Observational Study

Medicine

OPEN

Analysis of hydrogen sulfide-related signatures of breast cancer identified 2 distinct subtypes

Implications for individualized therapeutics

Congyan Yu, BD[a], Yong Shen, PhD[b], Xuen Li, BD[a,*]

Abstract

Hydrogen sulfide (H_2S) is a vital gasotransmitter involved in breast cancer (BC) pathogenesis. This study aims to employ hydrogen sulfide-related genes (HSRGs) for molecular classification of BC and, accordingly, to establish a robust prognostic risk signature. Transcriptomic, clinical, and mutational data of BC patients were collected from the cancer genome atlas and gene expression omnibus databases. Prognostic relevance was evaluated using Cox regression analysis, while consensus clustering analysis was employed for molecular subtyping. Gene expression profiles, prognosis, immune infiltration patterns, drug sensitivity, and response to immunotherapy were compared between subtypes. Multiple-gene prognostic features were developed and assessed along with a nomogram. The gene expression was validated in clinical samples using quantitative polymerase chain reaction. Among 282 HSRGs, 46 exhibited significant correlations with BC prognosis. Consensus clustering identified 2 distinct molecular subtypes (C1 and C2). C1 displayed significantly improved prognosis compared to C2, accompanied by increased infiltration of B cells, T cells, monocytes, and mast cells but decreased macrophage infiltration. Moreover, C1 demonstrated higher drug sensitivity and immunotherapeutic response relative to C2. Enrichment analysis revealed suppressed immune-related processes and pathways in C2 while cell cycle regulation and chromosomal processes were significantly activated. Additionally, a risk feature comprising 6 differentially expressed genes between subtypes was constructed; this feature performed well in prognostic prediction. Integration of this feature with other clinical parameters (radiotherapy/chemotherapy status, clinical stage, N stage) into a nomogram further enhanced prognostic accuracy. Clinical samples further validated the high expression of ATP13A5, LRTM2, MAFA, and SPDYC and the low expression of CYP4F12 and TNN in BC. Our findings highlight the clinical relevance of HSRGs in BC, providing a basis for precise molecular classification and prognosis evaluation. The developed risk feature and nomogram offer practical tools for guiding personalized treatment strategies in clinical practice.

Abbreviations: ATP13A5 = ATPase 13A5, ATR = ATR serine/threonine kinase, AUC = area under the curve, BC = breast cancer, BMDMs = bone marrow-derived macrophages, CAF = cancer-associated fibroblast, CCLE = cancer cell line encyclopedia, CDF = cumulative distribution function, CDH1 = cadherin 1, CREB1 = cAMP responsive element binding protein 1, CTH = cystathionine gamma-lyase, CXCL14 = C-X-C motif chemokine ligand 14, CYP4F12 = cytochrome P450 family 4 subfamily F member 12, DEGs = differentially expressed genes, FCGBP = Fc gamma binding protein, FDR = false discovery rate, FPKM = fragments per kilobase millio, GEO = gene expression omnibus, H_2S = hydrogen sulfide, HSRGs = H_2S-related genes, IPS = immunophenoscore, KLF4 = KLF transcription factor 4, LRTM2 = leucine rich repeats and transmembrane domains 2, MAFA = MAF bZIP transcription factor A, MAP2K1 = ltogen-activated protein kinase kinase 1, MAPK14 = mitogen-activated protein kinase 14, NOS1 = Nitric oxide synthase 1, NRF2 = Nuclear factor erythroid 2-related factor 2, OTUB1 = OTU deubiquitinase, ubiquitin aldehyde binding 1, SLC7A11 = Solute carrier family 7 member 11, SPDYC = speedy/RINGO cell cycle regulator family member C, TCGA = the cancer genome atlas, TCIA = the cancer immunome atlas, TIDE = tumor immune dysfunction and exclusion, TNN = tenascin N.

Keywords: breast cancer, hydrogen sulfide, immunotherapy, nomogram, prognosis

This study was supported by the Ningbo Science and Technology Development Special Fund (2022S076).

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Supplemental Digital Content is available for this article.

[a] Department of General Surgery, Longshan Hospital of Cixi City, Ningbo, Zhejiang, P.R. China, [b] Department of Surgical Oncology, The Second Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, Zhejiang, P.R. China.

** Correspondence: Xuen Li, Department of General Surgery, Longshan Hospital of Cixi City, No. 1200 Lingfeng Road, Ningbo 315201, Zhejiang, P.R. China (e-mail: lxe20085280@126.com).*

How to cite this article: Yu C, Shen Y, Li X. Analysis of hydrogen sulfide-related signatures of breast cancer identified 2 distinct subtypes: Implications for individualized therapeutics. Medicine 2025;104:14(e42076).

Received: 9 October 2023 / Received in final form: 19 March 2025 / Accepted: 20 March 2025

http://dx.doi.org/10.1097/MD.0000000000042076

1. Introduction

Breast cancer (BC) is the most prevalent malignancy among women and a leading cause of cancer-related mortality.[1,2] Recent estimates suggest that approximately 2.3 million individuals are diagnosed with BC annually, resulting in around 450,000 deaths each year.[3,4] The incidence of BC continues to rise steadily, posing a significant threat to women's health. Despite ongoing research efforts, the etiology of BC remains incompletely understood. However, several risk factors have been identified, including age, family history, hormone levels, and lifestyle choices.[5–7] Early detection and timely treatment play a crucial role in improving the survival rates of BC patients.[8] Substantial advancements have been made in screening techniques, diagnostic methods, and treatment strategies for BC within the medical field.[9] Nonetheless, numerous challenges and concerns persist. For instance, accurately diagnosing early-stage BC remains challenging; certain high-risk groups lack effective preventive measures; and personalized treatment plans are needed to enhance efficacy.[10,11] Therefore, exploring more effective biomarkers could help address these issues.

Endogenous hydrogen sulfide (H_2S) primarily originates from cystathionine beta-synthase, cystathionine gamma-lyase (CTH), and 3-mercaptopyruvate sulfurtransferase.[12] H_2S serves as a vital signaling gasotransmitter involved in regulating various physiological and pathological processes such as gastrointestinal diseases,[13] cardiovascular diseases,[14] oxidative stress-related diseases,[15] and inflammation.[16] Furthermore, studies have indicated that endogenous H_2S plays a role in the occurrence and progression of cancers.[17,18] In terms of cellular protective responses in cancer research, H_2S is considered a bidirectional target. On 1 hand, endogenous H_2S and low levels of exogenous H_2S promote angiogenesis, accelerate the cell cycle, inhibit apoptosis, and ultimately facilitate tumor formation.[19–21] On the other hand, H_2S donors stimulate high concentrations of endogenous H_2S production to prevent tumor formation.[22,23] In BC, inhibiting endogenous H_2S production significantly could suppress cancer cell growth,[24] suggesting that targeting endogenous H_2S may hold promise as a therapeutic approach for BC cells. Therefore, deciphering the expression patterns of hydrogen sulfide-related genes (HSRGs) can provide insights into the signaling strength of H_2S in BC and offer potential targets for H_2S-based treatment strategies.

This study aimed to explore the prognostic relevance of HSRGs in BC, establish molecular subtypes based on these genes, and identify potential targets for personalized treatment strategies. The workflow schematic of the present study was depicted in Figure 1.

2. Materials and methods

2.1. Data acquisition and preprocessing

The mRNA expression profiles along with the corresponding clinicopathologic information (including TNM staging, clinical staging details, as well as chemotherapy and radiotherapy records) pertaining to the cancer genome atlas (TCGA)-breast invasive carcinoma (BRCA) cohort were procured from TCGA (https://portal.gdc.cancer.gov/, accessed September 20, 2023). Standardized fragments per kilobase million data were obtained and further log2 transformed before using the data for model training and testing. After excluding cases with insufficient prognostic follow-up periods of <30 days and those lacking complete clinicopathologic information, a total of 869 BC patients were deemed eligible and thus incorporated into the study.

