimate the survival consequences and determine the broad applicability of the predictive signature (Figure 7B). Patients in the high-risk group had significantly shorter OS than those in the low-risk group for cases with N0 ($p = 0.044$), N1+3 ($p = 0.001$), clinical stages I and II ($p = 0.001$), clinical stages III and IV ($p = 0.024$), T1+2 ($p = 0.001$), T3+4 ($p = 0.015$), age ≤ 65 ($p < 0.001$), and M0 ($p < 0.001$) (Figure 7B).

Calculation of TME and checkpoints between the high- and low-risk groups

We assessed the correlation between risk scores and the abundance of immune cells. The association of biomarkers of



risk scores and immune cells is illustrated in [Figures 8A, B](#) Forty immune checkpoint genes showed a significant correlation with risk score. As shown in [Figures 8C, D](#), the risk scores were positively correlated with the monocytic lineage and negatively correlated with cytotoxic lymphocytes, the B lineage, neutrophils,

endothelial cells, and fibroblasts. The association between the proposed model and the numbers of immune cells was investigated using the Xcell, Cibersort, Cibersort-ABS, EPIC, MCPcounter, TIMER, and QuantiSeq algorithms, which showed similar results ([Figure 8E](#)). CD47 and CTLA4 were

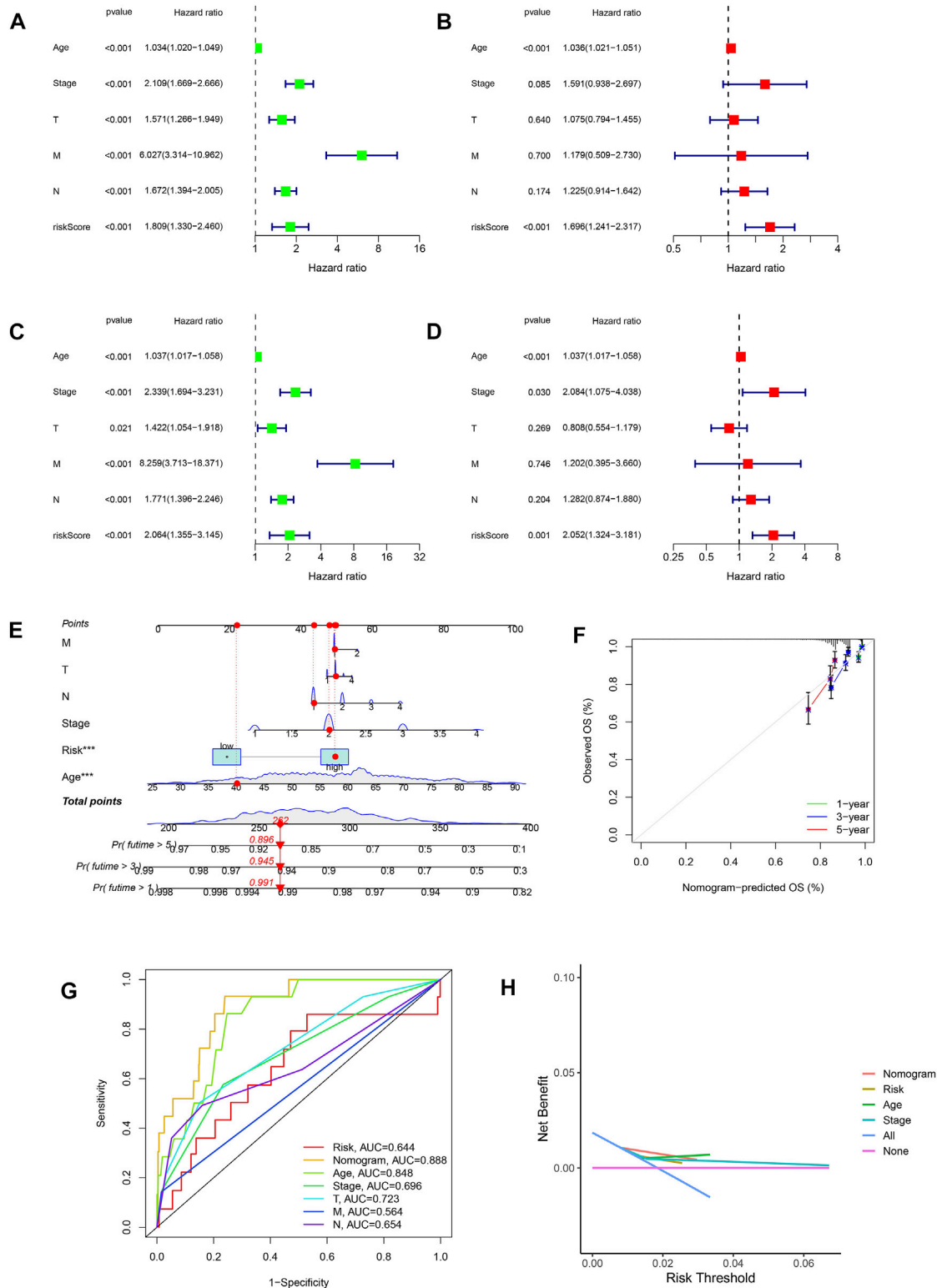


FIGURE 5 Relationship between risk scores and clinical features. (A,B) Univariate and multivariate Cox analyses considering risk score, age, tumor stage, and TNM stage in the training cohort. (C,D) Univariate and multivariate Cox analyses considering risk score, age, tumor stage, and TNM stage in the test cohort. (E) Nomogram combining risk score and clinicopathological factors. (F) Calibration plots established to compare the proposed nomogram with an ideal model. (G) AUCs showing that this nomogram had higher accuracy in OS than other factors. (H) Net benefit of the proposed nomogram, risk group, age, and tumor stage.

