

Identification of Breast Cancer Subtypes by Integrating Genomic Analysis with the Immune Microenvironment

Ran Ding, Qiwei Liu, Jing Yu, Yongkang Wang, Honglei Gao, Hongxing Kan,* and Yinfeng Yang*

Cite This: *ACS Omega* 2023, 8, 12217–12231

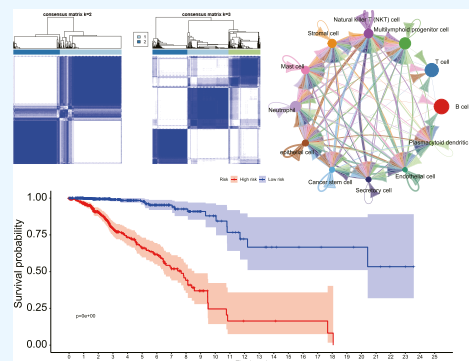
Read Online

ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: *Objectives:* We aim to identify the breast cancer (BC) subtype clusters and the crucial gene classifier prognostic signatures by integrating genomic analysis with the tumor immune microenvironment (TME). *Methods:* Data sets of BC were derived from the Cancer Genome Atlas (TCGA), METABRIC, and Gene Expression Omnibus (GEO) databases. Unsupervised consensus clustering was carried out to obtain the subtype clusters of BC patients. Weighted gene coexpression network analysis (WGCNA), least absolute shrinkage and selection operator (LASSO), and univariate and multivariate regression analysis were employed to obtain the gene classifier signatures and their biological functions, which were validated by the BC dataset from the METABRIC database. Additionally, to evaluate the overall survival rates of BC patients, Kaplan–Meier survival analysis was carried out. Moreover, to assess how BC subtype clusters are related to the TME, single-cell analysis was performed. Finally, the drug sensitivity and the immune cell infiltration for different phenotypes of BC patients were also calculated by the CIBERSORT and ESTIMATE algorithms. *Results:* TCGA–BC samples were divided into three subtype clusters, S1, S2, and S3, among which the prognosis of S2 was poor and that of S1 and S3 were better. Three key pathways and 10 crucial prognostic-related gene signatures are screened. Finally, single-cell analysis suggests that S1 samples have the most types of immune cells, S2 with more sensitivity to tumor treatment drugs are enriched with more neutrophils, and more multilymphoid progenitor cells are involved in subtype cluster S3. *Conclusions:* Our novelty was to identify the BC subtype clusters and the gene classifier signatures employing a large-amount dataset combined with multiple bioinformatics methods. All of the results provide a basis for clinical precision treatment of BC.



1. INTRODUCTION

Breast cancer (BC), a heterogeneous tumor, has many etiologies and clinical features.^{1,2} Studies have found that multiple subtypes have been identified in the development and progression of BC.³ Based on the 2013 St. Gallen International Breast Cancer Congress, more than five molecular subtypes of BC can be defined, *i.e.*, luminal A, luminal B, human epidermal growth factor receptor 2 (HER2) + B2, HER2 overexpression, basal-like triple-negative breast cancer (TNBC), and other subtypes.⁴ The classical classification methods of breast cancer subtypes include immunohistochemical-based estrogen receptor (ER)-, progesterone receptor (PR)-, and HER2 and intrinsic gene expression-based classification methods.⁵ Correspondingly, the resulting tumor subtypes with distinct genomic alterations show different clinical outcomes and treatment responses. Identification of BC subtypes can provide insight into the pathogenesis of the disease and contribute to the precise diagnosis of BC and personalized targeted therapy.

Clinical phenotypes such as disease stage, metastasis, and tumor resectability, as well as histological subtypes, are frequently used in current clinical approaches for outcome prediction and treatment decision-making for BC treatment. However, the development of molecular profiling more recently

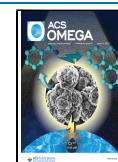
has given rise to the possibility of quantitative tumor analysis based on profiles of protein expression, mutations, and/or genome-wide gene transcription. These assays promise to characterize tumor subtypes with more precision and accuracy and to more correctly forecast how particular tumor types will respond to various therapies. For instance, Perou et al. divided BC into five intrinsic molecular subtypes based on gene expression signatures—basal-like, normal-like, HER2-enriched, and luminal A and B, revealing variations in the genesis of tumor cells as well as disparate patterns of progression.⁶ Another study classified BC into 1q/16q, amplifier, and complex subtypes based on the CNA pattern.⁷

Additionally, it is widely known that the cross-talk between the tumor cells and tumor microenvironment (TME) influences genetic and epigenetic changes that affect the formation,

Received: December 27, 2022

Accepted: March 14, 2023

Published: March 21, 2023



incidence, and therapeutic resistance of cancer. In BC patients, a better clinical result has been related to immune infiltration levels.^{8,9} Particularly, higher CD8+ T-cell infiltration is strongly associated with better overall survival (OS) in ER-negative patients.^{10,11} Moreover, high immune infiltration is additionally associated with a better response to adjuvant chemotherapy.¹² Many studies in recent years have shown that the TME can be explained by transcriptomic data.^{13–18} These findings indicate that a lower risk of breast cancer recurrence is correlated with a higher expression of leukocyte-related genes.^{13,16,19,20} Notably, Ali et al. and Bense et al. recently reported in a metastudy how specific immune cell types affect breast cancer prognosis.^{13,21} However, the role and clinical relevance of immunity in BC still need to be elucidated by more comprehensive analysis.

Presently, three BC prognostic subtype clusters and a set of genetic signatures that can be used to classify BC subtypes were identified using BC data in public databases, which may provide effective evidence for precise clinical treatment of BC and help explore BC heterogeneity between patients. The single-cell sequence analysis of BC suggests progressive immune infiltration of clinically relevant immune clusters. Through characterization of the immune components of the TME, the immune infiltration and poor prognosis group identified as tumor-related. Further subtype analysis revealed that most types of immune cells are enriched in subtype cluster S1 samples, more neutrophils are distributed in the subtype S2 with more sensitivity to tumor treatment drugs, and more multilymphoid progenitor cells are involved in subtype cluster S3. All of the results may provide an effective basis for the clinical precision treatment of BC.

2. MATERIALS AND METHODS

2.1. Data Source. Presently, a total of three publicly available data sets, *i.e.*, cohorts A, B, and C, are used in this study, which are respectively derived from the Cancer Genome Atlas (TCGA), Molecular Taxonomy of Breast Cancer International Consortium (METABRIC), and Gene Expression Omnibus (GEO) databases. Among them, the gene expression profiling data of cohort A obtained from TCGA was used as a training set for model construction, which contains 103 paracancerous samples and 903 BC samples.²² The gene expression data of cohort B derived from METABRIC, which included 154 paracancerous samples and 1826 BC samples, were used to assess the reliability of the constructed model.²³ Further, to explore the underlying biology of BC, single-cell analysis was performed on cohort C (accession number GSE118389), which includes six primary TNBC patients with gene expression profile data.²⁴

2.2. Clustering Analysis for Determining the BC Subtypes. To determine the molecular subtypes of BC in patients, the consensus clustering analysis was carried out by employing the ConsensusClusterPlus package of R.²⁵ Before running the program, the “max cluster number (maxK),” “cluster algorithm (clusterAlg),” and “distance” parameters in this package were set to “5,” “hc,” and “Pearson,” respectively. In order to ensure the stability and reliability of the classification, all cluster analyses were repeated 1000 times. Additionally, by calculating the consensus matrix, cumulative distribution function (CDF), and relative change in the area under the CDF curve, the optimal number of subtypes (K) was obtained.

2.3. Construction of the Coexpression Network. To identify the modules associated with clinical features, weighted gene coexpression network analysis (WGCNA)²⁶ was carried

out on BC cohort A data by the corresponding package of R. At first, a topological overlap matrix (TOM) was obtained by the adjacency matrix. Then, all genes were split into various gene modules by the TOM dissimilarity measure. The outliers were identified and removed by clustering analysis and the optimal value of soft threshold β was employed to ensure a scale-free network. When minModuleSize = 30 and merge CutHeight = 0.25, the gene module was identified as a key module. Finally, the gene significance (GS) and module membership (MM) of the genes in the module were calculated, which represent the correlation between genes and subtypes. Herein, the hub genes are the genes with GS > 0.5 and MM > 0.9.

2.4. Functional Enrichment Analysis of Genes in Key Modules. To investigate the potential biological functions performed by the acquired genes in key modules, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were carried out by employing several R packages, which contain “org. Hs.Eg.Db,” “ggplot2,” and “clusterProfiler.”^{27,28} Subsequently, the enriched GO or pathway terms with significant differences were screened by a threshold *p*-value < 0.05.

2.5. BC Subtype-Related Gene Classification Signature Screening. To screen the key genes classification signature associated with BC subtypes, we first built the gene classifier as previously described.²⁹ In brief, employing the logistic regression, the candidate genes associated with BC subtype distinction (with *p* < 0.05) were identified. Finally, using the glmnet and decision tree packages³⁰ in R, the LASSO regression was carried out to narrow the variables and obtain the final gene classification signature.

2.6. Establishment of the Risk Prognostic Model. To investigate the impact of gene classification signatures on the patient’s survival outcome and survival time, the BC gene expression data of cohort A were also used to establish a clinical prediction model. Employing the gene classifiers obtained above, the risk prognostic model was built based on the following formula.

$$\text{risk score} = \sum_{i=1}^n \text{coef}_i \times \text{id}_i \quad (1)$$

where *coef_i* denotes the coefficient and *id_i* denotes the normalized count for each gene.