As a validation set, the GSE96058 dataset comprising 3409 BC patients was downloaded from the gene expression omnibus

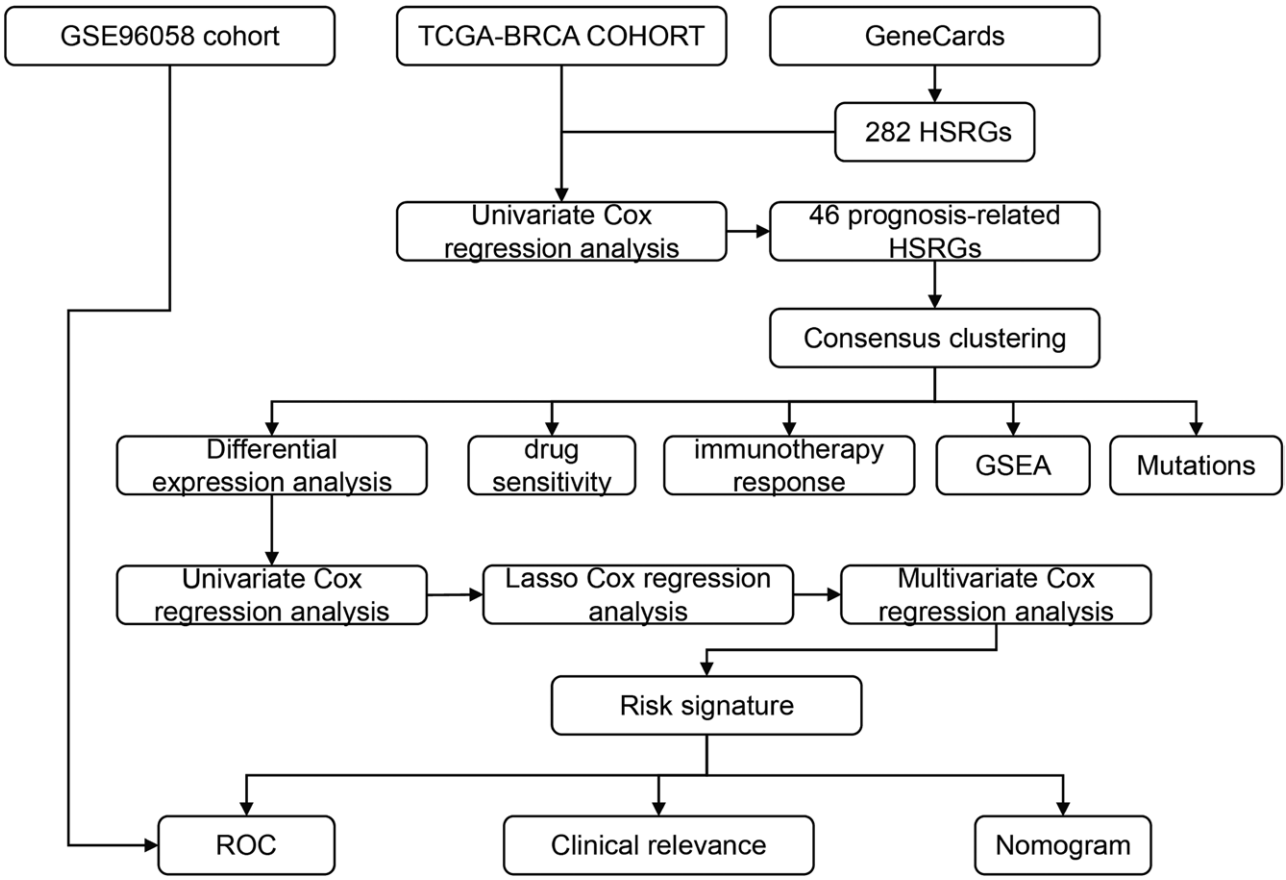

Figure 1. The workflow of the present study.

(GEO, https://www.ncbi.nlm.nih.gov/geo/) database. To account for potential technical variations between datasets, batch correction was performed using the ComBat algorithm from the sva R package. The clinicopathologic features of cohorts were summarized in Table 1. The GSE93601 and GSE14999 datasets were retrieved from the GEO database to validate the expression patterns of signature-related genes. Additionally, gene expression data for BC cell lines and noncancerous cell lines were obtained from the cancer cell line encyclopedia (CCLE, https://sites.broadinstitute.org/ccle/datasets) database for further validation and comparison.

Hydrogen sulfide-related genes were collected from the GeneCards database (https://www.genecards.org/, accessed September 20, 2023), yielding 282 genes (Table S1, Supplemental Digital Content, https://links.lww.com/MD/O639).

2.2. Consensus clustering

The association between HSRGs and BC prognosis was assessed through univariate Cox regression analysis. Genes exhibiting a significance level of $P < .05$ were subjected to Consensus clustering analysis. The ConsensusClusterPlus R package was employed for this purpose, utilizing the partitioning around medoids clustering algorithm and Euclidean distance metric. Each sampling iteration involved 89% of the dataset, with a maximum cluster number set at 6. Subsequently, the TCGA-BRCA cohort underwent molecular subtyping based on the aforementioned analysis, with determination of the optimal number of clusters relying on item-consensus plot and relative changes observed in the cumulative distribution function (CDF) curve. The CDF curve represented the cumulative proportion of sample pairs assigned to the same cluster across different clustering iterations.

Table 1

The clinicopathologic features of the TCGA-BRCA and GSE96058 cohorts.

	TCGA-BRAC cohort (n = 869)	GSE96058 cohort (n = 3409)
Age		
<60	469	1284
≥60	400	2125
Outcomes		
Alive	779	3056
Dead	90	353
T stage		
T1	239	
T2	488	
T3	114	
T4	28	
N stage		
N0	410	
N1	299	
N2	91	
N3	69	
M stage		
M0	726	
M1	143	
Chemotherapy		
Yes	661	2017
No	208	1371
N/A	0	21
Radiotherapy		
Yes	468	
No	401	
Clinical stage		
I	163	
II	491	
III	198	
IV	17	

BRCA = breast invasive carcinoma.

2.3. Gene set enrichment analysis

Gene set enrichment analysis was conducted utilizing the clusterProfiler R package to ascertain notable disparities in biological processes and Kyoto Encyclopedia of Genes and Genomes pathways across distinct subtypes. The Benjamini–Hochberg correction method was implemented, whereby a significance level of $P < .05$ denoted statistical significance.

2.4. Immune infiltration analysis

Utilizing the CIBERSORT R package, an immune infiltration analysis was conducted to evaluate the extent of infiltration by 22 distinct immune cell types within tumor tissues of TCGA-BRCA. This investigation employed standardized transcriptomic data derived from BC samples. Furthermore, a comparative analysis was performed to ascertain variations in immune infiltration across different molecular subtypes.

2.5. Treatment response analysis

A total of 8 commonly used chemotherapy drugs were carefully selected to assess the drug sensitivity in BC patients, employing the sophisticated pRRophetic R package. To ensure data integrity, batch correction was meticulously performed using the ComBat algorithm. The evaluation of immunotherapy response was conducted by leveraging both the tumor immune dysfunction and exclusion (TIDE) algorithm as well as the immunophenoscore (IPS). Specifically, the TIDE algorithm, accessible at https://tide.dfci.harvard.edu/, was employed to derive comprehensive scores including TIDE, cancer-associated fibroblast (CAF), Merck18, Dysfunction, and Exclusion through the importation of standardized transcriptomic data. Additionally, IPS scores were acquired via The cancer immunome atlas online tool (https://tcia.at/home), utilizing TCGA-BRCA patient IDs as input. Ultimately, a comparative analysis was conducted to elucidate disparities in drug sensitivity and immunotherapy response among distinct molecular subtypes.

2.6. Mutation analysis

The somatic mutation data of BC patients were acquired from the TCGA database and subjected to analysis and visualization using the maftools R package. Oncoplots were generated to visualize the prognostically relevant HSRGs ($P < .05$) and the top 10 most frequently mutated genes in different subtypes. The differences in tumor mutational burden between subtypes were compared.

2.7. Differential expression analysis

Differential expression analysis between molecular subtypes was conducted using the limma R package. Linear models were fitted using the lmFit function, followed by empirical Bayes moderation with the eBayes function to compute moderated *t*-statistics and *P*-values. Multiple hypothesis testing correction was performed using the Benjamini–Hochberg procedure to control false discovery rate. Differentially expressed genes (DEGs) were defined by thresholds of log2-fold change > 1 and adjusted *P*-value < .05. Visualization of DEGs was accomplished through volcano plots.

2.8. Risk feature construction and evaluation

Univariate cox regression analysis was employed to identify DEGs between subtypes that are significantly associated with BC prognosis ($P < .05$). To mitigate overfitting, LASSO Cox regression analysis was utilized to reduce the number of prognostic genes (coefficient ≠ 0). Finally, multivariate Cox

regression analysis was applied to select independent prognostic genes ($P < .05$), and the risk score was computed according to the following formula: riskscore $=\sum(\beta i * coef i)$, ("i" denotes the count of prognostic genes, "βi" represents the expression level of genes, "coefi" signifies the coefficient assigned to genes). BC patients were divided into high-risk and low-risk cohorts based on their median risk score values. Survival analysis was performed utilizing both the survival and survminer R packages, while receiver operating characteristic (ROC) curves were employed to assess the predictive performance of the risk score in terms of prognosis evaluation.

2.9. Nomogram construction and evaluation

Univariate and multivariate Cox regression analyses were conducted to assess the correlation between riskscore, clinical–pathological features, and prognosis of BC. The nomogram model was constructed using the rms R package. Calibration curves, decision curves, and ROC curves were employed to evaluate the prognostic capabilities of the nomogram.