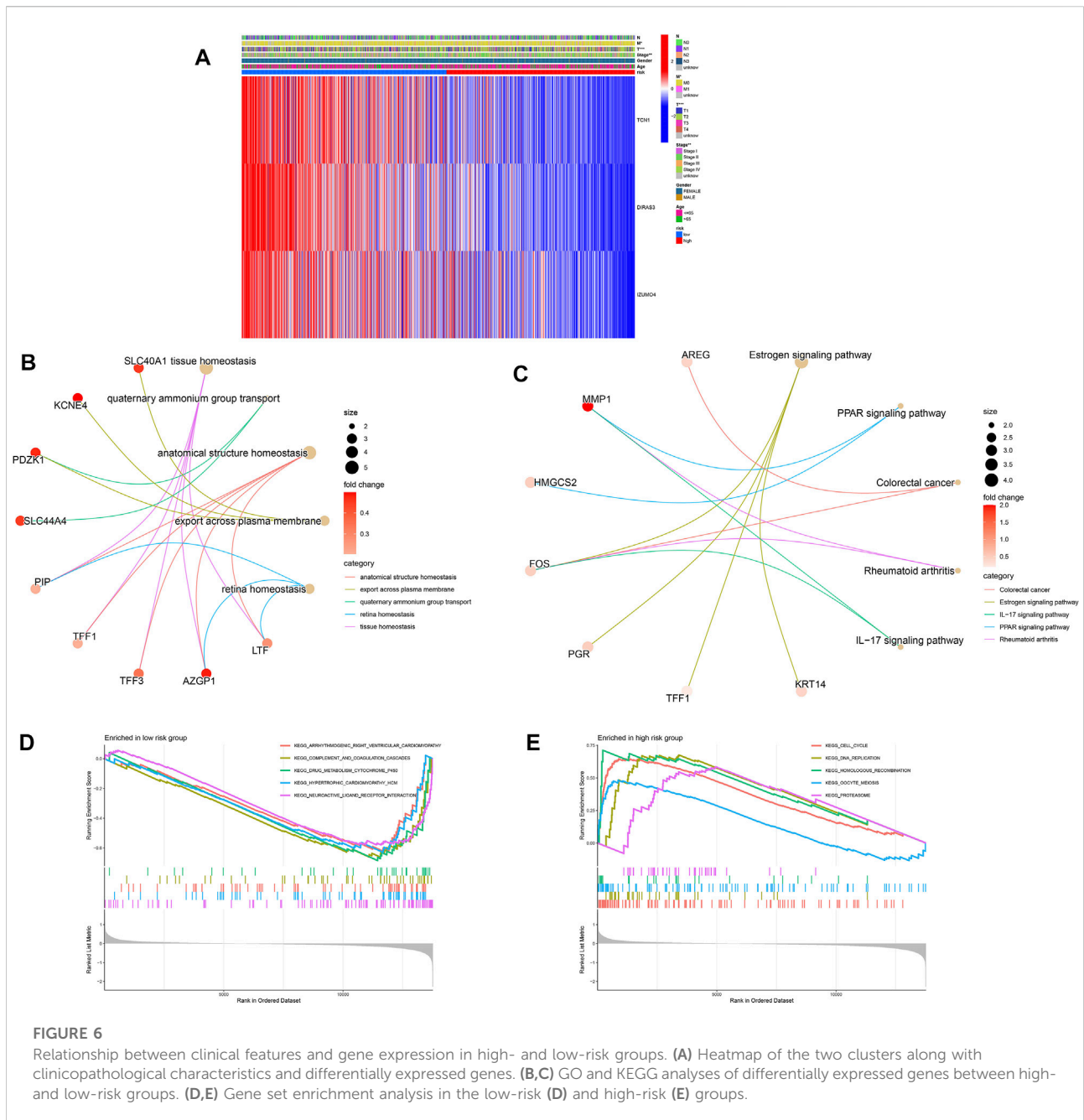
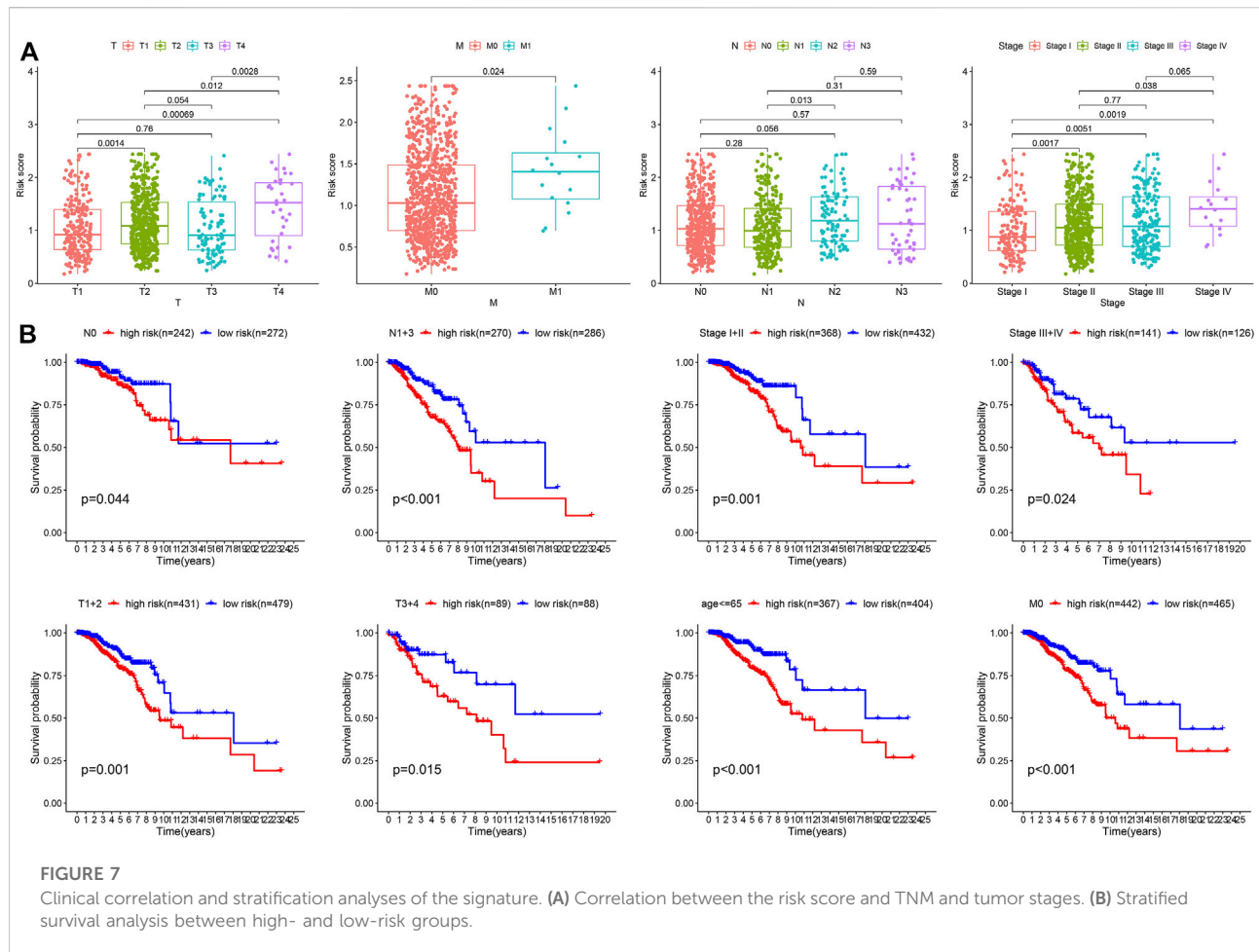