Additionally, the TCGA–BC patients were split into high- and low-risk groups based on the median risk score. The overall survival rates of two patient groups were assessed using Kaplan–Meier survival analysis. Additionally, employing the survival package³¹ in R, the ability of the classifier to predict the tumor lymph node metastasis (TNM) status, stage, age, and risk was assessed by the time-dependent receiver operating characteristic (ROC) curve. Moreover, area under the curve (AUC) values were computed to determine the classifier’s or pertinent clinical factors’ predictive potential. Finally, the correlation between TCGA-based breast cancer clinical immunohistochemical subtypes and previous high- and low-risk subtypes was evaluated and visualized using the ggstatsplot package.³²

2.7. Validation of the Classification Model. To further confirm the robustness of gene classified signatures, the data of cohort B derived from METABRIC are employed for external validation. By calculating the individual risk score of high- and low-risk BC patients, their individual survival status was evaluated. In addition, the molecular intrinsic subtypes of BC patients were determined by the PAM50 classifier algorithm.³³

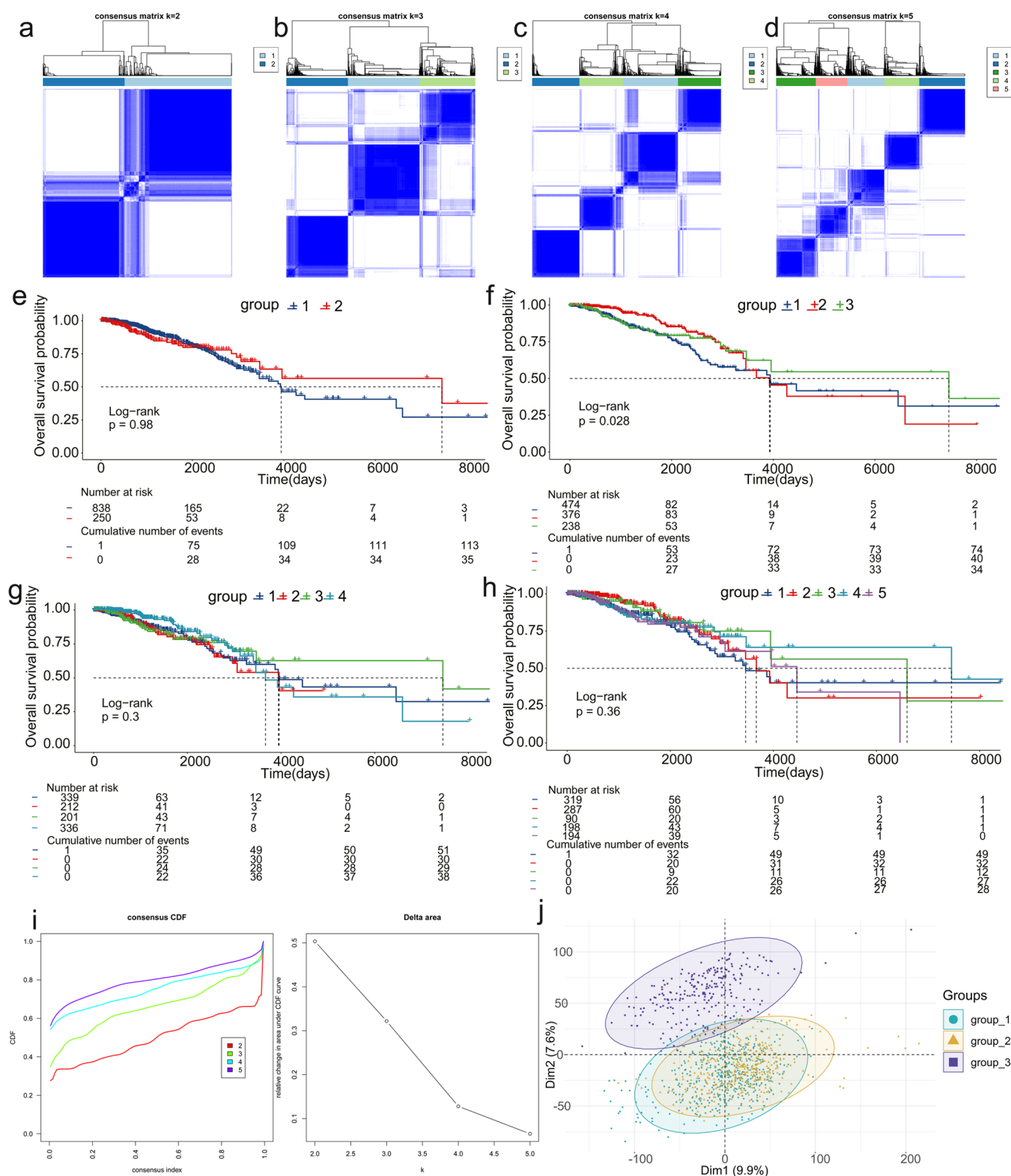


Figure 1. Consensus clustering for TCGA-BC. (a–d) Consistency matrix heat map at $k = 2$ –5. (e–h) Kaplan–Meier curves for OS of BC patients stratified by consensus clustering (2–5). (i) Cumulative distribution function curve tracking graph. (j) PCA of expression when subtype classification number is 3.

Also, to compare the overall survival among subtypes, log-rank tests were carried out to assess the correlation between the basic classification clusters and tumor stage.

2.8. Evaluation Heterogeneity among Different Subtype Clusters. To further evaluate the heterogeneity of the

molecular intrinsic subtypes of BC patients, single-cell analysis was performed on cohort C (GSE118389). To standardize the expression matrix of cohort C and to perform cell cluster analysis, the Seurat package, FindNeighbors, and FindClusters are used. In addition, the differentially expressed genes (DEGs)

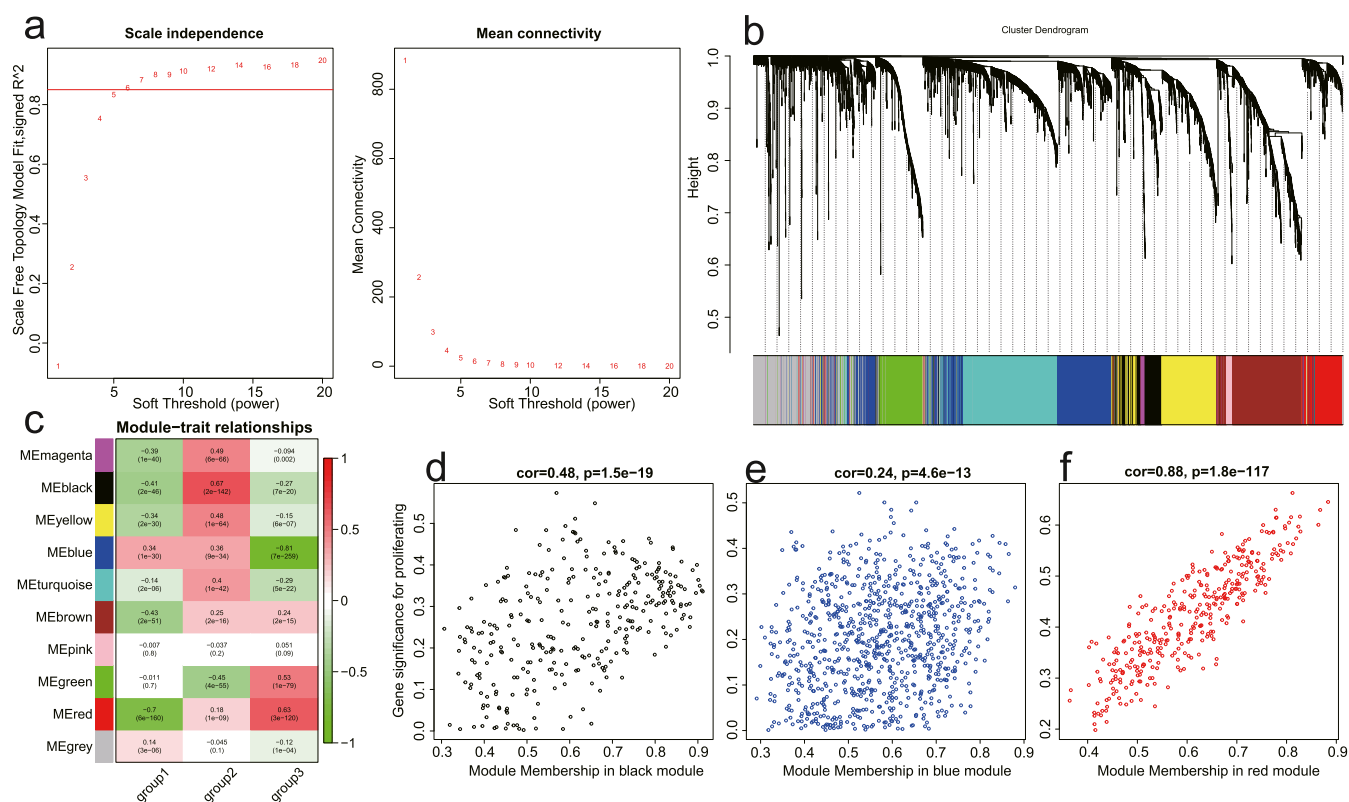


Figure 2. WGCNA analysis of TCGA-BC. (a) Scale-free fit index analysis for different soft thresholds. (b) Cluster dendrogram of TCGA-BC. (c) Correlation between each module and different subtype clusters. (d–f) Correlation between module gene expression and subtype clusters.

among cell clusters were obtained by FindAllMarkers in the Seurat package. Herein, the criteria for screening DEGs were $|\log \text{fold change (FC)}| \geq 1$ and the adjusted p -value ≤ 0.05 . Finally, the cell subgroups determined by the prognostic gene classifier signatures were observed by violin plots, and the interaction between cells was analyzed by the R cellchat (1.6.1) package.³⁴

2.9. Prognostic Nomogram Construction. For further exploring the prognostic role of gene classification signatures, univariate and multivariate regression analysis was also performed to filter the significant prognostic factors. To determine which gene classifier signatures were significantly associated with OS, the univariate regression models were used and p -values were calculated. The gene classifier signatures with a p -value < 0.05 are significantly related to clinicopathologic features. Subsequently, a prognostic nomogram was constructed to improve the prediction ability by integrating important features derived from multiple regression analysis with the survival and RMS packages.³⁵

2.10. Evaluation of Drug Sensitivity and Immune Cell Infiltration. For investigating the molecular characteristics associated with drug sensitivity and drug resistance, the Genomics of Drug Sensitivity in Cancer (GSDC) database (<https://www.cancerrxgene.org/>) was employed, which is a public resource for the discovery of biomarkers in cancer cells. Using the pRRophetic package³⁶ of R, the sensitivity data of three BC subtype clusters to different drugs were counted, and the drug sensitivity of different phenotypes of BC patients from the gene expression data was also predicted.

Additionally, for exploring the differences of different subtypes of immune cells based among multiple clusters of TCGA-BC samples, CIBERSORT,^{37,38} a tool to deconvolute

the expression matrix of immune cell subtypes, was used. Besides, to explore the degree of immune cell infiltration between BC subgroups, the ESTIMATE³⁹ algorithm was also performed.