2.10. Clinical validation

Ten pairs of BC tissues and their adjacent non-tumor tissues were obtained from the Longshan Hospital of Cixi City (Table S2, Supplemental Digital Content, https://links.lww.com/MD/O640). Total RNA was extracted using the RNeasy Mini Kit (QIAGEN, Germany) according to the manufacturer's instructions. Subsequently, reverse transcription was performed using the high-capacity cDNA synthesis kit (Takara, China). After reverse transcription, mRNA expression levels were measured by quantitative real-time PCR. All measurements were performed in biological triplicates, and each experiment was independently repeated 3 times to ensure reproducibility. The primers used for experiments are listed in Table 2. Relative gene expression levels were calculated using the $2^{\Delta\Delta Ct}$ method.

2.11. Statistical analysis

Data analysis and statistical procedures were carried out employing R version 4.2.2, and key R packages used in this study included ConsensusClusterPlus, clusterProfiler, CIBERSORT, pRRophetic, ggplot2, maftools, limma, survival, survminer, rms, and sva. The comparison of between-group disparities was accomplished utilizing the Wilcoxon test, whereas survival analysis was executed through the utilization of Kaplan–Meier plots and log-rank tests. A statistically significant distinction was determined as a P-value <.05.

3. Results

3.1. Association between HSRGs and prognosis of breast cancer

Univariate Cox regression analysis demonstrated that among the 282 HSRGs, a total of 46 genes exhibited significant associations with the prognosis of BC. Specifically, there were 22 risk genes (Hazard ratio > 1) and 24 protective genes (Hazard ratio < 1) identified (Fig. 2A). The genomic positions of these genes are depicted in Figure 2B. Among these prognostic genes, 20 were found to harbor mutations, with cadherin 1 (CDH1) mutations being the most prevalent and diverse, followed by Fc gamma binding protein (FCGBP) and ATR serine/threonine kinase (ATR) (Fig. 2C).

3.2. Molecular subtyping of breast cancer based on prognostic HSRGs

Moreover, we conducted consensus clustering analysis on the TCGA-BRCA cohort utilizing the aforementioned prognostic-associated HSRGs, leading to the identification of 2 distinct molecular subtypes of BC (Fig. 3A–C). Significant differences were observed in the expression of prognostic-associated HSRGs between the C1 and C2 subtypes (Fig. S1A, Supplemental Digital Content, https://links.lww.com/MD/O641). Notably, 6 of these HSRGs also exhibited significant differential expression between patients at early (I/II) and advanced (III/IV) stages (Fig. S1B, Supplemental Digital Content, https://links.lww.com/MD/O641). Subsequent survival analysis revealed significant disparities in prognosis between these 2 subtypes (Fig. 3D, $P = .00021$). Additionally, immune infiltration analysis showcased elevated levels of infiltrating immune cells, including naive B cells, CD8 T cells, resting CD4 memory T cells, follicular helper T cells, resting/activated NK cells, monocytes, M1 macrophages, resting dendritic cells, and resting mast cells within the favorable prognosis C1 subtype compared to the C2 subtype. Conversely, M0/M2 macrophages exhibited lower abundance in the C1 subtype relative to the C2 subtype (Fig. 3E). The sankey diagram illustrated the relationships between different subtypes and clinical-pathological characteristics (Fig. 3F). Furthermore, correlation analysis demonstrated a significant association between these prognostic-associated HSRGs and immune cell infiltration (Fig. 3G).

3.3. Differential therapeutic responses between HSRGs-derived molecular subtypes

The sensitivity of the C1 and C2 subtypes to various conventional chemotherapy agents was examined. The findings indicated that the C1 subtype displayed enhanced sensitivity towards these drugs relative to the C2 subtype, implying a greater likelihood of therapeutic efficacy (Fig. 4A). Furthermore, an evaluation was conducted to investigate the potential responsiveness of the C1 and C2 subtypes towards immunotherapy, utilizing IPS and TIDE algorithms. Notably, the outcomes revealed that the C1 subtype exhibited elevated IPS scores (Fig. 4B), as well as heightened scores for CAF, Merck18, and dysfunction (Fig. 4C–G), in comparison to the C2 subtype. Elevated IPS scores are typically indicative of robust immune responses and may correlate with improved immunotherapeutic outcomes, suggesting a potentially heightened immune response within the C1 subtype.

3.4. Biological processes, pathways, and mutation characteristics differences between HSRGs-derived molecular subtypes

Gene set enrichment analysis unveiled noteworthy attenuation of immune-related processes within the C2 subtype relative to the C1 subtype, encompassing antigen receptor-mediated signaling pathway, humoral immune response mediated by circulating immunoglobulin, complement activation, and B cell receptor signaling pathway. Conversely, chromosome-associated processes were substantially upregulated in the C2 subtype (Fig. 5A). Additionally, significant suppression of pathways such as cytokine-cytokine receptor interaction and Th17 cell differentiation was observed in the C2 subtype, while pathways including cell cycle and neutrophil extracellular trap formation were markedly activated (Fig. 5B). The top 10 most frequently mutated genes in the C1 and C2 subtypes are illustrated in Figures 5C and D, respectively. Notably, CDH1 exhibited frequent mutations solely within the C1 subtype but not in the C2 subtype. Compared to the C1 subtype, C2 subtype showed significant higher tumor mutational burden (Fig. 5E).

3.5. Construction of multi-gene risk features for breast cancer

Differential expression analysis successfully identified a total of 559 genes exhibiting differential expression between the

Table 2

The sequences of the primers used in this study.

Gene symbol	Forward (5'-3')	Reverse (5'-3')
ATP13A5	CATCGGGCTTTGCTCAACC	GCTTTCCGTACATTGTGGTCC
CYP4F12	TCCAGTGTTTCCCACAGC	CGATGAAGGGGATGATGG
LRTM2	CAGGTCGGGGATAACCCCT	CCCTCGGTAGGAGAACCACT
MAFA	GCTTGGAGGAGCGCTTCTCCGACG	CTGCTCCACCTGGCTCTGGAGCTG
SPDYC	TCAGCCTTCTGGAGGACAGT	CACCATGGCCAGGAGATACT
TNN	CAGTGGGAGCAGCAGGCAGAC	GTATGGACGTTGTGGATTTCAGTA
GAPDH	GGTGAAGGTCGGTGTGAACGGA	TGTTAGTGGGGTCTCGCTCCTG