FIGURE 6 Relationship between clinical features and gene expression in high- and low-risk groups. **(A)** Heatmap of the two clusters along with clinicopathological characteristics and differentially expressed genes. **(B,C)** GO and KEGG analyses of differentially expressed genes between high- and low-risk groups. **(D,E)** Gene set enrichment analysis in the low-risk **(D)** and high-risk **(E)** groups.

highly expressed in the high-risk group compared with the low-risk group, implying that CD47 and CTLA4 expression strongly correlated with higher risk (Figures 9A, B). A low-risk score was also associated with high TIDE, dysfunction, exclusion, TAM M2, and CAF scores, whereas a high-risk score was associated with a high MDSC score (Figures 9C–I). Groups with low- and high-risk samples were significantly correlated with immunotherapy responses; high-risk BC patients showed better therapy responses (Figure 9J).

Drug sensitivity analysis of high- and low-risk groups

We evaluated the sensitivities of patients in the low- and high-risk groups to chemotherapy drugs currently used to treat BC. Rapamycin had a low IC50 value in patients in the high URG score group, whereas axitinib, AZD6244, erlotinib, GDC0941, GSK.650394, GSK269962A, lapatinib, and PD0325901 had low IC50 values in patients in the low URG score group. The results



confirmed the validity of the ubiquitination-related signature in predicting drug sensitivity (Figure 10).

RT-qPCR

Three genes that constituted the ubiquitination-related signature were confirmed as prognostic biomarkers for BC (Figure 11A). Low expression of *DIRAS3*, *IZUMO4*, and *TCN1* correlated with poor BC prognosis. RT-qPCR was used to measure transcription levels and compare the mRNA expression of *DIRAS3*, *IZUMO4*, and *TCN1* in BC. The results demonstrated that *DIRAS3*, *IZUMO4*, and *TCN1* had a lower expression in tumor cells than in normal breast epithelial cells (Figure 11B). The present analysis shows that *DIRAS3*, *IZUMO4*, and *TCN1* are suppressor oncogenes.

Discussion

BC is characterized by high heterogeneity, which may affect its prognosis (Roulot et al., 2016). Early and small BC with

complicated clinical features makes diagnosis and treatment difficult (Unger-Saldaña, 2014). At present, the major challenges in treating BC are primary diagnosis, accurate forecasting of tumor progression, and current treatment. Therefore, it is essential to examine the biological features of the disease and identify specific biomarkers that could enhance the accuracy of prognosis prediction, inform treatment personalization, and improve survival rates. With the development of next-generation sequencing technology, we can effectively study the related genetic characteristics and determine their effect on risk (Behjati and Tarpey, 2013). This will facilitate selection of the optimal treatment for individual patients and help them achieve their treatment goals. To improve the outcomes of patients with BC, it will be necessary to conduct clinical trials to evaluate the benefits of incorporation of preclinical results and application of molecular-guided treatment. URGs are regulators of tumor cell cycle, gene expression, and progression (Deng et al., 2020). However, there is little understanding of the role of protein ubiquitination in the BC microenvironment. Thus, URG in the TME needs to be studied further.