3. RESULTS

3.1. Identification of Clusters Associated with Subtypes of BC Patients. Employing the ConsensusClusterPlus package of R, the gene expression data allowed the 1006 samples in the TCGA-BC cohort A to be divided into various groups. With the changes in the consensus matrix (CM) K value, the resulting clusters in CM (Figure 1a–d) and their corresponding prognosis (Figure 1e–h) are also different. Particularly, when the $\text{max}K$ value is equal to 3, three clean clusters associated with molecular subtypes of BC in patients are obtained, which are named cluster S1 ($n = 382$), cluster S2 ($n = 328$), and cluster S3 ($n = 296$) subgroups (Figure 1b). Subsequently, Kaplan–Meier survival analysis was performed on the gene expression data of the patients using the log-rank test (Figure 1e–h). Interestingly, only when $K = 3$, the differences in prognosis among the three clustering subgroups are statistically significant (log-rank test $P = 0.028 < 0.05$, Figure 1f), which is consistent with the above results. In detail, the prognosis of cluster S2 patients was the worst, and that of cluster S3 patients was the best (Figure 1f). In contrast, moderate OS was observed in patients of cluster 1 (Figure 1f). Besides, we evaluated the relative changes of areas under the CDF curve, and the results showed that when $K = 3$, the clustering results tend to be stable (Figure 1i). Further, the principal component analysis of gene expression data in these three clusters is depicted in Figure 1j. Consequently, a total of three subtype clusters of BC with significantly different

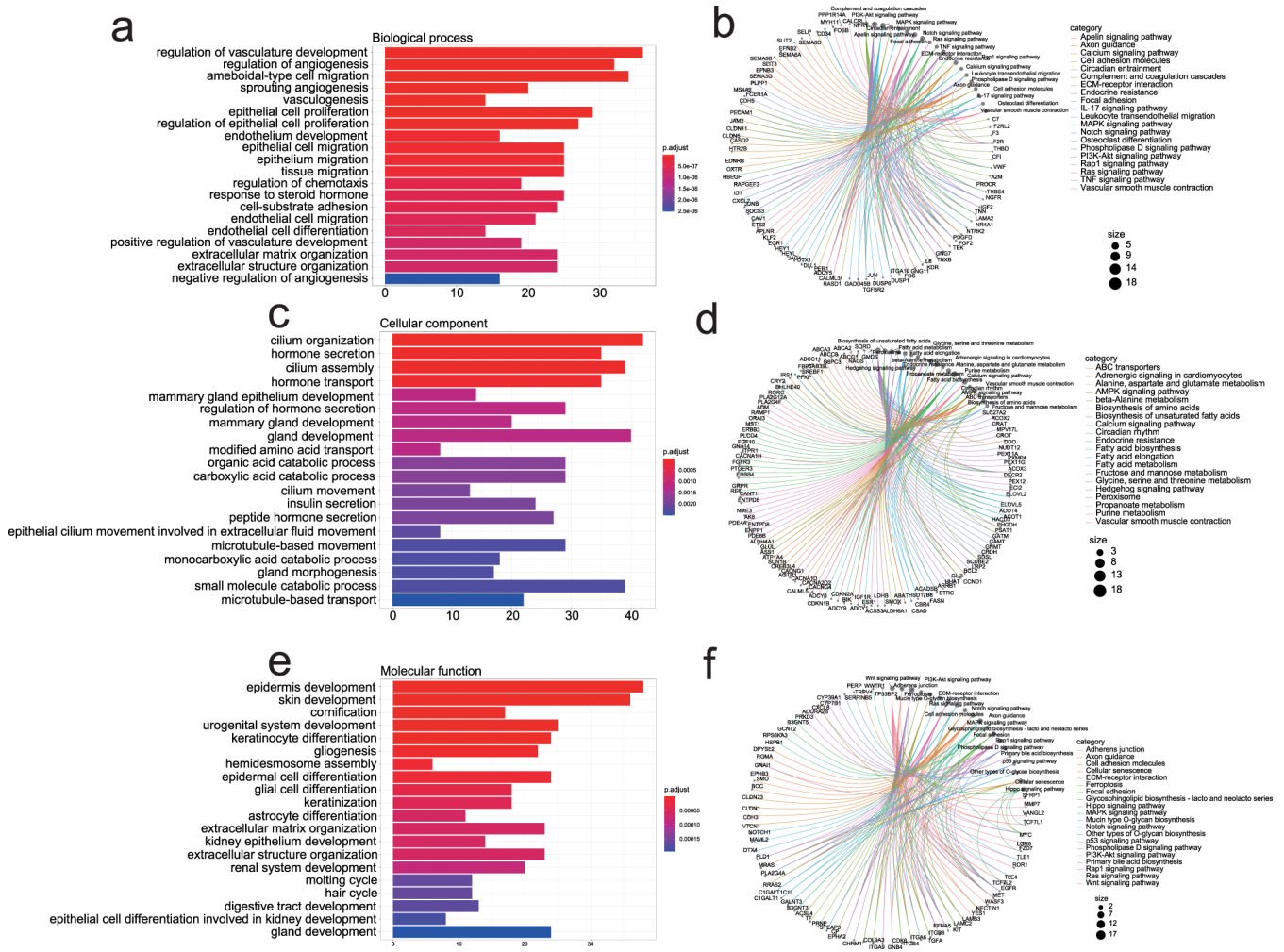


Figure 3. GO bioprocess functional enrichment and pathway analysis of genes in (a, b) subtype cluster S1, (c, d) subtype cluster S2, and (e, f) subtype cluster S3 modules.

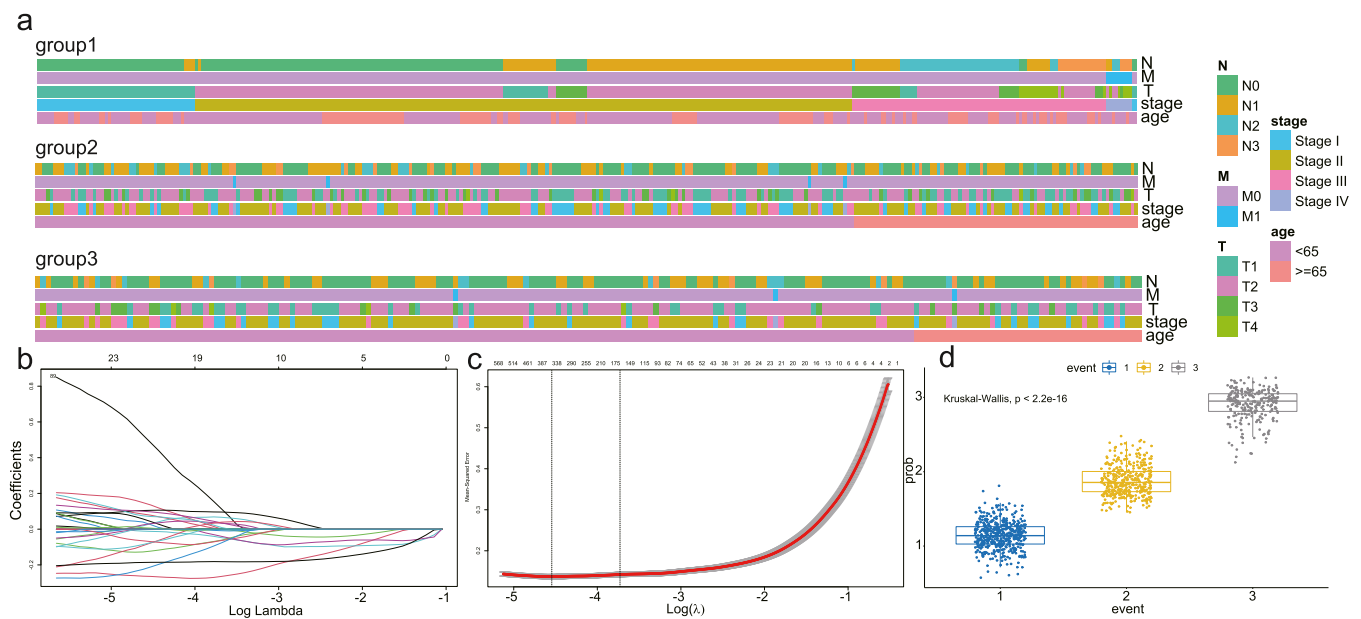


Figure 4. Identification and evaluation of the gene classifier for BC subtyping. (a) Distribution of follow-up data of breast cancer patients in three subtypes. (b) LASSO coefficient profiles of the module genes. (c) Selection of the penalty parameter (λ) in the LASSO model. (d) Gene classifier is represented by the boxplot in the three subtypes.

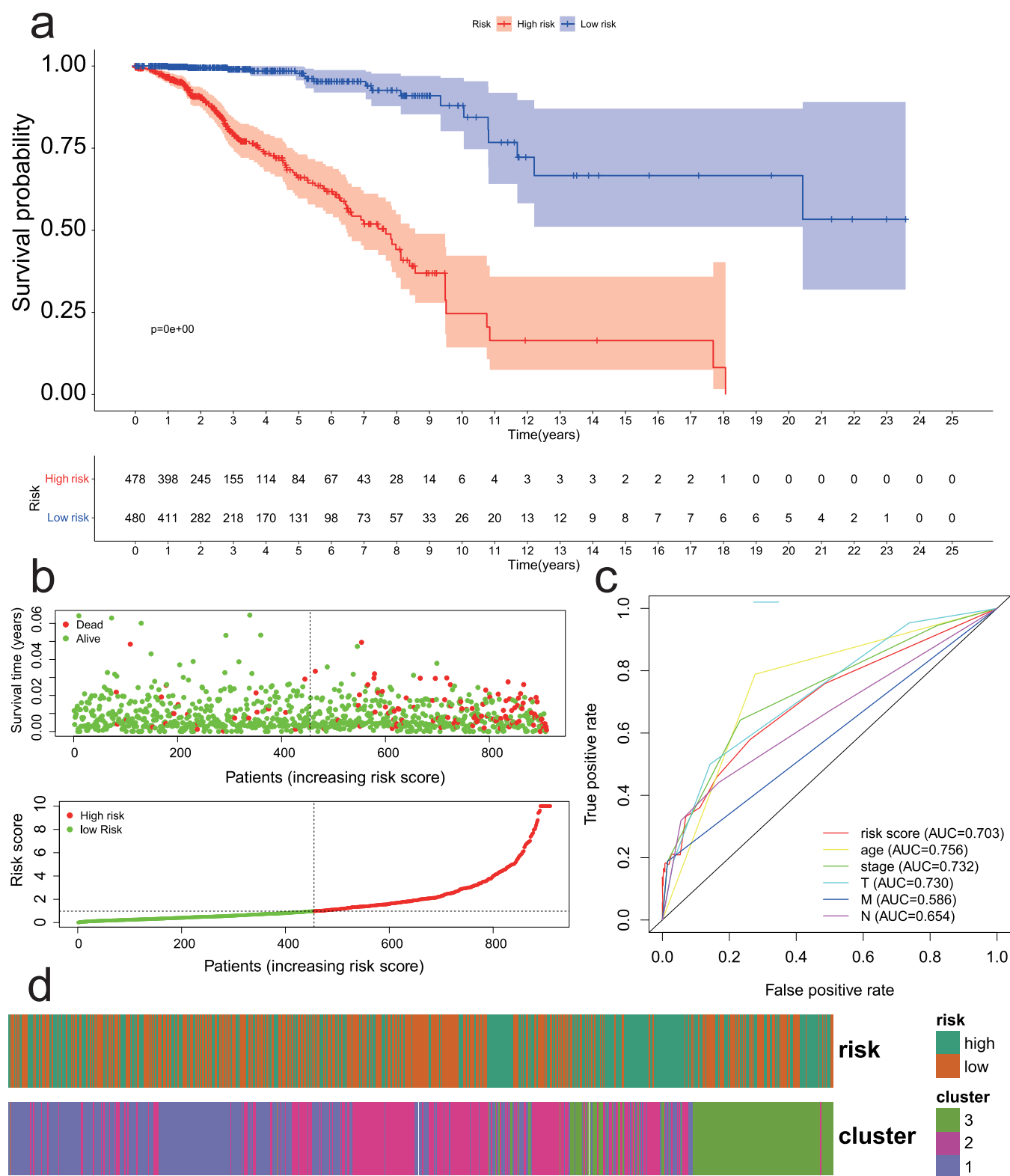


Figure 5. Survival evaluation of gene classifier signatures in TCGA. (a) Kaplan–Meier survival curve of the patients in high- and low-risk scores. (b) Patient subtypes were grouped based on the median level of gene classifier labels in TCGA. (c) Results of ROC analysis for the clinical factors. (d) Distribution of subtypes in high- and low-risk levels.

prognoses were classified by unsupervised consistent clustering, which are used for further research.