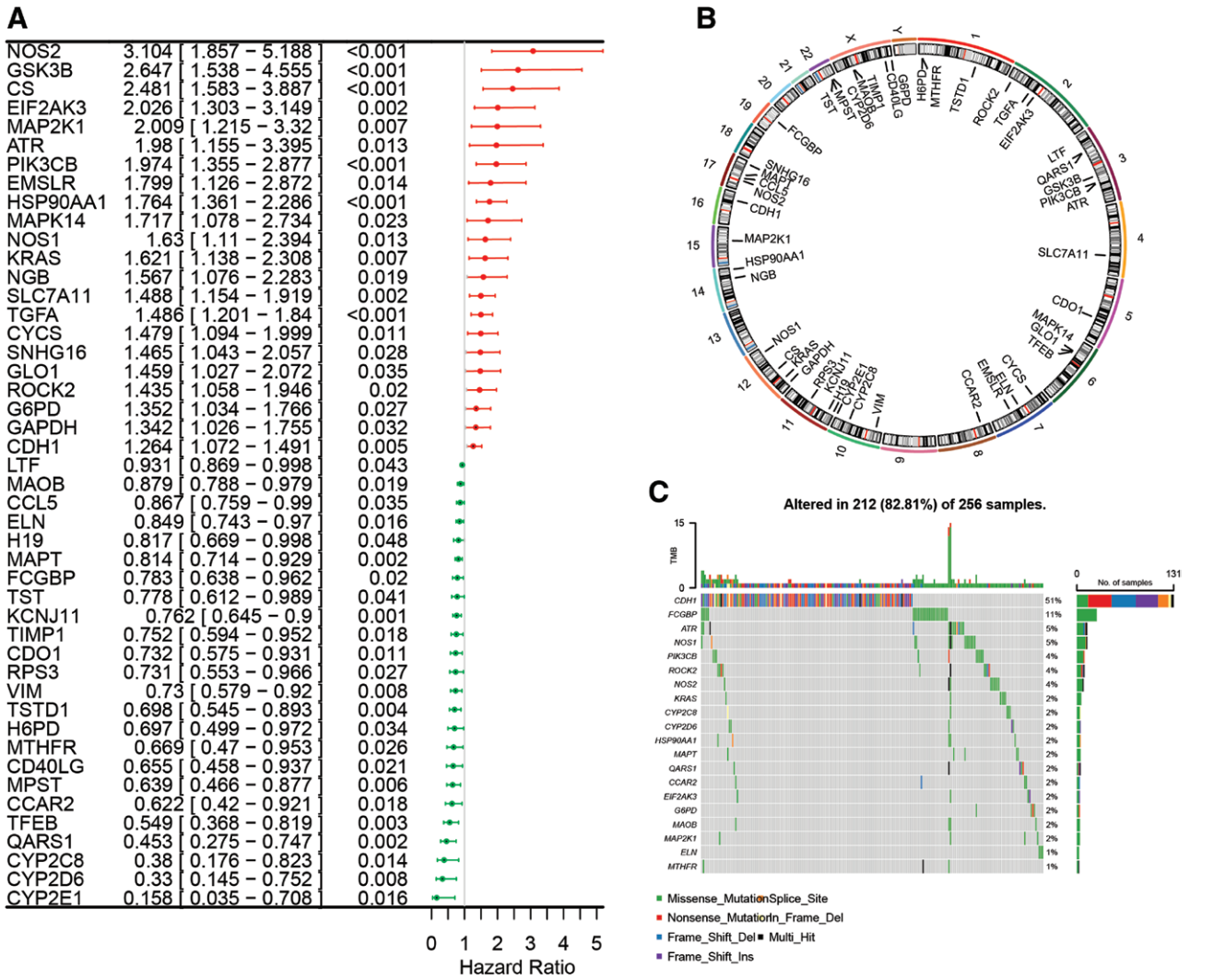

Figure 2. Expression and mutations of HSRGs in BC. (A) Univariate cox regression analysis identified HSRGs associated with BC prognosis. (B) Chromosomal localization of HSRGs significantly associated with the prognosis of BC. (C) Mutations of prognostic HSRGs in BC. BC = breast cancer, HSRGs = hydrogen sulfide-related genes.

C1 and C2 subtypes, with 468 genes being downregulated and 91 genes upregulated (Fig. 6A). Subsequently, univariate Cox regression analysis was conducted to assess the association between these DEGs and BC prognosis. Remarkably, 56 genes displayed significant correlations with BC prognosis ($P < .05$, Table S3, Supplemental Digital Content, https://links.lww.com/MD/O642). To establish prognostic risk features, LASSO Cox regression analysis was performed, resulting in the identification of a set of 23 prognostic genes (Fig. 6B and C). Finally, through multivariate cox regression analysis, 6 genes, including ATPase 13A5 (ATP13A5), cytochrome P450 family 4 subfamily F member 12 (CYP4F12), leucine rich repeats and transmembrane domains 2 (LRTM2), MAF bZIP transcription factor A (MAFA), speedy/RINGO cell cycle regulator family member C (SPDYC), and tenascin N(TNN), were incorporated into the construction of the risk feature (Fig. 6D). The final risk score formula is as follows: riskscore = 0.3668871 * ATP13A5–0.8047951 * CYP4F12 + 0.3826746 * LRTM2 + 1.2772720 * MAFA + 0.1688637 * SPDYC –0.3079934 * TNN.

3.6. Assessment of the prognostic risk feature

In the TCGA-BRCA and GSE96058 cohorts, patients were stratified into high-risk and low-risk groups based on risk

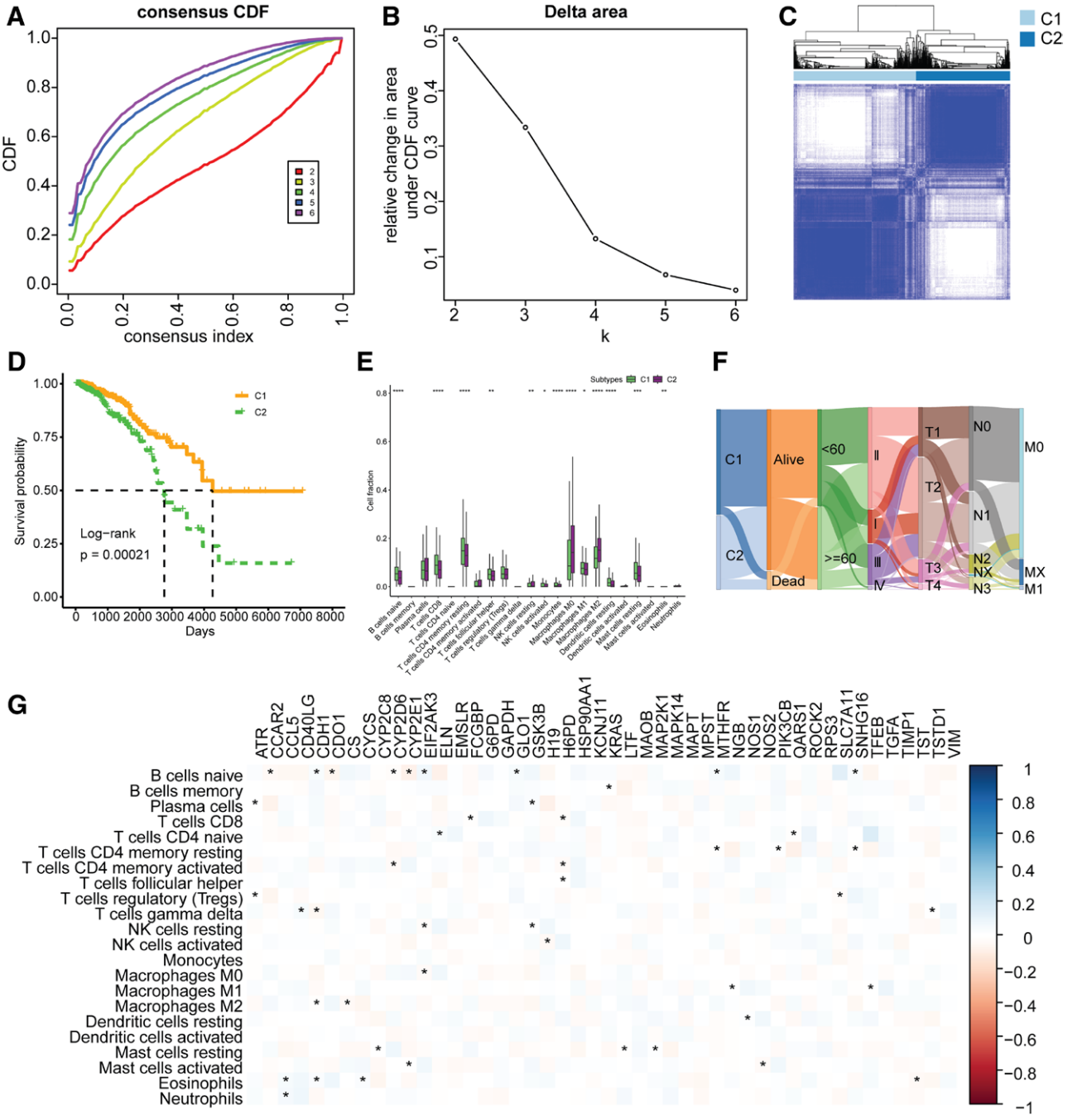

Figure 3. Molecular subtyping based on prognostic HSRGs in BC using consensus clustering method. (A) Consensus CDF curves for different subtypes, showing the distribution of a specific metric across the consensus index. (B) Relative change in area under the CDF curve as a function of k, indicating the sensitivity of the CDF to changes in k. (C) Heatmap representing the expression levels of genes in subtypes C1 and C2. The dendrogram at the top shows the hierarchical clustering of samples. (D) Survival analysis between molecular subtypes derived from HSRGs in BC. (E) Differences in immune cell infiltration between C1 and C2 subtypes were identified by the Wilcoxon test. (F) The sankey diagram illustrated the relationships between different subtypes and clinical-pathological characteristics. (G) Correlation between prognostic HSRGs and immune cell infiltration were identified by the pearson correlation test. * $P < .05$, ** $P < .01$, *** $P < .001$, **** $P < .0001$. BC = breast cancer, CDF = cumulative distribution function, HSRGs = hydrogen sulfide-related genes.

scoring, as depicted in Figure 7A. The expression levels of 6 genes are illustrated in Figure 7B. Survival analysis revealed that patients in the low-risk group exhibited significantly superior prognosis compared to those in the high-risk group, both in the TCGA-BRCA cohort and GSE96058 cohort (Fig. 7C and D, $P < .0001$). ROC curve analysis demonstrated that the area under the curve (AUC) of this risk score for predicting overall survival at 1, 3, and 5 years among BC patients in the TCGA-BRCA cohort were calculated as 0.781, 0.776, and 0.758, respectively (Fig. 7E). Similarly, in the GSE96058 cohort, the AUC of this risk model for predicting overall survival at 1, 3, and 5 years among BC patients were determined to be 0.753, 0.664, and 0.642 correspondingly (Fig. 7F). Compared to the high-risk groups, the low-risk group exhibited significantly elevated expression of TNN, CYP4F12 and down-regulated expression of *SPDYC*, MAFA, LRTM2, and ATP13A5 (Fig. 7G and H) in all cohorts. These findings strongly suggest that this risk feature exhibits excellent performance in prognostic evaluation of BC patients.