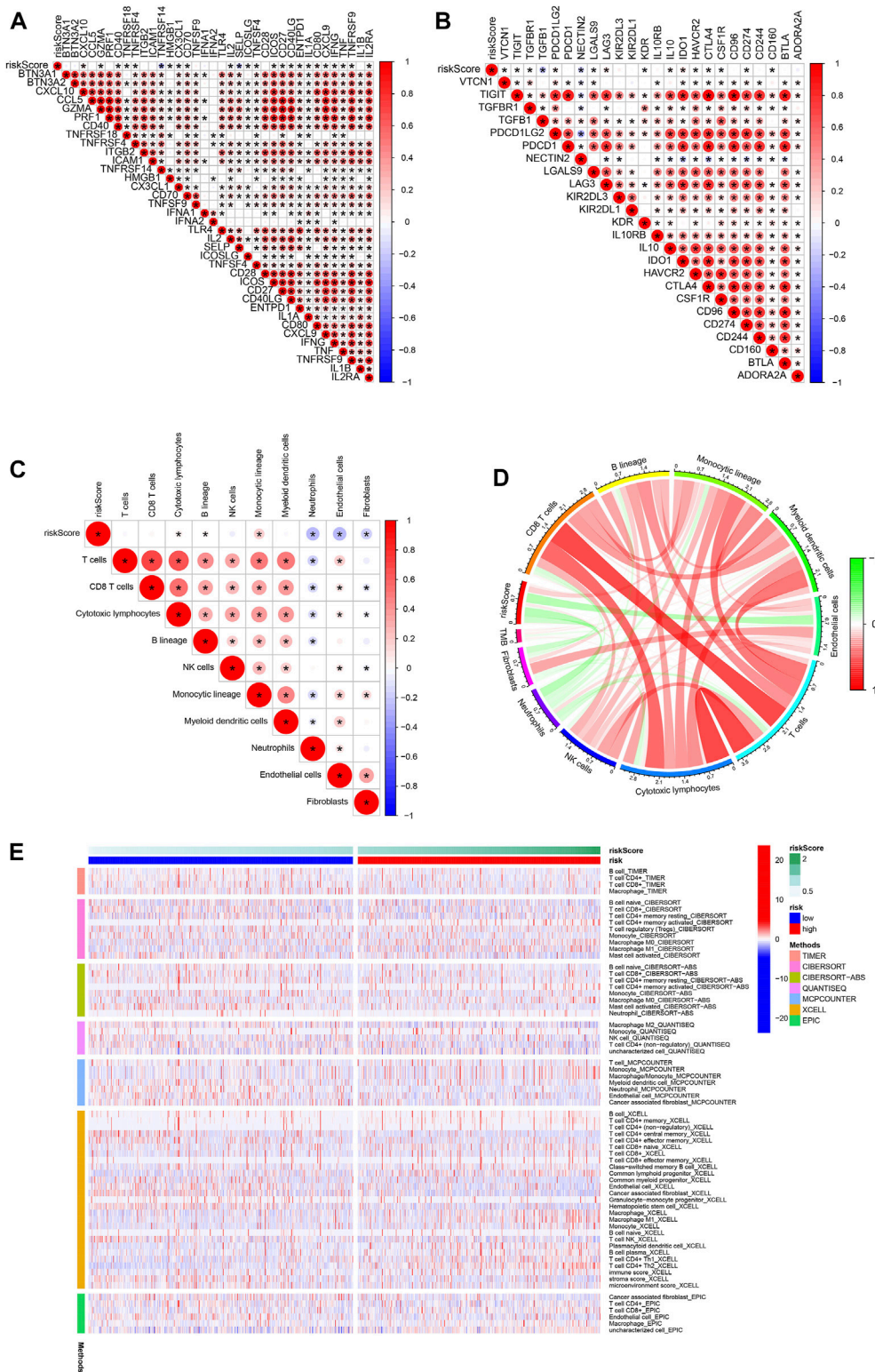
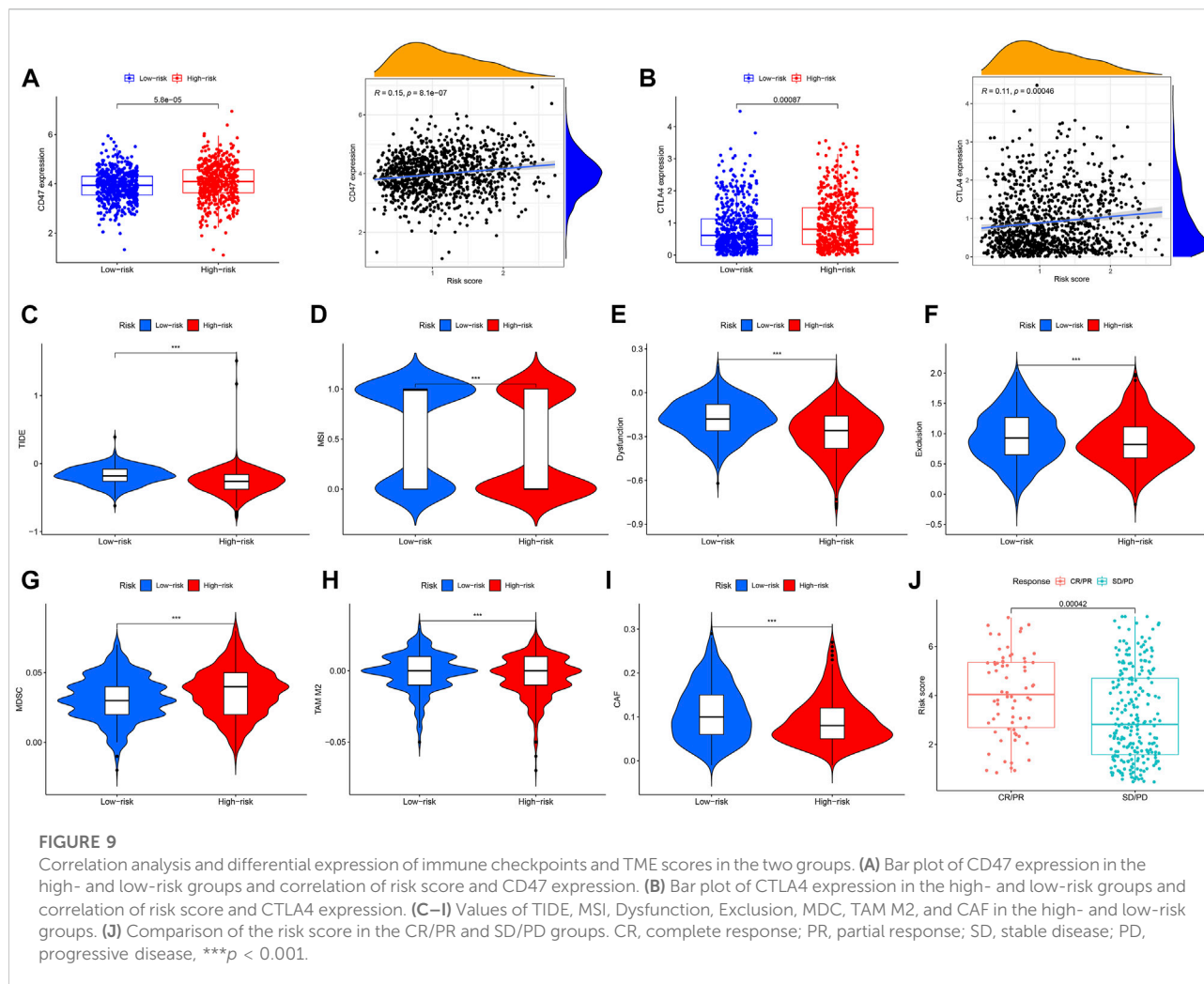