3.2. WGCNA Analysis. Presently, WGCNA was carried out on the TCGA–BC cohort A to establish a weighted coexpression network. Clearly, the connection between genes

in the network satisfies the scale-free distribution when the soft threshold $\beta = 6$ (scale-free $R^2 = 0.9$) (Figure 2a). Subsequently, the phylogenetic tree was used to retrieve the coexpression modules (cut height = 0.25) (Figure 2a). Additionally, the hierarchical cluster analysis showed that modules with similar

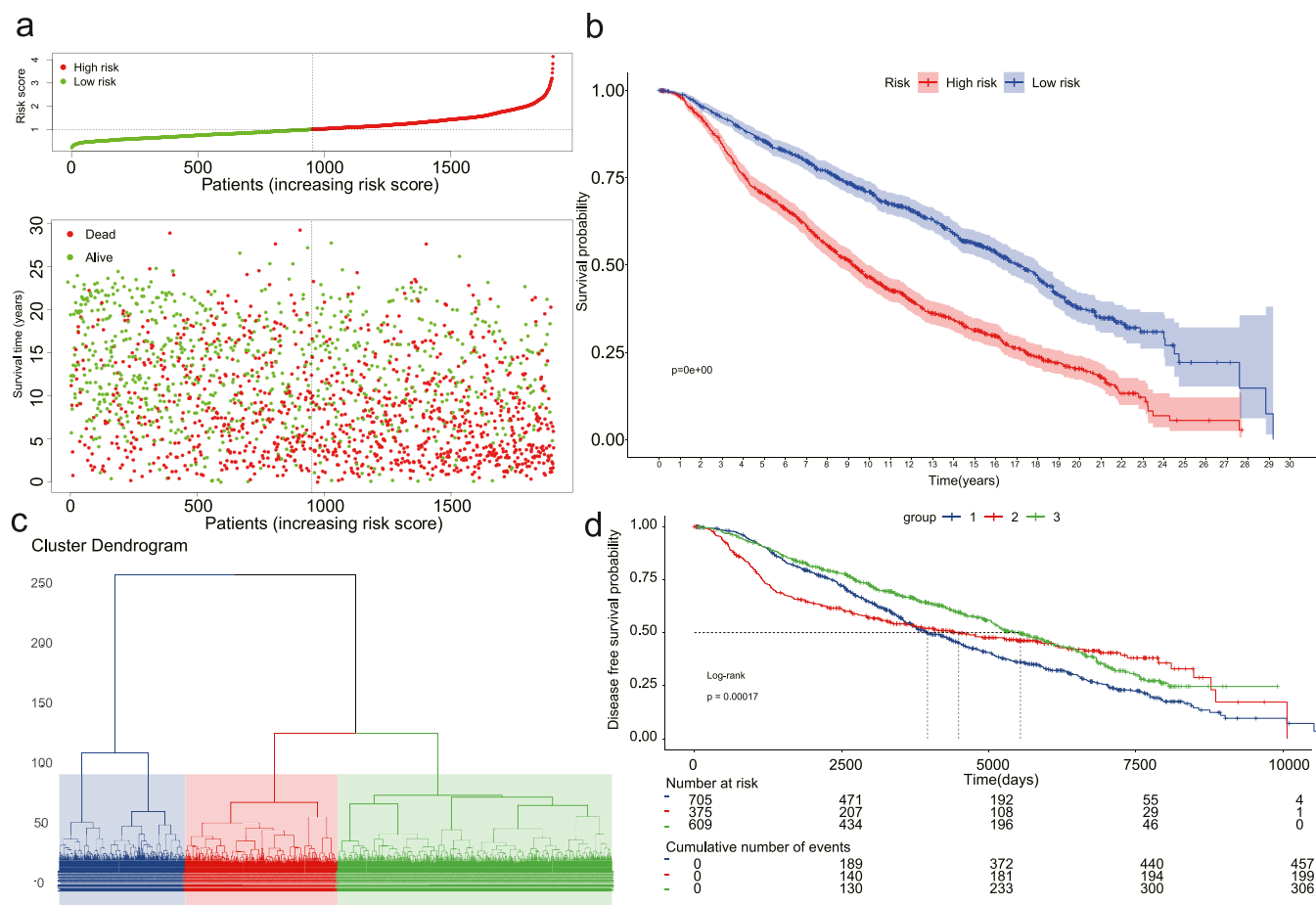


Figure 6. Survival evaluation of gene classifier signatures in METABRIC. (a) Patient subtypes were grouped based on the median level of gene classifier labels in METABRIC. (b) Kaplan–Meier survival curve of OS in different groups for the BC patients in METABRIC. (c) K-means algorithm for molecular subtype classification in the BC dataset of METABRIC ($n = 1904$). (d) Kaplan–Meier survival curve of OS between subtypes in METABRIC.

gene expression patterns were clustered on the same branch (Figure 2b). Figure 2c shows the calculated and visible correlations between modules and the subtype clusters, in which the red color module had a significantly negative association with subtype cluster S1 ($r = 0.70, p < 0.01$). Besides, the black module had the strongest association with subtype cluster S2 ($r = 0.67, p < 0.01$), while the blue module had the strongest association with subtype cluster S3 ($r = -0.81, p < 0.01$). Finally, the correlation between genes and subtypes is depicted in Figure 2d–f. Through WGCNA, a total of three subtype-related gene modules were found, which can be considered as gene signatures specific to each BC subtype.

3.3. BC Subtype Cluster-Specific Enrichment Analysis.

For investigating the biological pathways and biological processes related to BC subtype clusters, enrichment analysis of significant genes in modules was performed. Figure 3a,b shows that most of the biological processes and pathways associated with subtype cluster S1 are related to the regulation of angiogenesis and the PI3K–Akt pathway. The occurrence of these tumor vessels may be related to the tumor progression, which is consistent with the moderate survival time of subtype cluster S1 (Figure 1f). Additionally, the genes in the subtype cluster S2 module show significant enrichment for hormone secretion and the AMPK signaling pathway (Figure 3c,d). Actually, these are linked to the body's active immune metabolic processes, which indicate the body's immunological response as

cancer progresses.⁴⁰ After observing Figure 3e,f, we found that the genes in subtype cluster S3 are mainly enriched in epidermis development and the Wnt signaling pathway. Overall, through enrichment analysis of the gene modules related to the three subtypes, combined with the analysis of the prognosis of BC, the accuracy of unsupervised consistent clustering for BC subtyping is further confirmed, and the basis for clinical treatment of these subtypes is also provided.

3.4. Identification of Gene Classifier Signatures Associated with BC Subtype Clusters.

The distribution of different BC stages and ages in the three subtype clusters is depicted in Figure 4a. Obviously, patients with N and M stages had no significant difference in age, while lower stage and T stages were more common in subtype clusters S1 and S3 (Figure 4a). To avoid overfitting the subtyping characteristics, the LASSO regression on the gene expression of TCGA–BC patients was performed. As a result, a total of 160 genes that are related to BC subtype clusters are obtained (Figure 4b) and the optimal value of penalty parameters through 10 rounds of cross-validation is also determined (Figure 4c). Additionally, the boxplots of the three stages also suggest obvious differences among the three BC subtype clusters, revealing the robustness of constructed gene classifier signatures (Figure 4d).

3.5. Assessment of the Gene Classifier Prognostic Signature for BC.

By analyzing the Kaplan–Meier curve, we found that the OS of BC samples in the high-risk group had a