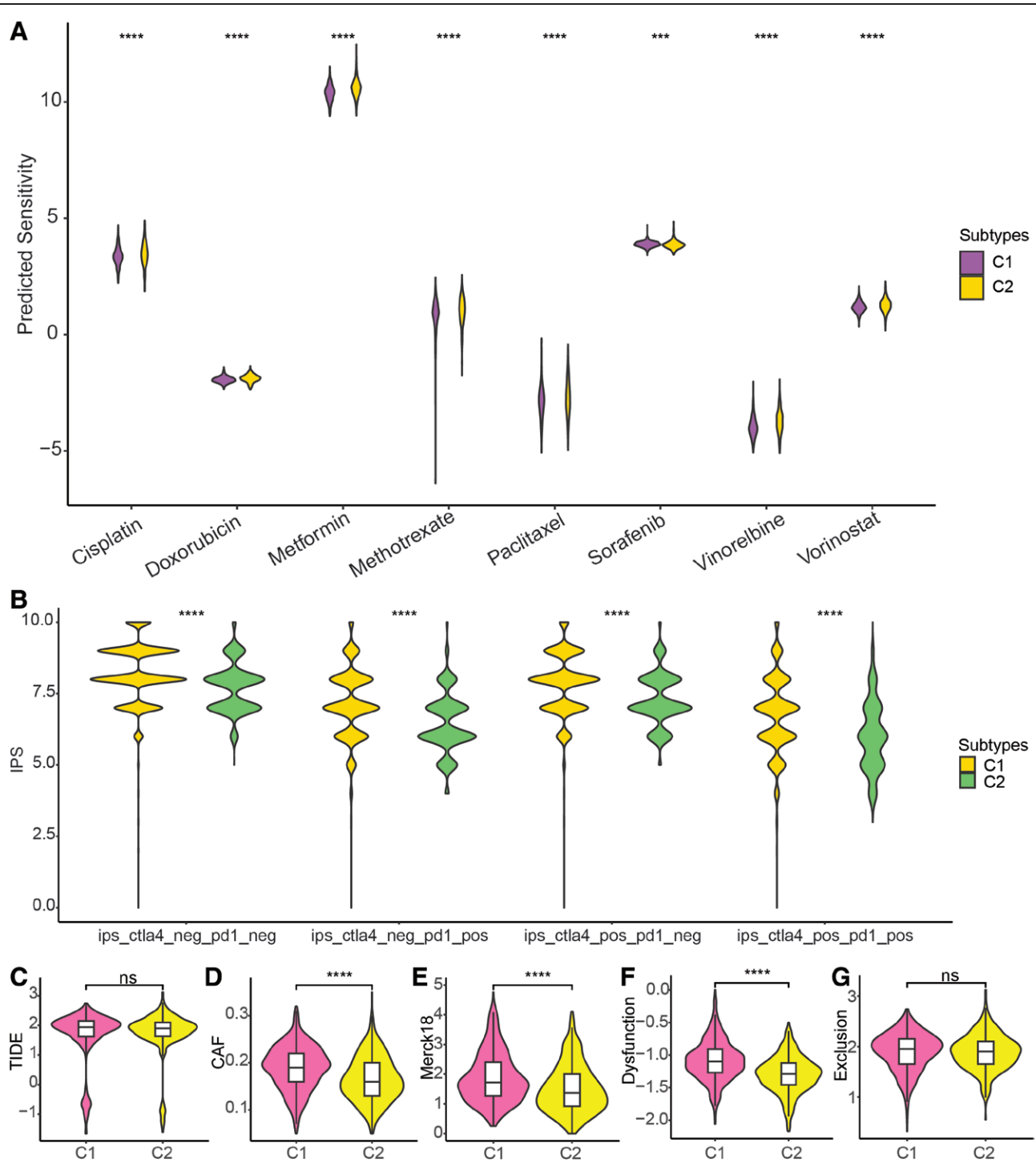

Figure 4. Differences in treatment response among HSRGs-derived molecular subtypes were identified by the Wilcoxon test. (A) Differential sensitivity to chemotherapy drugs among molecular subtypes derived from HSRGs. (B) Differences in IPS scores between C1 and C2 subtypes. Comparison of TIDE (C), CAF (D), Merck18 (E), dysfunction (F), and exclusion scores (G) between C1 and C2 subtypes. ns: not significant, * $P < .05$, ** $P < .01$, *** $P < .001$, **** $P < .0001$. CAF = cancer-associated fibroblast, HSRGs = hydrogen sulfide-related genes, IPS = immunophenoscore, TIDE = tumor immune dysfunction and exclusion.

3.7. Clinical relevance of the prognostic risk features

Figure 8A illustrated the expression of biomarker genes with clinicopathologic annotations. In order to assess the association between the aforementioned risk feature and clinical-pathological characteristics in BC patients, a comparative analysis was conducted on variations in risk scores across different age groups, survival outcomes, radiotherapy/chemotherapy status, clinical stages, and TNM stages. The findings illustrated in Figure 8B indicated that older patients (aged ≥ 60) exhibited higher risk scores compared to their younger counterparts. Furthermore, deceased patients demonstrated higher risk scores relative to surviving patients. Additionally, an elevated clinical stage was associated with higher risk scores when compared to lower clinical stages. Conversely, no significant disparities were observed in risk scores among different radiotherapy groups and TNM stage groups.

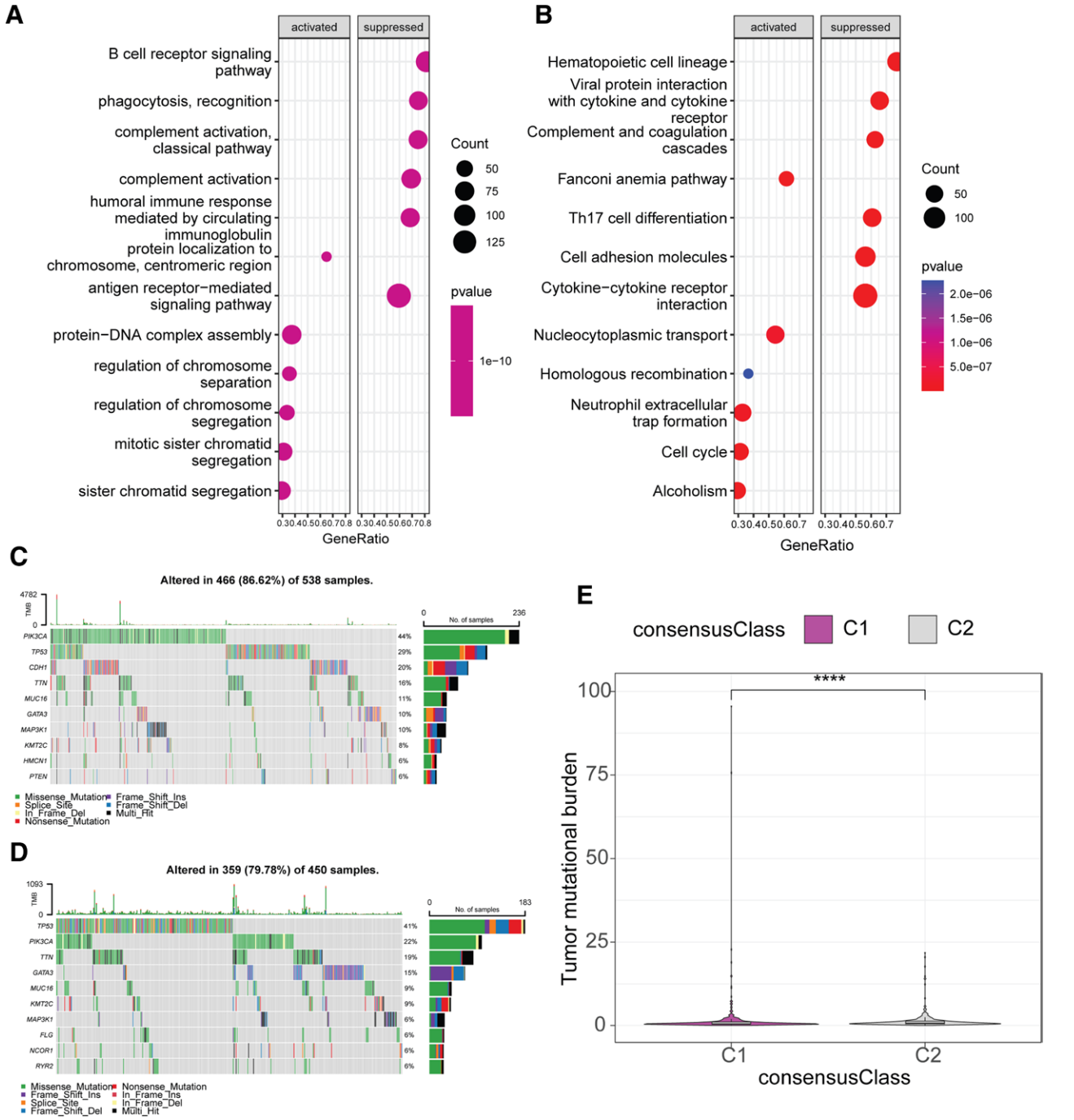

Figure 5. Differences in the biological processes, KEGG pathways, and gene mutation characteristics between C1 and C2 subtypes. (A) Biological processes significantly activated and inhibited in C2 subtype compared to C1 subtype. (B) KEGG pathways significantly activated and inhibited in C2 subtype compared to C1 subtype. (C) Oncoplot showing the top 10 most frequently mutated genes in the C1 subtype. (D) Oncoplot showing the top 10 most frequently mutated genes in the C2 subtype. (E) Comparison of the difference in the tumor mutational burden between the C1 and C2 subtypes using the Wilcoxon test. **** $P < .0001$. KEGG = Kyoto Encyclopedia of Genes and Genomes.