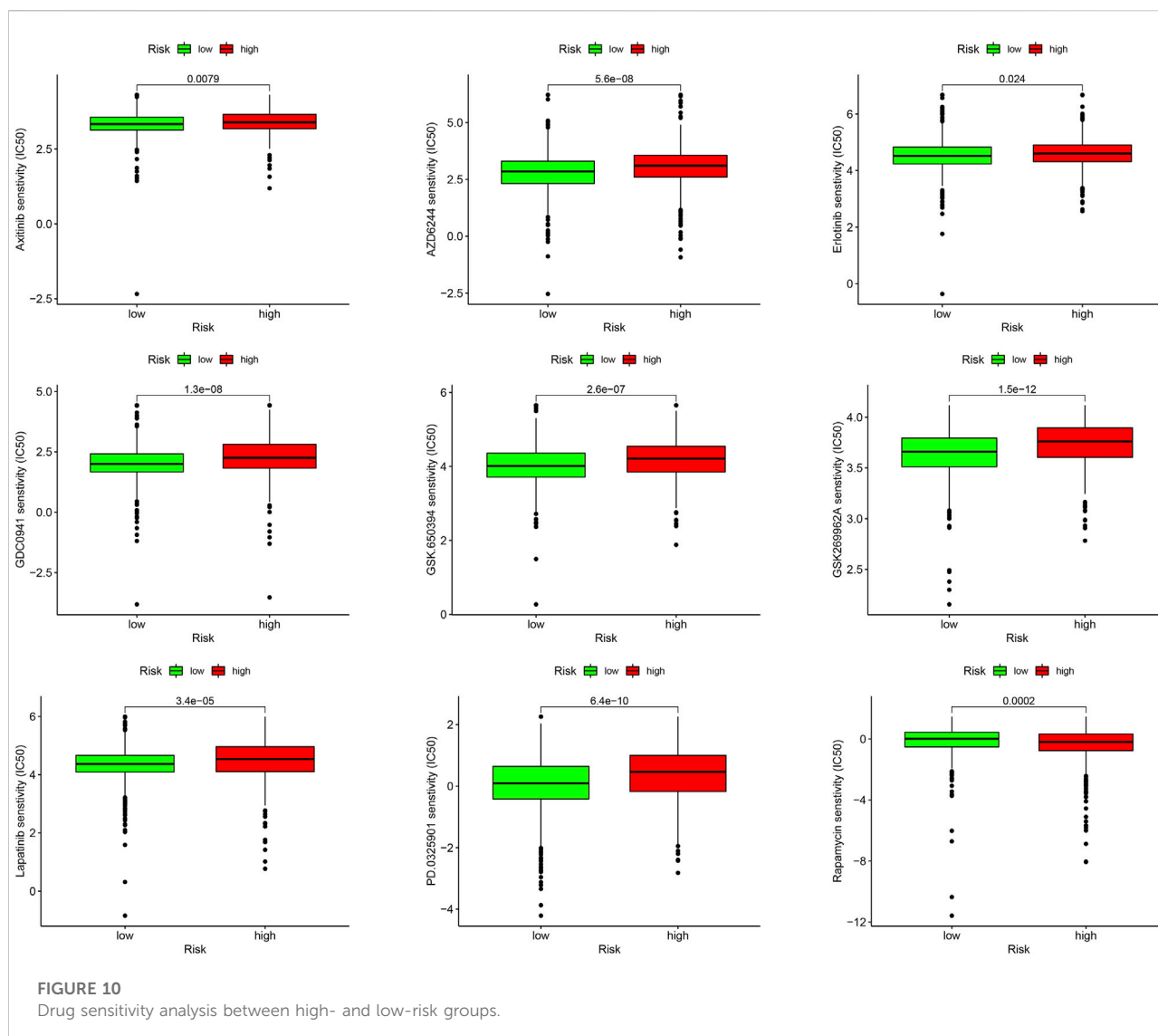
FIGURE 8 Association of immune signatures and immune cell infiltration with risk score. (A,B) Correlation of immune signatures and risk score. (C,D) Correlation of immune cell infiltration and risk score. (E) Clustering pattern of immune cell type in the high- and low-risk groups.



Ubiquitination is a post-translational modification that plays a major role in pathological and physiological processes (Dougherty et al., 2020). Abnormally expressed genes have been investigated in several studies. *DERL1* acts as a vital regulator via interaction with *UBE2C* during oncogenic activities in BC (Zeng et al., 2020). One study uncovered pivotal functions of Derlin-1 in the regulation of ER stress-induced preventive control and ubiquitination by the HRD1 E3 ubiquitin ligase prior to deprivation, suggesting an important role of *DERL1* in homeostasis (Kadowaki et al., 2018). The present study showed that *DERL1* was significantly upregulated in BC and corresponded with a poor prognosis. Furthermore, we found that *UBA1*, *UBE2A*, *UBE2C*, *UBE2G1*, *UBE2T*, and *WAC* were significantly positively associated with poor BC prognosis.

In our study, two ubiquitination clusters were constructed and verified to improve BC outcome prediction. Cluster 1 was associated with better BC prediction than cluster 2. Higher N and T stages were observed in cluster 2 than in cluster 1, indicating

that the cluster 2 group may be associated with a higher tumor stage of BC. Comparing immune cell infiltration in the two clusters, we found that cluster 1 had the most infiltration of endothelial cells, T cells, neutrophils, and myeloid dendritic cells. The robust ubiquitination-related gene signature was then verified by LASSO Cox regression analysis. Patients with a low-risk score had better OS than those with high-risk scores; 1-, 3-, and 5-year ROC curves confirmed the positive performance of the signature. Additional stratified existence analyses of various medical subgroups also established the signature's robust predictive power. Multivariate and univariate Cox analyses revealed that the signature had independent predictive value. Using a nomogram model, we improved the efficacy of the analytical signature by considering age, risk scores, tumor stage, and TNM. This important model showed excellent accuracy in predicting the survival rates of BC patients. There were significant differences in immunocyte penetration and immunologic function between the low- and high-risk groups, and TME scores, immune checkpoints, and