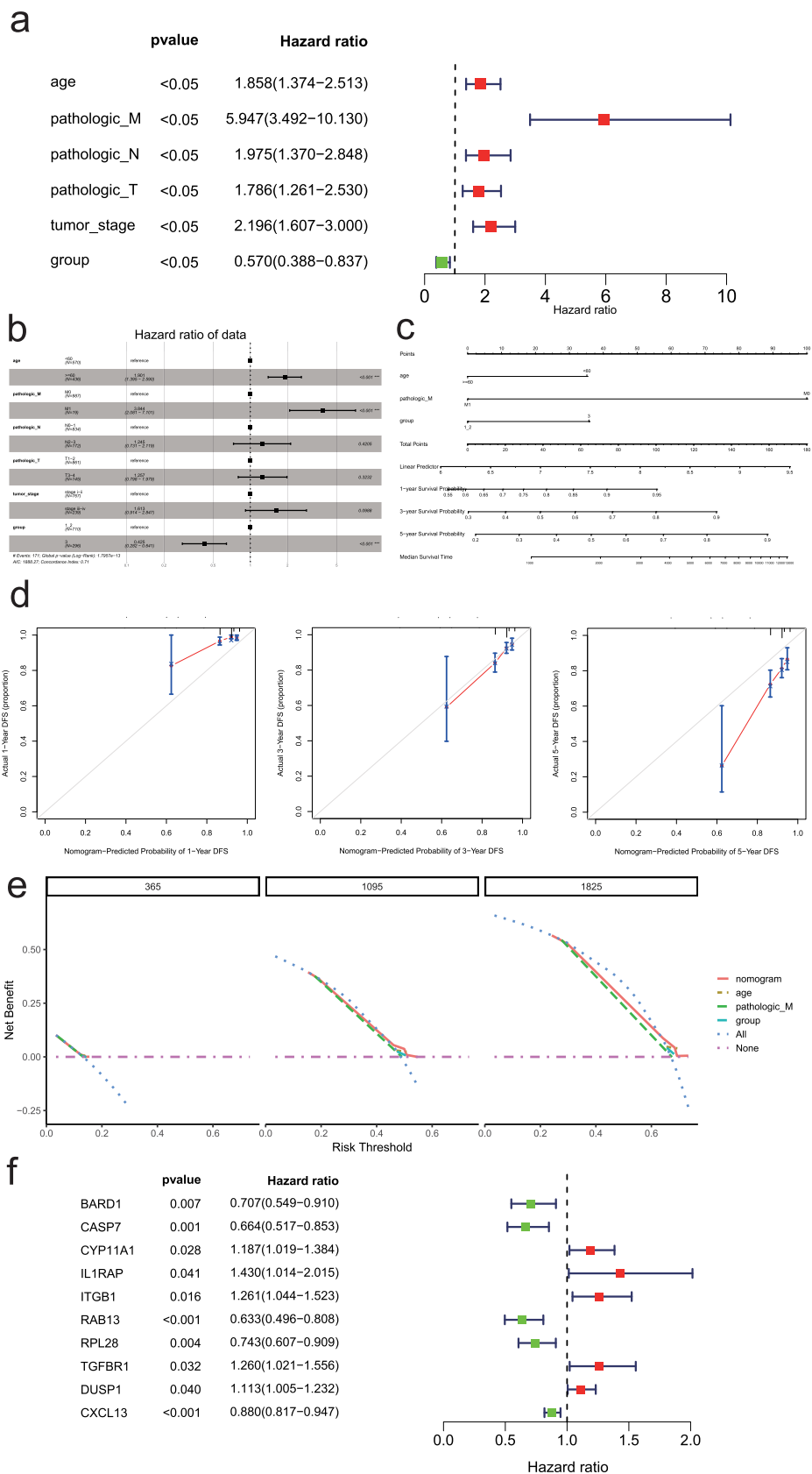


Figure 7. Correlation between the gene classifier signatures and clinicopathological features of TCGA–BC samples. (a, b) Univariate and multivariate analysis including risk status and clinical factors. (c) Nomogram for comprehensively predicting the 1-, 3-, and 5-year OS of BC patients in the TCGA database. (d) Calibration plots for estimating the 1-, 3-, and 5-year OS for TCGA–BC samples. (e) Decision curve plots of the “nomogram,” “group,” “pathology_M,” “age,” “All,” and “None” models by Cox regression analysis. (f) Significantly prognostic genes screened by univariate analysis.

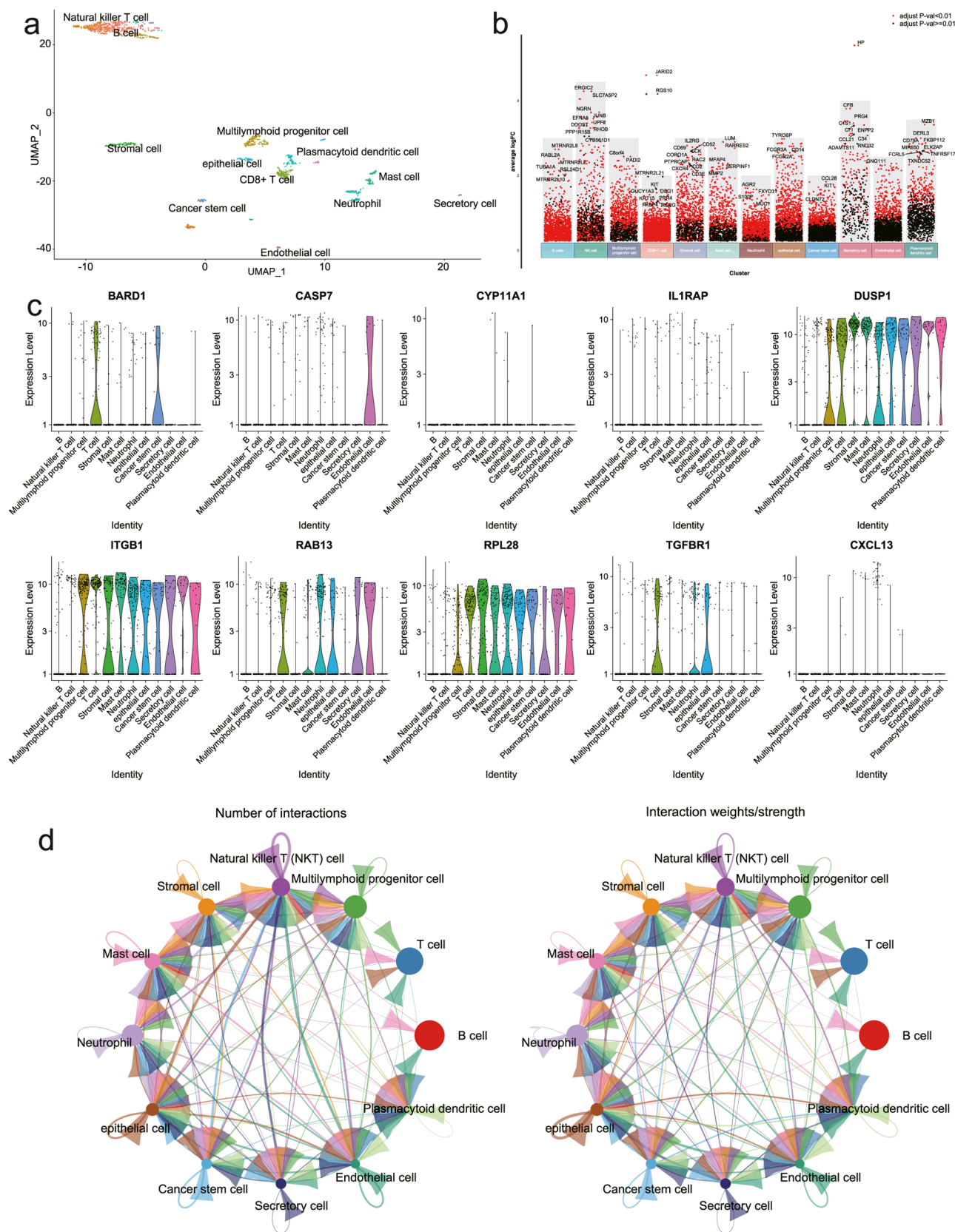


Figure 8. Single-cell sequencing analysis. (a) Cells are divided into thirteen types, including CD8+ T-cells, endothelial cells, mast cells, stromal cells, neutrophil, natural killer T-cells, multilymphoid progenitor cells, epithelial cells, secretory cells, B-cells, plasmacytoid dendritic cells, and cancer stem cells. (b) Volcano of DEGs between different cell populations. (c) Violin plots displaying the expression of representative gene classifier signatures across the cell types identified in the single-cell dataset. The y-axis shows the normalized read count. (d) Cellular communication analysis among tumor cells, NK-T cells, epithelial cells, and secretory cells.

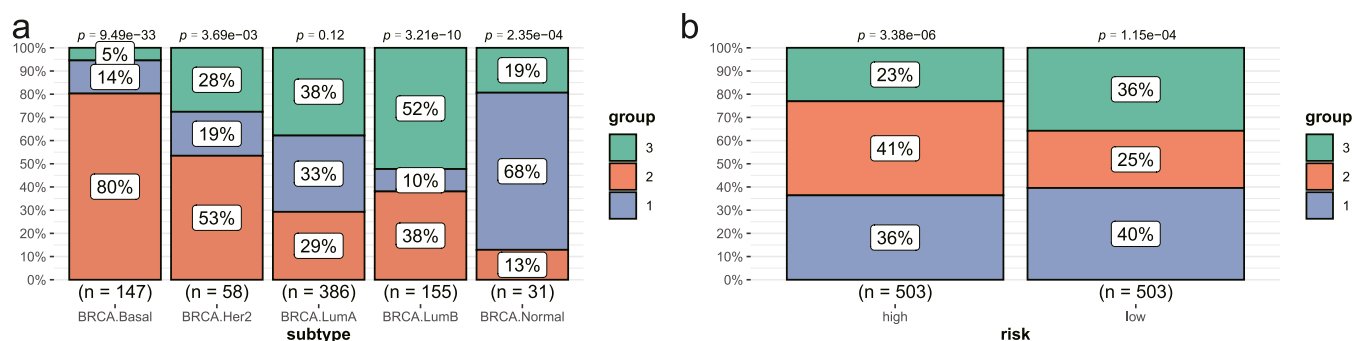


Figure 9. Correlation analysis of the clinical type and subtype clusters. (a) Histogram of the correlation between clinical immunohistochemical typing and subtype clusters. (b) Histogram of the correlation between high- and low-risk typing and subtype clusters.

lower value than those in the low-risk group, indicating that the prognostic marker of the risk score was effective ($p < 0.05$) (Figure 5a). In addition, the risk curve and scatter plot indicate that the risk factors and mortality of the high-risk group were higher than those of the low-risk group (Figure 5b). Furthermore, ROC analysis reveals that the AUC of the risk score was 0.703 (Figure 5c), suggesting the reliability of the obtained gene classifier signatures of BC. At the same time, the survival analysis results of BC subtype clusters obtained by using the unsupervised consistent clustering were relatively consistent with the prognosis results in high- and low-risk groups obtained from the prognostic model. In detail, the subtype cluster S3 with a better survival rate was similar to that in the low-risk group and the subtype cluster S2 with poor survival was close to that in the high-risk group (Figure 5d). All of these findings imply that the gene signature is a predictive factor unique to BC patients.

3.6. Validation of the Prognostic Signature for BC. For further validating the stability of gene classifier signatures, the patients of BC in the METABRIC database were classified into two different groups, namely, high- and low-risk groups. Then, the OS of patients in high- and low-risk groups was analyzed by Kaplan–Meier survival analysis. As a result, samples from high-risk groups have higher OS than those from low-risk groups, suggesting that prognostic markers of risk scores were valid ($P < 0.05$) (Figure 6a). The generated risk curves and scatterplots indicate that compared with the risk factor and mortality in low-risk groups, those in high-risk groups were high (Figure 6b). All of these results suggest that the gene classifiers are highly reliable in marking prognostic markers of BC. Moreover, the molecular intrinsic subtypes of BC patients were determined by the PAM50 classifier algorithm. Clearly, according to gene classified signatures, three subtype clusters were identified (Figure 6c) and there were also significant differences in prognosis ($p < 0.05$) among three subtype clusters (Figure 6d), which is in agreement with the above results. Overall, all of the findings suggest that the genetic classification markers are independent prognostic factors in BC patients.

3.7. Independent Prognostic Ability of Gene Classifier Signatures. For evaluating the independent prognostic ability of the obtained gene classifier signatures, we also investigated the correlation between the gene classifier signatures and clinicopathological features of TCGA–BC samples through univariate regression analysis and multivariate Cox regression analysis. The clinically common factors like age, tumor stage, group, and TNM status were registered as covariates for analysis. As a result, age, M status, and group are independent factors that can be used to predict the prognosis of BC patients (Figure 7a,b). In addition, the clinicopathological information and

independent prognostic factors are combined and then a nomogram involving three items in TCGA was constructed, which is used as a clinically relevant quantitative method for predicting the mortality of BC patients. Obviously, we found that by adding points for each prognostic parameter, each patient will be assigned a total score value and higher total scores corresponded to poorer patient outcomes (Figure 7c). Moreover, calibration plots for the TCGA cohort A reveal a similar performance of the nomogram to the developed prognostic model (Figure 7d). In addition, the DCA curve also indicated that the nomogram had good stability and reliability net benefit curve in age compared with other clinical factors (Figure 7e). Finally, we further screened the gene signatures that are associated with the prognosis by employing univariate regression analysis. As a result, a total of ten gene signatures significantly related to prognosis were obtained, which contain BARD1, CASP7, CYP11A1, IL1RAP, ITGB1, RAB13, RPL28, TGFBR1, DUSP1, and CXCL13 (Figure 7f).