3.8. Prognostic nomogram derived from the risk features

Univariate cox regression analysis revealed that the risk score, age, radiotherapy/chemotherapy status, clinical stage, and N stage exhibited prognostic significance in BC patients (Fig. 9A). Moreover, with the exception of age, these variables were identified as independent prognostic factors through multivariate cox regression analysis (Fig. 9B). Consequently, a prognostic nomogram model for BC assessment was constructed incorporating the risk score, radiotherapy/chemotherapy status, clinical stage, and N stage (Fig. 9C). The calibration curve demonstrated a high level of concordance between predicted overall survival and actual overall survival as determined by this nomogram (Fig. 9D), while the decision curve indicated that this nomogram provided greater net benefit compared to individual prognostic factors in predicting 1-year overall survival in BC cases (Fig. 9E). Furthermore, the ROC curve displayed AUC values of 0.87, 0.859, and 0.782 for predicting overall survival at 1, 3, and 5 years respectively in BC patients using this nomogram (Fig. 9F). These findings imply that the incorporation of risk features and clinical-pathological features into a nomogram holds potential value for prognostic evaluation of BC.

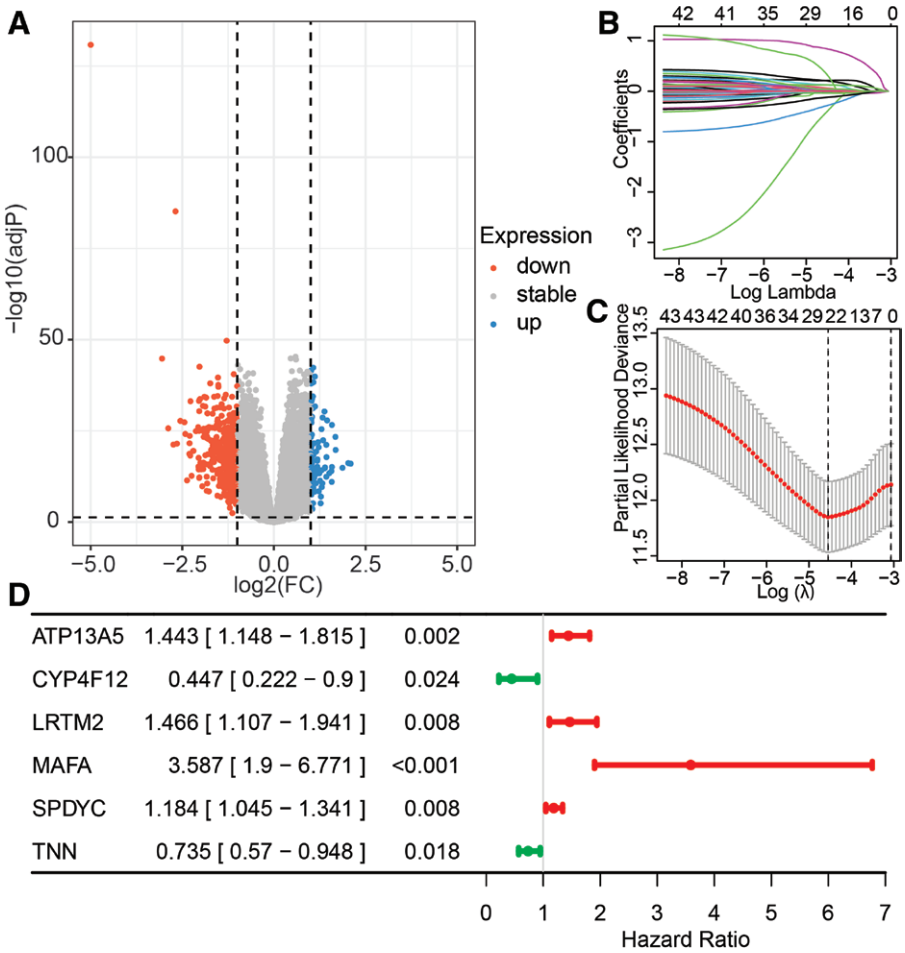

Figure 6. Development of multi-gene risk features for BC using the LASSO cox regression analysis. (A) Volcano plot showing the distribution of gene expression changes. Red dots indicate genes with significantly down-regulated expression, blue dots indicate significantly upregulated expression, and gray dots indicate stable expression. (B) Coefficient profiles as a function of the log (lambda) sequence. The vertical dashed line indicates the lambda value chosen by cross-validation. (C) Partial likelihood deviance as a function of the log(lambda) sequence. The red line indicates the deviance, and the shaded area represents the standard error. (D) Multivariate Cox regression analysis identified independent prognostic DEGs for constructing a risk score. BC = breast cancer, DEG = differentially expressed genes.

3.9. Validation of signature-related genes in external cohorts

To validate the expression of signature-related genes in BC, we initially retrieved the expression data of 6 signature-related genes from the CCLE database, comparing their levels between BC cell lines and noncancerous cell lines. The results demonstrated that these signature-related genes exhibited aberrant expression patterns in BC cell lines compared to noncancerous cell lines (Fig. 10A). Furthermore, our analysis of the GSE93601 cohort revealed that TNN and LRTM2 were significantly down-regulated in BC tissues compared to normal tissues. Similarly, in the GSE14999 cohort, *CYP4F12* and *TNN* showed reduced expression in tumor tissues relative to adjacent non-tumor tissues (Fig. 10B and C). Moreover, clinical sample validation revealed that the expression levels of 4 genes – ATP13A5, *LRTM2*, *MAFA*, and SPDYC – were significantly upregulated in BC tissues compared to adjacent non-tumor tissues, while the expression of CYP4F12 and TNN was significantly decreased (Fig. 10D). These findings collectively support the potential of these signature-related genes as candidate biomarkers for BC.

4. Discussion

Hydrogen sulfide is a gas molecule renowned for its distinctive biological activity and diverse physiological functions within the human body. Our research findings unveil the significant role of HSRGs in the heterogeneity of BC and personalized treatment strategies, which aligns with previously reported functions of hydrogen sulfide in cancer progression and immune regulation,[18,21,25] but also provides novel evidence that distinct molecular subtypes characterized by differential expression of HSRGs exhibit unique gene expression profiles, immune infiltration statuses, and therapeutic response properties. Unlike prior studies, our work not only identifies 2 unique molecular subtypes of BC but also underscores the distinctive value of HSRGs in prognosis assessment and treatment strategy formulation for BC patients.

The identification of these subtypes has broad clinical relevance. Subtype C1, characterized by improved prognosis, enhanced immune cell infiltration, and greater sensitivity to drugs and immunotherapy, suggests that patients in this group may benefit from more aggressive or targeted treatment strategies. Conversely, subtype C2, with its poorer prognosis and suppressed immune responses, may require alternative therapeutic approaches, such as combination therapies or personalized medicine tailored to overcome resistance mechanisms. The 6-gene prognostic risk signature we developed further enhances the clinical utility of HSRGs by providing a robust tool for predicting patient outcomes and guiding treatment decisions.