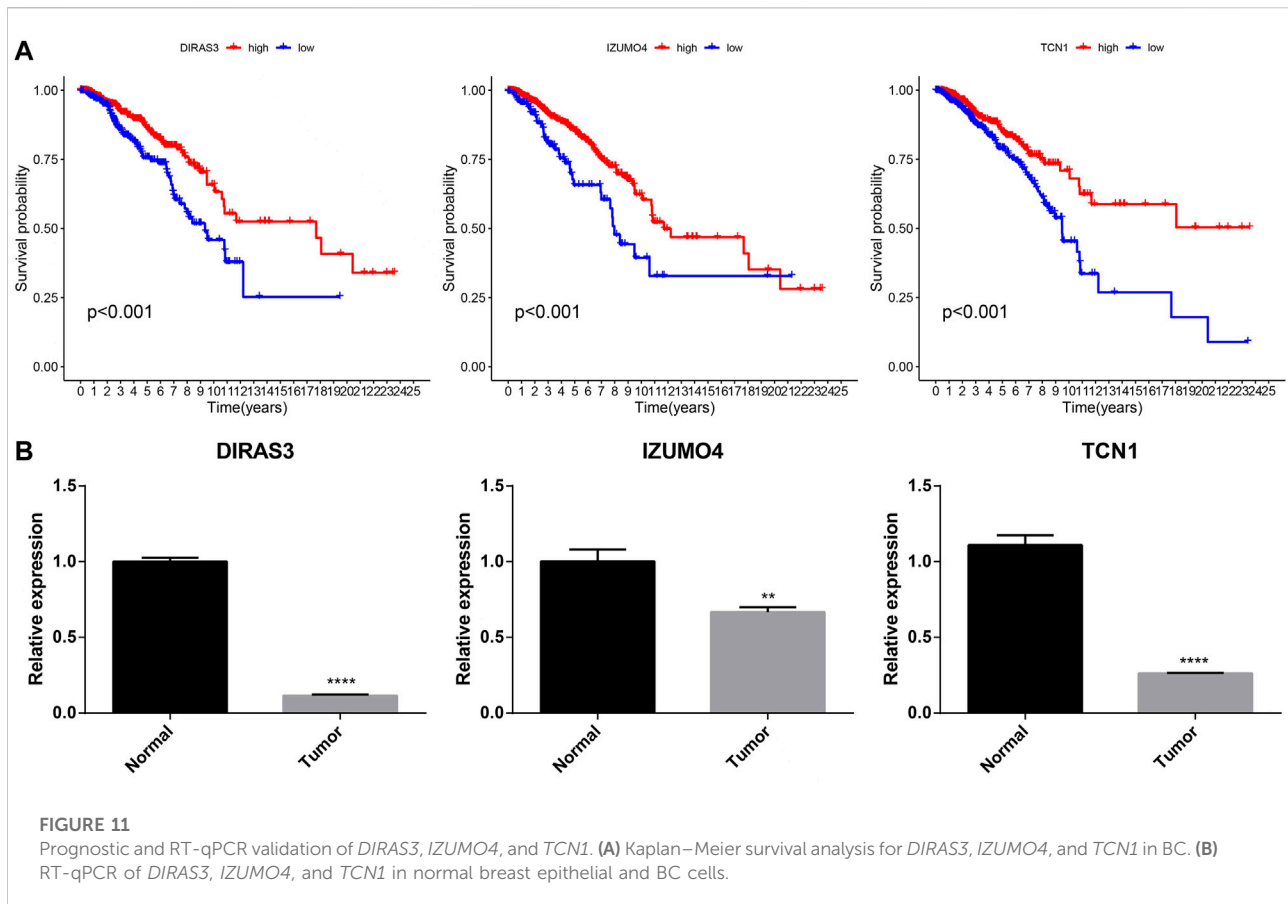


drug susceptibilities differed significantly between the two groups.

In another experiment, we identified three specific genes associated with the prognostic signature: *TCN1*, *DIRAS3*, and *IZUMO4*. A prognostic model for lung adenocarcinoma based on eight genes (*TCN1*, *COL1A1*, *SPOK2*, *PCP4*, *S100P*, *CAV2*, *GPX3*, and *ASPM*) has been previously constructed (Tu et al., 2021). Additionally, bioinformatics analyses have confirmed that the mRNA expression of *TCN1* was upregulated in common colon cancer (Liu et al., 2020); both Yang's study and our model identified this gene. The expression of *TCN1* was significantly higher in the low-risk/high-survival group, than in the high-risk group. Low *TCN1* expression was associated with malignancy in BC, suggesting that it should be further evaluated for its specific predictive value. The expression levels of two other genes, *DIRAS3* and *IZUMO4*, were significantly higher in the low-

risk group than in the high-risk group and were also useful in prognosis.

There are diverse models for breast cancer, based on metabolic heterogeneity, stromal immune phenotype, diverse cell-death patterns, cell necroptosis, plasmalogen deficiency, overactive fatty acid elongation biomarkers, and 4-mRNA metastasis-related genes (Xie et al., 2018; Zheng et al., 2020; Tomida et al., 2021; Yu et al., 2021; Xie et al., 2022; Zou et al., 2022). Yu et al. demonstrated that energy-related metabolic features in BC were related to glycolytic activity and survival. The clustering reflected intertumoral metabolic heterogeneity and could be used to personalize therapeutic strategies (Yu et al., 2021). In another study, 237 patients with triple-negative breast cancer (TNBC) from real-world cases and 533 patients with TNBC from public datasets were used to determine a stromal immune phenotype for TNBC. According to the density of



stromal CD4⁺ T cells, $\gamma\delta$ T cells, monocytes, M1 macrophages, and M2 macrophages, TNBC patients were divided into immune phenotypes, A and B, with different immune activities and prognosis (Zheng et al., 2020). A BC gene signature based on 27 metastasis-associated DEGs was also identified. Its prognostic ability was demonstrated and a nomogram was constructed based on mRNA signature, T stage, and N stage (Xie et al., 2018). Our signature revealed the heterogeneity of BC from the standpoint of ubiquitination, and we confirmed that it could be used as a tool to predict survival, immune activity, and therapy response.

A previous study acknowledged a risk model group of eight URGs that forecast risk for glioma (Wang et al., 2021). Several tumor-related signaling pathways were augmented in the high-risk group. TFs were also predicted to adjust the risk model for CD4 T cells and B cells; URGs and neutrophils were related to the risk model. In another study, FBXL6 played an important role in promoting hepatocellular carcinoma owing to the ubiquitination and stabilization of HSP90AA1, which contributed to tumorigenesis in hepatocellular carcinoma (Shi et al., 2020). Various enzymes are involved in ubiquitination, which can lead to cancer development; however, the underlying mechanisms need to be researched further (Wang et al., 2022). Ubiquitination is reversible, and de-ubiquitination can be

used in common tumor treatment. In a study on BC, the E3 ubiquitin ligase, CHFR, was used to create double-strand breaks (DSBs) through poly(ADP-ribose) or PAR. Additionally, ALC1 (amplification in liver cancer-1) plays an important role in metastasis and invasion of BC, and poly(ADP-ribosyl)ated PARP1 stimulated ALC1 at DNA damage sites, emphasizing that the PBZ sphere of CHFR and PMD and Macro spheres of ALC1 are essential for the PAR interface. The common ubiquitination of ALC1 through CHFR is dependent on PARylation and directed to the degradation of PARylated ALC1 (Wang et al., 2019).

Several cancers, including bladder cancer and glioma, exhibited immune cell infiltration associated with URGs, which can provide novel therapeutic targets. The common ubiquitination-related subtypes showed significant differences in immune cell infiltration, stromal scores, ESTIMATE scores, and immune scores (Cai et al., 2021). According to the investigation of the immune microenvironment in gliomas, this risk grouping could be used as a guide for common glioma immunotherapy. An association study between risk ratings and immune cells showed that many immune cells were closely associated with risk scores. Lastly, high-risk BC patients showed higher expression levels of immune checkpoint genes than low-risk BC patients (Tang et al., 2022).

Six types of immune cells, including endothelial cells, fibroblasts, cytotoxic lymphocytes, neutrophils, monocytic cells, and B cells, were differentially expressed in different risk groups. Among these cells, monocytic lineages, which had a low survival rate, were significantly more abundant in the high-risk group than in the low-risk group. There were differences in MSI score, TIDE score, and immune checkpoint expression between the high- and low-risk groups. High-risk BC was significantly associated with low immune cell infiltration, which could support the discovery of new therapeutic methods. Despite recent advances in tumor immunotherapy, BC has not achieved satisfactory therapeutic benefits, and the mechanism underlying chemoresistance remains unclear; therefore, the development of multimodal therapies and bio-integration targets is necessary. The association between URG and BC requires further investigation.