3.8. Evaluation of Cellular Heterogeneity in the Subtype Cluster of BC Patients. To evaluate the heterogeneity of the molecular intrinsic subtypes of BC patients, single-cell analysis was performed on GSE118389. Employing the immune cell-specific marker genes obtained from the published article,⁴¹ more than 1500 cells from six primary BC patients were clustered into 13 clusters and 12 cell types, *i.e.*, CD8⁺ T-cells, endothelial cells, mast cells, stromal cells, neutrophil, natural killer T-cells, multilymphoid progenitor cells, epithelial cells, secretory cells, B-cells, plasmacytoid dendritic cells, and cancer stem cells (Figure 8a). Subsequently, the DEGs between different cell types were identified and visualized by the volcano plots (Figure 8b). In addition, the signature distribution of key prognostic gene classifiers reveals that BARD1 is mainly expressed in T and cancer stem cells, while CASP7 is mainly expressed in endothelial cells (Figure 8c). Moreover, DUSP1, ITGB1, RAB13, and RPL28 are expressed in subsets other than T-cells and B-cells (Figure 8c). With respect to TGFBR1, it was mainly expressed in T-cells, neutrophils, and epithelial cells. In contrast, CYP11A1, IL1RAP, and CXCL13 were not specifically expressed (Figure 8c). Furthermore, to characterize the communication between key immune cells, tumor cells, and other microenvironmental cells, the significant ligand–receptor interactions between 12 cell types were detected and analyzed through the CellChat analysis. The results show that there is extensive cellular communication among tumor cells, NK-T cells, epithelial cells, and secretory cells (Figure 8d).

3.9. Distribution of Subtype Clusters in BC Intrinsic Profiles and Risk Subgroups. To investigate the association

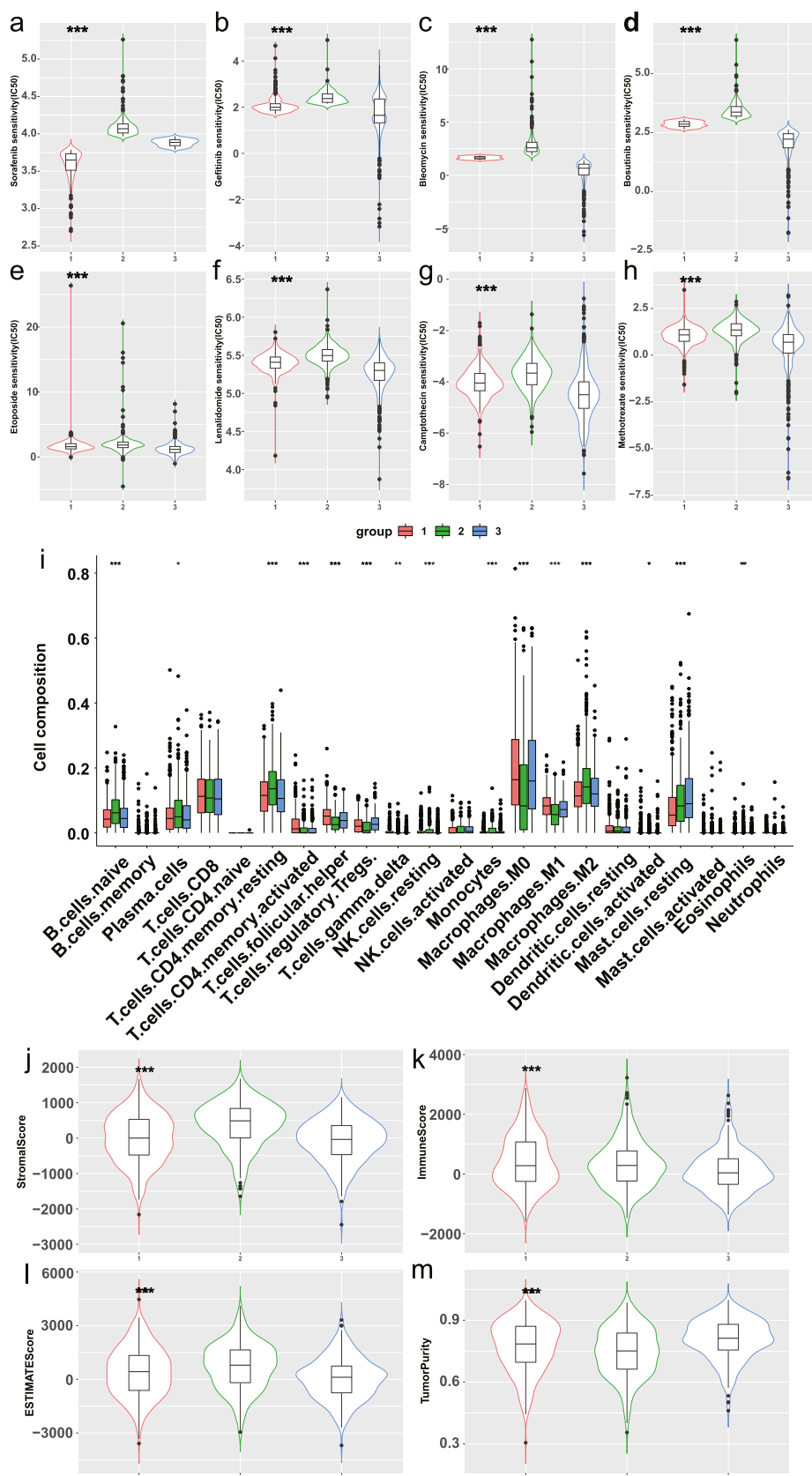


Figure 10. Drug sensitivity analysis as well as immune cell infiltration evaluation. (a–h) Scatter plot shows that drug sensitivity prediction scores are distributed differently among subtypes. (i) Comparison of immune cell proportions between subtypes (TCGA). (j–m) Scatter plot for the immune score, stromal score, estimated score, and tumor purity, which are distributed differently among subtypes.

between the subtype clusters and the clinicopathological characteristics and risk subgroups of BC patients, we assessed the proportion of subtype clusters in BC molecular intrinsic profiles (*i.e.*, basal, Her2, LumA, LumB, and normal) and the risk subgroups (*i.e.*, high, low) using the chi-square analysis.⁴² Interestingly, all of the BC molecular intrinsic profiles were identified in the three subtype clusters, but the proportions are different (Figure 9a). For example, the proportion of subtype cluster S2 with poor prognosis in the basal-like profile was 80% ($p = 9.49 \times 10^{-33}$), whereas subtype cluster S1 (68%) showed a BC normal-enriched profile ($p = 2.35 \times 10^{-4}$, Figure 9a). Of note, 52% of subtype cluster S3 exhibited a BC LumB-enriched profile ($p = 3.21 \times 10^{-10}$, Figure 9a). Additionally, compared with the proportion of subtype cluster S2 patients in the low-risk group, that in the high-risk group was significantly higher ($p = 3.38 \times 10^{-6}$), whereas the proportion of samples in subtype clusters S1 and S3 in the low-risk group are higher ($p = 1.15 \times 10^{-4}$, Figure 9b). All of the results indicate that the classified subtype clusters and the clinicopathological features of BC patients exhibit a certain correlation.

3.10. Drug Sensitivity Analysis and Immune Cell Infiltration Evaluation. Using the “pRRophetic” package of R, the IC50 differences of eight targeted and chemotherapeutic drugs in three different subtypes of clustering were studied to predict their sensitivities to drug therapy. Figure 10a–h, respectively, shows the drug sensitivity results of Sorafenib, Gefitinib, Bleomycin, Bosutinib, Etoposide, Lenalidomide, Camptothecin, and Methotrexate in three subtypes of BC clusters. The statistical results demonstrated that subtype cluster S2 samples are generally more sensitive to tumor treatment drugs than the other two subtype clusters. In contrast, subtype cluster S3 is less sensitive to most drugs (Figure 10a–h).

Additionally, the differences of 22 subtypes of immune cells among TCGA–BC samples in three subtype clusters are investigated using the CIBERSORT algorithm. According to the results, there are significant differences in immune infiltration between the three subtype clusters for B-cells naive, plasma cells, T-cells CD4 memory resting, T-cells CD4 memory activated, T-cells follicular helper, T-cells regulatory Tregs, T-cells γ delta, NK cells resting, monocytes, macrophages M0, M1, M2, dendritic cells activated, mast cells resting, and eosinophils (Figure 10i), taking up a sizable portion of immune cell invasion. Besides, compared to other clusters with a poorer prognosis, the subtype cluster S1 with better survival showed a higher number of M1 and M2 macrophages and activated NK cells (Figure 10i). Finally, the tumor purity for all BC samples in three subtype clusters as well as the stromal and immune scores were calculated. As a result, subtype cluster S2 had better immune, stromal, and estimate scores than the others, while subtype clusters S1 and S3 had more pure tumors (Figure 10j–m).

4. DISCUSSION

With the rapid development of high-throughput sequencing and the accumulation of massive biomedical data, a large number of bioinformatics methods have been used to identify new gene signatures related to the progression, survival, or prognosis of malignant carcinoma.^{43–45} Particularly, the critical role of the TME in tumor biology has made targeted TME therapy for tumors a research hotspot. BC, as one of the most common malignant tumors in women, has attracted more attention from researchers due to its poorer prognosis.⁴⁶ As a biologically heterogeneous disease, BC has different molecular subtypes

according to the gene expression profiles. Presently, we attempt to identify the BC subtype clusters by integrating genomic analysis with the TME. As far as we know, a large-amount dataset combined with multiple bioinformatics methods was first used to identify the intrinsic subtypes and the gene classifier signatures for BC patients.