Within the context of HSRGs related to BC prognosis, certain members have been implicated in the modulation of endogenously produced H_2S, whereas others demonstrate

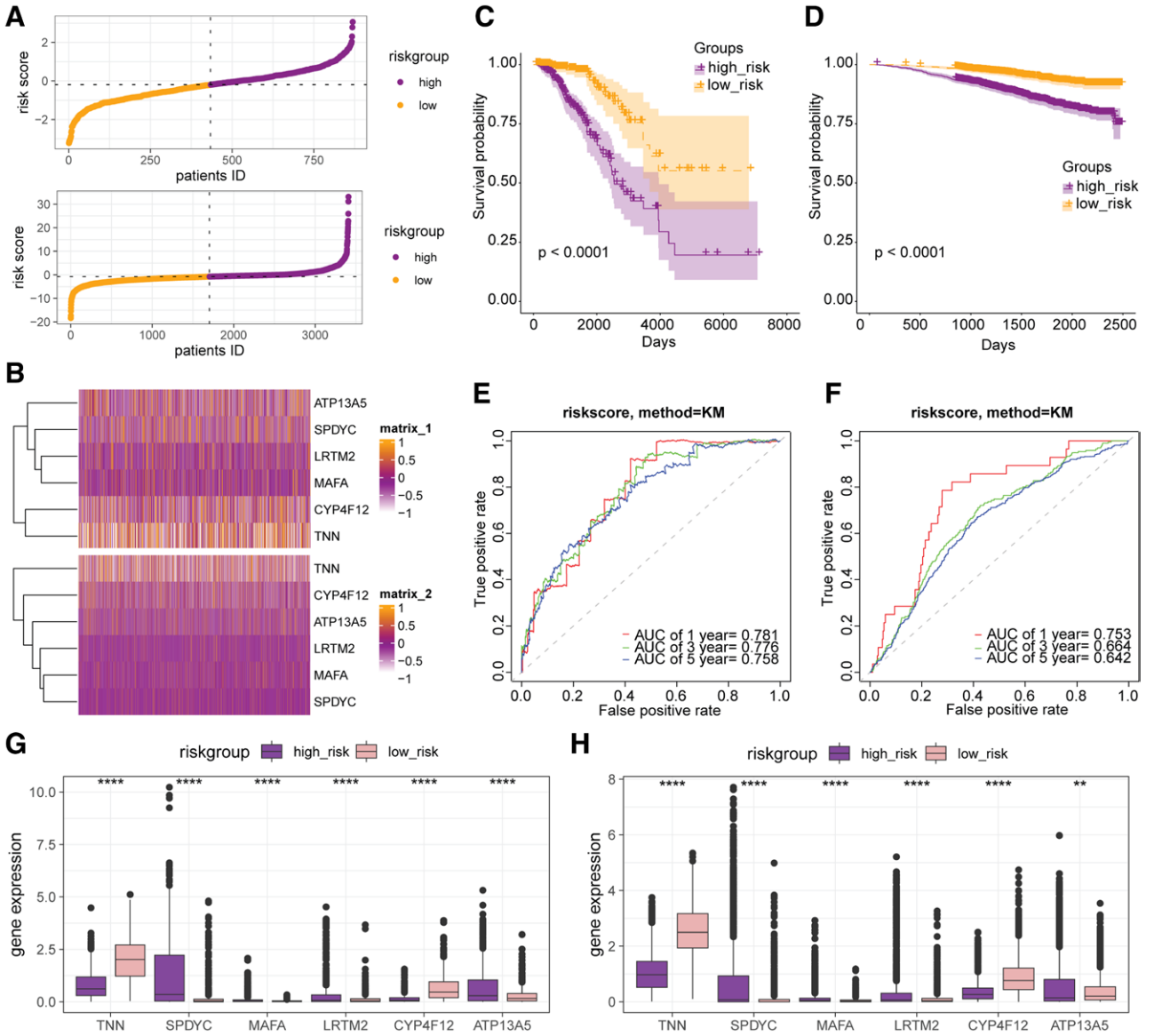

Figure 7. Evaluation of multi-gene risk features in BC. (A) Risk scores plotted against patient IDs, showing the distribution of risk scores across patients. (B) Heatmap representing the expression levels of selected genes across different samples. (C) Kaplan–Meier survival curve comparing high-risk and low-risk groups in the TCGA-BRCA cohort, with a log-rank test of $P < .0001$. (D) Kaplan–Meier survival curve comparing high-risk and low-risk groups in the GSE96058 cohort, with a log-rank test of $P < .0001$. ROC curves for the risk score model in the TCGA-BRCA (E) and GSE96058 (F) cohorts, showing the true positive rate vs the false positive rate at various thresholds. The AUC values for 1, 3, and 5 yr are provided. Comparison of the expression of biomarker genes between the high-risk and low-risk groups in the TCGA-BRCA (G) and GSE96058 cohorts (H) using the Wilcoxon test. Statistical significance is indicated by asterisks. ** $P < .01$, **** $P < .0001$. AUC = area under the curve, BC = breast cancer, BRCA = breast invasive carcinoma, ROC = receiver operating characteristic, TCGA = the cancer genome atlas.

responsiveness to exogenous H_2S in regulating cellular and systemic functionalities pertinent to the disease progression and outcome. For instance, there exists a mutual regulation between H_2S bioavailability and ATR kinase, whereby ATR inhibition decreases intracellular H_2S concentration, while the intracellular H_2S concentration modulates ATR kinase activity through ATR serine 435 phosphorylation.[26,27] Previous investigations have demonstrated that elevated transcription levels of *KRAS* are linked to unfavorable prognosis in BC,[28] and *KRAS* mutations activate nuclear factor erythroid 2-related factor 2 (NRF2), promoting *CTH* transcription.[29] Furthermore, the aberrant overexpression of solute carrier family 7 member 11 (SLC7A11) in BC is associated with poor prognosis. Studies have shown that exogenous H_2S donors upregulate the expression of SLC7A11, while endogenous H_2S regulates the stability of SLC7A11 through persulfidation at cysteine 91 mediated by *OTU* deubiquitinase, ubiquitin aldehyde binding 1 (OTUB1).[30,31] Consequently, the BC molecular subtypes derived from HSRGs may represent distinct levels of endogenous hydrogen sulfide or different patterns of response to exogenous hydrogen sulfide, resulting in significant phenotypic variations.

H_2S has been implicated in modulating the immune response within the tumor microenvironment, leading to immune cell suppression and facilitating immune evasion by tumors. Mitogen-activated protein kinase kinase 1 (MAP2K1) is identified as a prognostic risk gene in BC, and studies have demonstrated that inhibition of MAP2K1 strongly suppresses T cell proliferation and activation.[32] Conversely, in vitro stimulation with H_2S induces phosphorylation and activation of MAP2K1,[33] thereby promoting tumor immunity and inhibiting tumor progression. In bone marrow-derived macrophages, inhibition of phosphorylation of mitogen-activated protein

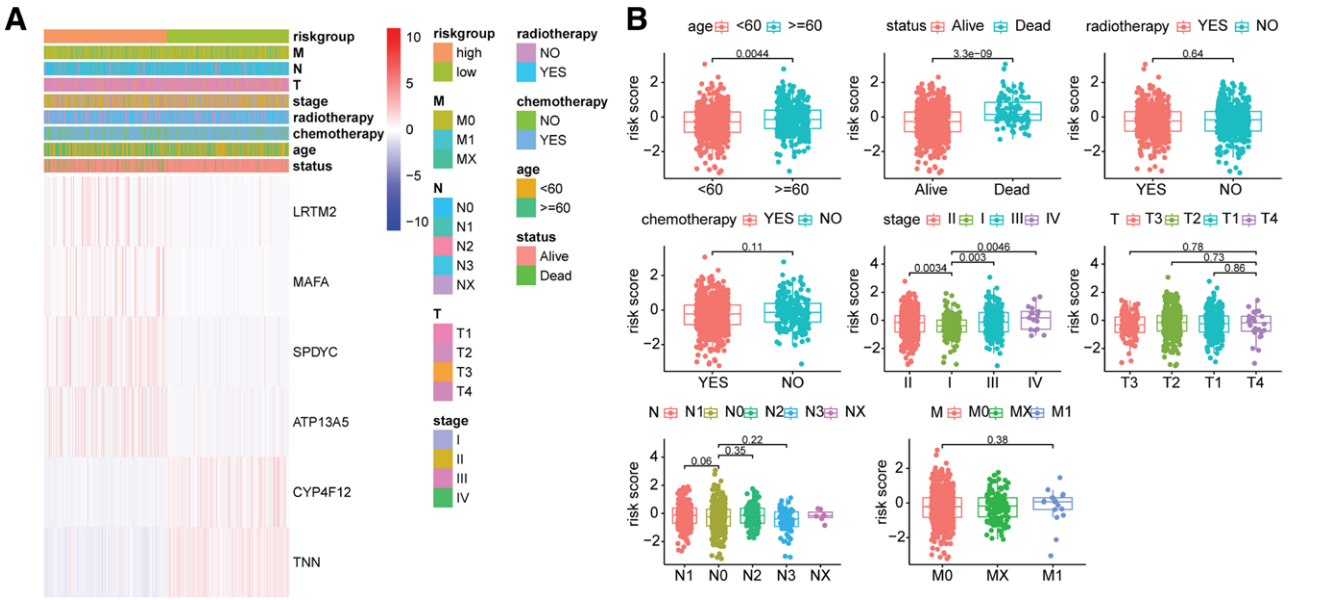

Figure 8. Association between the risk features and clinical-pathological features. (A) Heatmap showing the expression levels of selected genes across different clinical and pathological features. (B) Comparison of risk scores among different age groups, survival outcomes, radiotherapy/chemotherapy status, staging, T staging, N staging, and M staging using the Wilcoxon test.