In summary, our study identified a three-URG signature associated with immune infiltration and drug sensitivity. We confirmed the analytical model using training and test sets. The results showed that this signature could be used as a novel biomarker in many fields and provide a personalized treatment strategy for BC. The mechanism underlying URGs is still unknown, however, and needs to be explored in future studies.

Data availability statement

The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Author contributions

KH, QC, and YZ drafted the manuscript and designed the study. XC, KC, and YD conducted statistical analyses. YM and

YG reviewed the manuscript. All the authors have read and approved the final manuscript.

Funding

This study was supported by the Ningbo University Institute of Geriatrics (LNBYJS-2021).

Acknowledgments

We thank all the members who contributed to this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1038207/full#supplementary-material>

References

- Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18, 220. doi:10.1186/s13059-017-1349-1
- Austin, P. C., Harrell, F. E., Jr., and van Klaveren, D. (2020). Graphical calibration curves and the integrated calibration index (ICI) for survival models. *Stat. Med.* 39, 2714–2742. doi:10.1002/sim.8570
- Balar, A. V., Galsky, M. D., Rosenberg, J. E., Powles, T., Petrylak, D. P., Bellmunt, J., et al. (2017). Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: A single-arm, multicentre, phase 2 trial. *Lancet* 389, 67–76. doi:10.1016/S0140-6736(16)32455-2
- Baldasici, O., Balacescu, L., Cruceriu, D., Roman, A., Lisencu, C., Fetica, B., et al. (2022). Circulating small EVs miRNAs as predictors of pathological response to neo-adjuvant therapy in breast cancer patients. *Int. J. Mol. Sci.* 23, 12625. doi:10.3390/ijms232012625
- Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 17, 218. doi:10.1186/s13059-016-1070-5
- Behjati, S., and Tarpey, P. S. (2013). What is next generation sequencing? *Arch. Dis. Child. Educ. Pract. Ed.* 98, 236–238. doi:10.1136/archdischild-2013-304340
- Belkina, A. C., Ciccolella, C. O., Anno, R., Halpert, R., Spidlen, J., and Snyder-Cappione, J. E. (2019). Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets. *Nat. Commun.* 10, 5415. doi:10.1038/s41467-019-13055-y
- Bergmann, A. (2010). The role of ubiquitylation for the control of cell death in *Drosophila*. *Cell Death Differ.* 17, 61–67. doi:10.1038/cdd.2009.70
- Britt, K. L., Cuzick, J., and Phillips, K. A. (2020). Key steps for effective breast cancer prevention. *Nat. Rev. Cancer* 20, 417–436. doi:10.1038/s41568-020-0266-x