First, we performed unsupervised consistent clustering on the mRNA profiles of TCGA–BC cohort A and a total of three BC subtype clusters were obtained, which are named cluster S1 ($n = 382$), cluster S2 ($n = 328$), and cluster S3 ($n = 296$) subtypes (Figure 1b). Then, three corresponding subtype-related gene modules were obtained from WGCNA, which are specific to each BC subtype cluster (Figure 2c). Further analysis of enrichment reveals that genes in subtype cluster S1 are principally involved in the regulation of angiogenesis and PI3K–Akt pathways, and genes in subtype cluster S2 exhibit significant enrichment for hormone secretion and AMPK signaling pathways (Figure 3a–d). Additionally, genes in subtype cluster S3 are mainly enriched in epidermis development and Wnt signaling pathways (Figure 3e,f). Actually, these biological processes and signaling pathways exhibit crucial roles in the development of BC. For instance, angiogenesis is a crucial process in the formation of inflammatory BC (IBC), and genes and proteins involved in angiogenesis are often overexpressed and overactivated in IBC.^{47,48} Recent studies have shown that hormone receptor (HR) + BC causes the majority of BC-related deaths.⁴⁹ Yu et al. have also reported that in BC patients with bone, lung, and brain metastases, actionable mutations frequently occur in the PI3K/AKT/mTOR pathway.⁵⁰ Inhibiting the activation of the PI3K/Akt pathway partially restores the sensitivity of trastuzumab-resistant HER2-positive BC cells.⁵¹

In addition, the genetic diversities of BC patients demonstrate that it is important to identify gene signatures with distinct survival and druggable targets. Presently, LASSO regression analysis on the gene expression profiles of TCGA–BC patients obtained a total of 160 gene classifiers related to the BC subtype (Figure 4b). Employing the obtained gene classifiers, we built the risk prognostic model, and this model was also then validated by the gene expression data of cohort B in the METABRIC database. The results show that the OS of samples from high-risk groups is higher than those in low-risk groups, suggesting that the risk model of prognostic markers was reliable ($P < 0.05$) (Figure 6a). Indeed, the important role of OS in multiple stages of tumorigenesis and development has been reported in many studies.^{52,53} Further, evaluation of the independent prognostic ability of obtained gene classifier signatures indicates that the BC patients in subtype cluster S2 have the poorest prognosis, while subtype clusters S1 and S3 exhibit good prognosis (Figure 5d). Subsequently, ten gene signatures significantly related to prognosis were obtained, which contain BARD1, CASP7, CYP11A1, IL1RAP, ITGB1, RAB13, RPL28, TGFBR1, DUSP1, and CXCL13 (Figure 7f). All of these results suggest that the independent prognostic factors of BC patients can be the genetic classifier signatures.

Besides, it is well-known that TME has an important role during the progression of human cancers, as well as the OS after the operation.^{54,55} Presently, the heterogeneity of the molecular subtypes of BC patients was also explored by the single-cell analysis on GSE118389. The results demonstrate that most of the key prognosis-related gene classification signatures are distributed in immune cells. Notably, there were more neutrophils in the subtype S2 with poor prognosis (Figure 8c), which is consistent with the previous study.⁵⁶ Single-cell

analysis also shows that subtype cluster S1 contains more types of immune cells, while subtype cluster S3 with a good prognosis includes more multilymphoid progenitor cells (Figure 8c). Moreover, the chi-square analysis exhibits that 80% of basal-like profile samples are found in subtype cluster S2 ($p = 9.49 \times 10^{-33}$), 68% of the normal-enriched profiles are subtype cluster S1 samples, and 52% of the subtype cluster S3 showed LumB enrichment ($p = 3.21 \times 10^{-10}$, Figure 9a), indicating the association among the classification subtype cluster and the clinicopathological features of patients with BC. Finally, drug sensitivity analysis demonstrates that subtype cluster S2 samples are generally more sensitive to tumor treatment drugs and subtype cluster S3 is less sensitive to most drugs (Figure 10). All of the findings may provide a basis for targeting the subtype-specific molecular and immunotherapy of BC. However, the gene signatures that characterize each subtype cluster of BC need further be studied and validated by biological experiments.

5. CONCLUSIONS

Taken together, three subtype clusters, S1, S2, and S3, are obtained from the gene expression profiles of TCGA-BC cohort A. WGCNA and pathway enrichment analysis revealed that PI3K-Akt, AMPK, and Wnt signaling pathways, respectively, participated in the subtype clusters S1, S2, and S3. Additionally, the risk prognostic model was constructed by the obtained 160 gene classifiers associated with the BC subtype and validated by the data in the METABRIC database. Clinical prognostic evaluation reveals the poor prognostic ability of subtype cluster S2 and the good prognostic ability of subtype clusters S1 and S3. Then, ten crucial prognostic-related gene signatures, *i.e.*, BARD1, CASP7, CYP11A1, IL1RAP, ITGB1, RAB13, RPL28, TGFBR1, DUSP1, and CXCL13, are screened. Finally, single-cell analysis demonstrates that most types of immune cells are enriched in subtype cluster S1 samples, more neutrophils are distributed in the subtype S2 with more sensitivity to tumor treatment drugs, and more multilymphoid progenitor cells are involved in subtype cluster S3. All of the results may provide an effective basis for the clinical precision treatment of BC.

AUTHOR INFORMATION

Corresponding Authors

Hongxing Kan – School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012, China; Anhui Computer Application Research Institute of Chinese Medicine, China Academy of Chinese Medical Sciences, Hefei 230013, China; Email: ffdkx@ahcm.edu.cn

Yinfeng Yang – School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012, China; Anhui Computer Application Research Institute of Chinese Medicine, China Academy of Chinese Medical Sciences, Hefei 230013, China; orcid.org/0000-0002-2557-4807; Email: yinfengyang@yeah.net

Authors

Ran Ding – School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012, China; Anhui Computer Application Research Institute of Chinese Medicine, China Academy of Chinese Medical Sciences, Hefei 230013, China

Qiwei Liu – School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012, China

Jing Yu – School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012, China
Yongkang Wang – School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012, China
Honglei Gao – School of Medical Informatics Engineering, Anhui University of Chinese Medicine, Hefei 230012, China

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.2c08227>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank the National Science Foundation for Young Scientists of China (No. 81904062), the Outstanding Youth Research Project of Anhui Department of Education (No. 2022AH020042), the Natural Science Key Project of Anhui University of Chinese Medicine (No. 2020zrz12), and the Natural Science Key Project of Anhui universities (Nos. KJ2021A0577 and KJ2020A0394).

REFERENCES

- (1) Stingl, J.; Caldas, C. Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nat. Rev. Cancer* **2007**, *7*, 791–799.
- (2) Dai, X.; Li, T.; Bai, Z.; Yang, Y.; Liu, X.; Zhan, J.; Shi, B. Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* **2015**, *5*, 2929–2943.
- (3) Liang, M.; Li, Z.; Chen, T.; Zeng, J. Integrative Data Analysis of Multi-Platform Cancer Data with a Multimodal Deep Learning Approach. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2015**, *12*, 928–937.
- (4) de Rinaldis, E.; Gazinska, P.; Mera, A.; Modrusan, Z.; Fedorowicz, G. M.; Burford, B.; Gillett, C.; Marra, P.; Grigoriadis, A.; Dornan, D.; et al. Integrated genomic analysis of triple-negative breast cancers reveals novel microRNAs associated with clinical and molecular phenotypes and sheds light on the pathways they control. *BMC Genomics* **2013**, *14*, No. 643.
- (5) Kim, S.; Kon, M.; DeLisi, C. Pathway-based classification of cancer subtypes. *Biol. Direct* **2012**, *7*, No. 21.
- (6) Perou, C. M.; Sorlie, T.; Eisen, M. B.; van de Rijn, M.; Jeffrey, S. S.; Rees, C. A.; Pollack, J. R.; Ross, D. T.; Johnsen, H.; Akslen, L. A.; et al. Molecular portraits of human breast tumours. *Nature* **2000**, *406*, 747–752.
- (7) Chin, K.; DeVries, S.; Fridlyand, J.; Spellman, P. T.; Roydasgupta, R.; Kuo, W. L.; Lapuk, A.; Neve, R. M.; Qian, Z.; Ryder, T.; et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiology. *Cancer Cell* **2006**, *10*, 529–541.
- (8) Yang, R.; Li, Y.; Wang, H.; Qin, T.; Yin, X.; Ma, X. Therapeutic progress and challenges for triple negative breast cancer: targeted therapy and immunotherapy. *Mol. Biomed.* **2022**, *3*, No. 8.
- (9) Chaudhary, L. N.; Wilkinson, K. H.; Kong, A. Triple-Negative Breast Cancer: Who Should Receive Neoadjuvant Chemotherapy? *Surg. Oncol. Clin. North Am.* **2018**, *27*, 141–153.
- (10) Collignon, J.; Lousberg, L.; Schroeder, H.; Jerusalem, G. Triple-negative breast cancer: treatment challenges and solutions. *Breast Cancer: Targets Ther.* **2016**, *8*, 93–107.
- (11) Manuel, M.; Tredan, O.; Bachelot, T.; Clapisson, G.; Courtier, A.; Parmentier, G.; Rabeony, T.; Grives, A.; Perez, S.; Mouret, J. F.; et al. Lymphopenia combined with low TCR diversity (divpenia) predicts poor overall survival in metastatic breast cancer patients. *Oncoimmunology* **2012**, *1*, 432–440.
- (12) Papatostas, A. E.; Lesnick, G. J.; Jenkins, G.; Aufses, A. H., Jr. The prognostic significance of peripheral lymphocyte counts in patients with breast carcinoma. *Cancer* **1976**, *37*, 164–168.
- (13) Ali, H. R.; Chlon, L.; Pharoah, P. D.; Markowitz, F.; Caldas, C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical

- Implications: A Gene-Expression-Based Retrospective Study. *PLoS Med.* **2016**, *13*, No. e1002194.
- (14) Aran, D.; Hu, Z.; Butte, A. J. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **2017**, *18*, No. 220.
- (15) Clancy, T.; Dannenfelser, R.; Troyanskaya, O.; Malmberg, K. J.; Hovig, E.; Kristensen, V. Bioinformatics Approaches to Profile the Tumor Microenvironment for Immunotherapeutic Discovery. *Curr. Pharm. Des.* **2017**, *23*, 4716–4725.
- (16) Dannenfelser, R.; Nome, M.; Tahiri, A.; Ursini-Siegel, J.; Vollan, H. K. M.; Haakensen, V. D.; Helland, Å.; Naume, B.; Caldas, C.; Børresen-Dale, A. L.; et al. Data-driven analysis of immune infiltrate in a large cohort of breast cancer and its association with disease progression, ER activity, and genomic complexity. *Oncotarget* **2017**, *8*, 57121–57133.
- (17) Galon, J.; Angell, H. K.; Bedognetti, D.; Marincola, F. M. The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures. *Immunity* **2013**, *39*, 11–26.
- (18) Newman, A. M.; Liu, C. L.; Green, M. R.; Gentles, A. J.; Feng, W.; Xu, Y.; Hoang, C. D.; Diehn, M.; Alizadeh, A. A. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **2015**, *12*, 453–457.
- (19) Alexe, G.; Dalgin, G. S.; Scandfeld, D.; Tamayo, P.; Mesirov, J. P.; DeLisi, C.; Harris, L.; Barnard, N.; Martel, M.; Levine, A. J.; et al. High expression of lymphocyte-associated genes in node-negative HER2+ breast cancers correlates with lower recurrence rates. *Cancer Res.* **2007**, *67*, 10669–10676.
- (20) Teschendorff, A. E.; Miremadi, A.; Pinder, S. E.; Ellis, I. O.; Caldas, C. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. *Genome Biol.* **2007**, *8*, No. R157.
- (21) Bense, R. D.; Sotiriou, C.; Piccart-Gebhart, M. J.; Haanen, J. B. A. G.; van Vugt, M. A. T. M.; de Vries, E. G. E.; Schröder, C. P.; Fehrmann, R. S. N. Relevance of Tumor-Infiltrating Immune Cell Composition and Functionality for Disease Outcome in Breast Cancer. *J. Natl. Cancer Inst.* **2016**, *109*, No. djw192.
- (22) Tomczak, K.; Czerwińska, P.; Wiznerowicz, M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Współczesna Onkol.* **2015**, *19*, 68–77.
- (23) Curtis, C.; Shah, S. P.; Chin, S. F.; Turashvili, G.; Rueda, O. M.; Dunning, M. J.; Speed, D.; Lynch, A. G.; Samarajiwa, S.; Yuan, Y.; et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **2012**, *486*, 346–352.
- (24) Karaayvaz, M.; Cristea, S.; Gillespie, S. M.; Patel, A. P.; Mylvaganam, R.; Luo, C. C.; Specht, M. C.; Bernstein, B. E.; Michor, F.; Ellisen, L. W. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat. Commun.* **2018**, *9*, No. 3588.
- (25) Wilkerson, M. D.; Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **2010**, *26*, 1572–1573.
- (26) Langfelder, P.; Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* **2008**, *9*, No. 559.
- (27) Ferreira, M. A.; Vonk, J. M.; Baurecht, H.; Marenholz, I.; Tian, C.; Hoffman, J. D.; Helmer, Q.; Tillander, A.; Ullemar, V.; van Dongen, J.; et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat. Genet.* **2017**, *49*, 1752–1757.
- (28) Ito, K.; Murphy, D. Application of ggplot2 to Pharmacometric Graphics. *CPT: Pharmacometrics Syst. Pharmacol.* **2013**, *2*, No. 79.
- (29) Mahmoud, S. M.; Paish, E. C.; Powe, D. G.; Macmillan, R. D.; Grainge, M. J.; Lee, A. H.; Ellis, I. O.; Green, A. R. Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. *J. Clin. Oncol.* **2011**, *29*, 1949–1955.
- (30) Engebretsen, S.; Böhlin, J. Statistical predictions with glmnet. *Clin. Epigenet.* **2019**, *11*, No. 123.
- (31) Heagerty, P. J.; Zheng, Y. Survival model predictive accuracy and ROC curves. *Biometrics* **2005**, *61*, 92–105.
- (32) Ramadan, F.; Fahs, A.; Ghayad, S. E.; Saab, R. Signaling pathways in Rhabdomyosarcoma invasion and metastasis. *Cancer Metastasis Rev.* **2020**, *39*, 287–301.
- (33) Pu, M.; Messer, K.; Davies, S. R.; Vickery, T. L.; Pittman, E.; Parker, B. A.; Ellis, M. J.; Flatt, S. W.; Marinac, C. R.; Nelson, S. H.; Mardis, E. R.; Pierce, J. P.; Natarajan, L. Research-based PAM50 signature and long-term breast cancer survival. *Breast Cancer Res. Treat.* **2020**, *179*, 197–206.
- (34) Jin, S.; Guerrero-Juarez, C. F.; Zhang, L.; Chang, I.; Ramos, R.; Kuan, C. H.; Myung, P.; Plikus, M. V.; Nie, Q. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **2021**, *12*, No. 1088.
- (35) Crago, A. M.; Denton, B.; Salas, S.; Dufresne, A.; Mezghir, J. J.; Hameed, M.; Gonen, M.; Singer, S.; Brennan, M. F. A prognostic nomogram for prediction of recurrence in desmoid fibromatosis. *Ann. Surg.* **2013**, *258*, 347–353.
- (36) Geeleher, P.; Cox, N.; Huang, R. S. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One* **2014**, *9*, No. e107468.
- (37) Chen, B.; Khodadoust, M. S.; Liu, C. L.; Newman, A. M.; Alizadeh, A. A. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol. Biol.* **2018**, *1711*, 243–259.
- (38) Kawada, J. I.; Takeuchi, S.; Imai, H.; Okumura, T.; Horiba, K.; Suzuki, T.; Torii, Y.; Yasuda, K.; Imanaka-Yoshida, K.; Ito, Y. Immune cell infiltration landscapes in pediatric acute myocarditis analyzed by CIBERSORT. *J. Cardiol.* **2021**, *77*, 174–178.
- (39) Bhattacharya, S.; Dunn, P.; Thomas, C. G.; Smith, B.; Schaefer, H.; Chen, J.; Hu, Z.; Zalocusky, K. A.; Shankar, R. D.; Shen-Orr, S. S.; Thomson, E.; Wiser, J.; Butte, A. J. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci. Data* **2018**, *5*, No. 180015.
- (40) Pearce, E. L.; Poffenberger, M. C.; Chang, C. H.; Jones, R. G. Fueling immunity: insights into metabolism and lymphocyte function. *Science* **2013**, *342*, No. 1242454.
- (41) Clarke, Z. A.; Andrews, T. S.; Atif, J.; Pouyababar, D.; Innes, B. T.; MacParland, S. A.; Bader, G. D. Tutorial: guidelines for annotating single-cell transcriptomic maps using automated and manual methods. *Nat. Protoc.* **2021**, *16*, 2749–2764.
- (42) Alavi, M.; Visentin, D. C.; Thapa, D. K.; Hunt, G. E.; Watson, R.; Cleary, M. Chi-square for model fit in confirmatory factor analysis. *J. Adv. Nurs.* **2020**, *76*, 2209–2211.
- (43) Wu, P.; Liu, J. L.; Pei, S. M.; Wu, C. P.; Yang, K.; Wang, S. P.; Wu, S. Integrated genomic analysis identifies clinically relevant subtypes of renal clear cell carcinoma. *BMC Cancer* **2018**, *18*, No. 287.
- (44) Li, S.; Liu, X.; Liu, T.; Meng, X.; Yin, X.; Fang, C.; Huang, D.; Cao, Y.; Weng, H.; Zeng, X.; Wang, X. Identification of Biomarkers Correlated with the TNM Staging and Overall Survival of Patients with Bladder Cancer. *Front. Physiol.* **2017**, *8*, No. 947.
- (45) Zhu, Z.; Jin, Z.; Deng, Y.; Wei, L.; Yuan, X.; Zhang, M.; Sun, D. Co-expression Network Analysis Identifies Four Hub Genes Associated With Prognosis in Soft Tissue Sarcoma. *Front. Genet.* **2019**, *10*, No. 37.
- (46) Siegel, R. L.; Miller, K. D.; Jemal, A. Cancer statistics, 2019. *CA: Cancer J. Clin.* **2019**, *69*, 7–34.
- (47) Lim, B.; Woodward, W. A.; Wang, X.; Reuben, J. M.; Ueno, N. T. Inflammatory breast cancer biology: the tumour microenvironment is key. *Nat. Rev. Cancer* **2018**, *18*, 485–499.
- (48) Van der Auwera, I.; Van Laere, S. J.; Van den Eynden, G. G.; Benoy, I.; van Dam, P.; Colpaert, C. G.; Fox, S. B.; Turley, H.; Harris, A. L.; Van Marck, E. A.; Vermeulen, P. B.; Dirix, L. Y. Increased angiogenesis and lymphangiogenesis in inflammatory versus non-inflammatory breast cancer by real-time reverse transcriptase-PCR gene expression quantification. *Clin. Cancer Res.* **2004**, *10*, 7965–7971.
- (49) Munoz, D.; Near, A. M.; van Ravesteyn, N. T.; Lee, S. J.; Schechter, C. B.; Alagoz, O.; Berry, D. A.; Burnside, E. S.; Chang, Y.; Chisholm, G.; de Koning, H. J.; Ali Ergun, M.; Heijnsdijk, E. A.; Huang, H.; Stout, N. K.; Sprague, B. L.; Trentham-Dietz, A.; Mandelblatt, J. S.; Plevritis, S. K. Effects of screening and systemic adjuvant therapy on ER-specific US breast cancer mortality. *JNCI: J. Natl. Cancer Inst.* **2014**, *106*, No. dju289.
- (50) Yu, Z.; Song, M.; Chouchane, L.; Ma, X. Functional Genomic Analysis of Breast Cancer Metastasis: Implications for Diagnosis and Therapy. *Cancers* **2021**, *13*, No. 3276.

(51) Qiu, N.; He, Y. F.; Zhang, S. M.; Zhan, Y. T.; Han, G. D.; Jiang, M.; He, W. X.; Zhou, J.; Liang, H. L.; Ao, X.; Xia, H. M.; Li, J.; Yang, Y. Y.; He, Z. M.; Zou, Z. Z.; Li, H. S. Cullin7 enhances resistance to trastuzumab therapy in Her2 positive breast cancer via degrading IRS-1 and downregulating IGFBP-3 to activate the PI3K/AKT pathway. *Cancer Lett.* **2019**, *464*, 25–36.

(52) Reuter, S.; Gupta, S. C.; Chaturvedi, M. M.; Aggarwal, B. B. Oxidative stress, inflammation, and cancer: how are they linked? *Free Radical Biol. Med.* **2010**, *49*, 1603–1616.

(53) Hecht, F.; Pessoa, C. F.; Gentile, L. B.; Rosenthal, D.; Carvalho, D. P.; Fortunato, R. S. The role of oxidative stress on breast cancer development and therapy. *Tumour Biol.* **2016**, *37*, 4281–4291.

(54) Di Nunno, V.; Franceschi, E.; Tosoni, A.; Gatto, L.; Bartolini, S.; Brandes, A. A. Glioblastoma Microenvironment: From an Inviolable Defense to a Therapeutic Chance. *Front. Oncol.* **2022**, *12*, No. 852950.

(55) Zhao, F.; Gao, S.; Qin, X.; Niu, R.; Li, Z.; Wang, C.; Li, S. Comprehensive Analysis of TRP Channel-Related Genes for Estimating the Immune Microenvironment, Prognosis, and Therapeutic Effect in Patients With Esophageal Squamous Cell Carcinoma. *Front. Cell Dev. Biol.* **2022**, *10*, No. 820870.

(56) Xiao, Y.; Cong, M.; Li, J.; He, D.; Wu, Q.; Tian, P.; Wang, Y.; Yang, S.; Liang, C.; Liang, Y.; Wen, J.; Liu, Y.; Luo, W.; Lv, X.; He, Y.; Cheng, D. D.; Zhou, T.; Zhao, W.; Zhang, P.; Zhang, X.; Xiao, Y.; Qian, Y.; Wang, H.; Gao, Q.; Yang, Q. C.; Yang, Q.; Hu, G. Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation. *Cancer Cell* **2021**, *39*, 423–437.e7.