- Cai, H., Chen, H., Huang, Q., Zhu, J. M., Ke, Z. B., Lin, Y. Z., et al. (2021). Ubiquitination-related molecular subtypes and a novel prognostic index for bladder cancer patients. *Pathol. Oncol. Res.* 27, 1609941. doi:10.3389/pore.2021.1609941
- Che, Y., Jiang, D., Xu, L., Sun, Y., Wu, Y., Liu, Y., et al. (2022). The clinical prediction value of the ubiquitination model reflecting the immune traits in LUAD. *Front. Immunol.* 13, 846402. doi:10.3389/fimmu.2022.846402
- Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12
- Deng, L., Meng, T., Chen, L., Wei, W., and Wang, P. (2020). The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduct. Target. Ther.* 5, 11. doi:10.1038/s41392-020-0107-0
- Dougherty, S. E., Maduka, A. O., Inada, T., and Silva, G. M. (2020). Expanding role of ubiquitin in translational control. *Int. J. Mol. Sci.* 21, 1151. doi:10.3390/ijms21031151
- Edge, S. B., Hortobagyi, G. N., and Giuliano, A. E. (2019). New and important changes in breast cancer TNM: Incorporation of biologic factors into staging. *Expert Rev. Anticancer Ther.* 19, 309–318. doi:10.1080/14737140.2019.1582335
- Ge, Z., Leighton, J. S., Wang, Y., Peng, X., Chen, Z., Chen, H., et al. (2018). Integrated genomic analysis of the ubiquitin pathway across cancer types. *Cell Rep.* 23, 213–226. e3. doi:10.1016/j.celrep.2018.03.047
- Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLOS ONE* 9, e107468. doi:10.1371/journal.pone.0107468
- Hajian-Tilaki, K. (2013). Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp. J. Intern. Med.* 4, 627–635.
- Hu, T., Zhao, X., Zhao, Y., Cheng, J., Xiong, J., and Lu, C. (2022). Identification and verification of necroptosis-related gene signature and associated regulatory Axis in breast cancer. *Front. Genet.* 13, 842218. doi:10.3389/fgene.2022.842218
- Iasonos, A., Schrag, D., Raj, G. V., and Panageas, K. S. (2008). How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* 26, 1364–1370. doi:10.1200/JCO.2007.12.9791
- Kadowaki, H., Satrimafitrah, P., Takami, Y., and Nishitoh, H. (2018). Molecular mechanism of ER stress-induced pre-emptive quality control involving association of the translocon, Derlin-1, and HRD1. *Sci. Rep.* 8, 7317. doi:10.1038/s41598-018-25724-x
- Kao, K., Chang, K., Hsu, H., and Huang, A. (2011). Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: Implications for treatment optimization. *BMC Cancer* 11, 143. doi:10.1186/1471-2407-11-143
- Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77, e108–e110. doi:10.1158/0008-5472.CAN-17-0307
- Liu, G. J., Wang, Y. J., Yue, M., Zhao, L. M., Guo, Y. D., Liu, Y. P., et al. (2020). High expression of TCN1 is a negative prognostic biomarker and can predict neoadjuvant chemosensitivity of colon cancer. *Sci. Rep.* 10, 11951. doi:10.1038/s41598-020-68150-8
- Morrow, J. K., Lin, H. K., Sun, S. C., and Zhang, S. (2015). Targeting ubiquitination for cancer therapies. *Future Med. Chem.* 7, 2333–2350. doi:10.4155/fmc.15.148
- Plattner, C., Finotello, F., and Rieder, D. (2020). Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. *Methods Enzymol.* 636, 261–285. doi:10.1016/bs.mie.2019.05.056
- Popovic, D., Vucic, D., and Dikic, I. (2014). Ubiquitination in disease pathogenesis and treatment. *Nat. Med.* 20, 1242–1253. doi:10.1038/nm.3739
- Racle, J., and Gfeller, D. (2020). Epic: A tool to estimate the proportions of different cell types from bulk gene expression data. *Methods Mol. Biol.* 2120, 233–248. doi:10.1007/978-1-0716-0327-7_17
- Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., and Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *Otolaryngol. Head. Neck Surg.* 143, 331–336. doi:10.1016/j.otohns.2010.05.007
- Ringnér, M. (2008). What is principal component analysis? *Nat. Biotechnol.* 26, 303–304. doi:10.1038/nbt0308-303
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007
- Roulot, A., Héquet, D., Guinebretière, J. M., Vincent-Salomon, A., Lerebours, F., Dubot, C., et al. (2016). Tumoral heterogeneity of breast cancer. *Ann. Biol. Clin.* 74, 653–660. English. doi:10.1684/abc.2016.1192
- Shi, W., Feng, L., Dong, S., Ning, Z., Hua, Y., Liu, L., et al. (2020). FBXL6 governs c-MYC to promote hepatocellular carcinoma through ubiquitination and stabilization of HSP90AA1. *Cell Commun. Signal.* 18, 100. doi:10.1186/s12964-020-00604-y
- Sun, Y. (2006). E3 ubiquitin ligases as cancer targets and biomarkers. *Neoplasia* 8, 645–654. doi:10.1593/neo.06376
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660
- Tang, J., Luo, Y., and Wu, G. (2020). A glycolysis-related gene expression signature in predicting recurrence of breast cancer. *Aging (Albany NY)* 12, 24983–24994. doi:10.18632/aging.103806
- Tang, Q., Chen, Z., Xie, J., Mo, C., Lu, J., Zhang, Q., et al. (2022). Transcriptome analysis and single-cell sequencing analysis constructed the ubiquitination-related signature in glioma and identified USP4 as a novel biomarker. *Front. Immunol.* 13, 915709. doi:10.3389/fimmu.2022.915709
- Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Stat. Med.* 16, 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3
- Tomida, S., Goodenowe, D., Koyama, T., Ozaki, E., Kuriyama, N., Morita, M., et al. (2021). Plasmalogen deficiency and overactive fatty acid elongation biomarkers in serum of breast cancer patients pre- and post-surgery-new insights on diagnosis, risk assessment, and disease mechanisms. *Cancers* 13, 4170. doi:10.3390/cancers13164170
- Tu, Z., He, X., Zeng, L., Meng, D., Zhuang, R., Zhao, J., et al. (2021). Exploration of prognostic biomarkers for lung adenocarcinoma through bioinformatics analysis. *Front. Genet.* 12, 647521. doi:10.3389/fgene.2021.647521
- Unger-Saldaña, K. (2014). Challenges to the early diagnosis and treatment of breast cancer in developing countries. *World J. Clin. Oncol.* 5, 465–477. doi:10.5306/wjco.v5.i3.465
- Van Calster, B., Wynants, L., Verbeek, J. F. M., Verbakel, J. Y., Christodoulou, E., Vickers, A. J., et al. (2018). Reporting and interpreting decision curve analysis: A guide for investigators. *Eur. Urol.* 74, 796–804. doi:10.1016/j.eururo.2018.08.038
- Wang, K., Liu, J., Li, Y. L., Li, J. P., and Zhang, R. (2022). Ubiquitination/de-ubiquitination: A promising therapeutic target for pten reactivation in cancer. *Biochim. Biophys. Acta. Rev. Cancer* 1877, 188723. doi:10.1016/j.bbcan.2022.188723
- Wang, L., Liu, Y., Xuan, C., Liu, Y., Shi, H., and Gao, Y. (2021). Identification of ubiquitination-related genes in human glioma as indicators of patient prognosis. *PLOS ONE* 16, e0250239. doi:10.1371/journal.pone.0250239
- Wang, Y., Wang, Y., Wang, N., You, L., Long, F., Shao, C., et al. (2019). Poly(ADP-ribose) polymerase 1/2 inhibitors decrease the ubiquitination of ALC1 mediated by CHFR in breast cancer. *Oncol. Rep.* 42, 1467–1474. doi:10.3892/or.2019.7242
- Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26, 1572–1573. doi:10.1093/bioinformatics/btq170
- Wu, F., Chen, W., Kang, X., Jin, L., Bai, J., Zhang, H., et al. (2020). A seven-nuclear receptor-based prognostic signature in breast cancer. *Clin. Transl. Oncol.* 23, 1292–1303. doi:10.1007/s12094-020-02517-1
- Xie, J., Tian, W., Tang, Y., Zou, Y., Zheng, S., Wu, L., et al. (2022). Establishment of a cell necroptosis index to predict prognosis and drug sensitivity for patients with triple-negative breast cancer. *Front. Mol. Biosci.* 9, 834593. doi:10.3389/fmolb.2022.834593
- Xie, X., Wang, J., Shi, D., Zou, Y., Xiong, Z., Li, X., et al. (2018). Identification of a 4-mRNA metastasis-related prognostic signature for patients with breast cancer. *J. Cell. Mol. Med.* 23, 1439–1447. doi:10.1111/jcmm.14049
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic* 16, 284–287. doi:10.1089/omi.2011.0118
- Yu, T., Ma, D., Liu, Y., Xiao, Y., Gong, Y., Jiang, Y., et al. (2021). Bulk and single-cell transcriptome profiling reveal the metabolic heterogeneity in human breast cancers. *Mol. Ther.* 29, 2350–2365. doi:10.1016/j.yjmt.2021.03.003
- Zeng, J., Tian, Q., Zeng, Z., Cai, J., Ye, M., Liu, Y., et al. (2020). Derlin-1 exhibits oncogenic activities and indicates an unfavorable prognosis in breast cancer. *Cell Biol. Int.* 44, 593–602. doi:10.1002/cbin.11259
- Zheng, S., Zou, Y., Xie, X., Liang, J., Yang, A., Yu, K., et al. (2020). Development and validation of a stromal immune phenotype classifier for predicting immune activity and prognosis in triple-negative breast cancer. *Int. J. Cancer* 147, 542–553. doi:10.1002/ijc.33009
- Zou, Y., Xie, J., Zheng, S., Liu, W., Tang, Y., Tian, W., et al. (2022). Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. *Int. J. Surg.* 107, 106936. doi:10.1016/j.ijsu.2022.106936