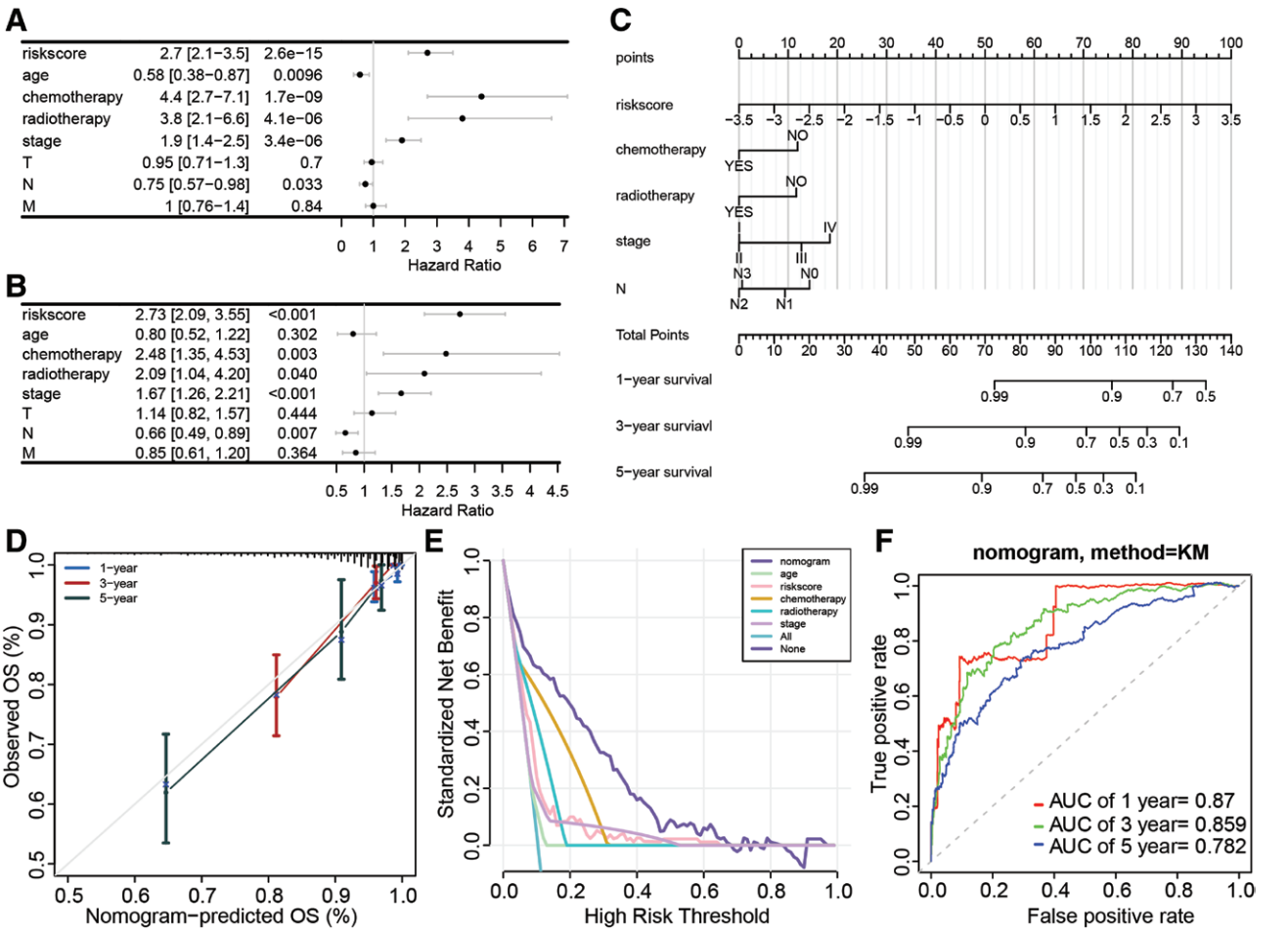

Figure 9. Development and evaluation of BC prognostic nomogram. (A) Univariate and (B) multivariate Cox regression analyses of risk scores and clinical-pathological features. (C) Nomogram composed of risk score, radiotherapy/chemotherapy status, staging, and N staging. (D) Calibration curve, (E) decision curve, and (F) receiver operating characteristic curve evaluating the predictive performance of the nomogram. BC = breast cancer.

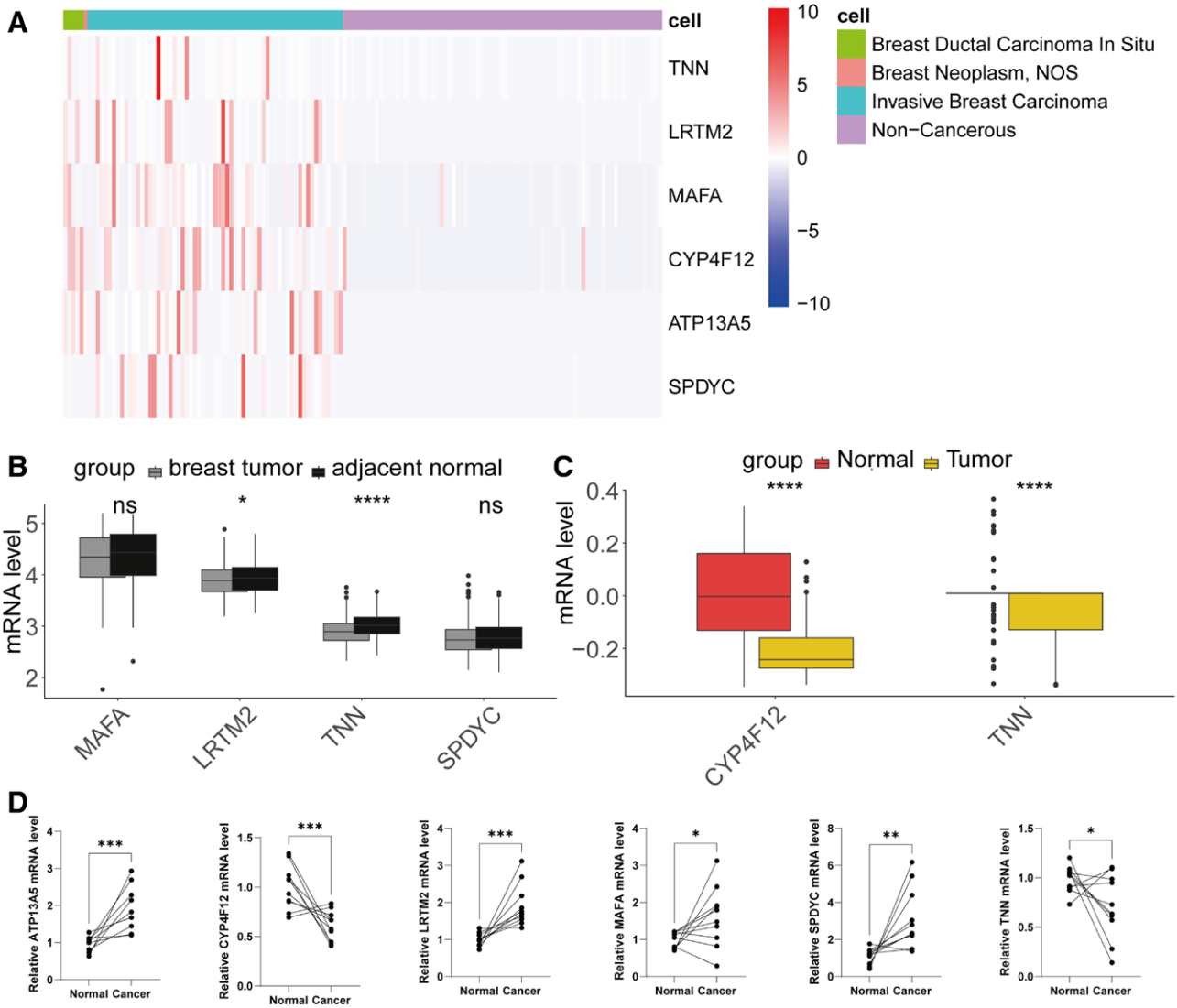

Figure 10. Validation of the expression of HSRGs in BC cells and tissue. (A) The heatmap of the expression of HSRGs in BC cells and noncancerous cells obtained from the CCLE database. Comparison of the expression of HSRGs between the BC tissue and adjacent tissue in the (B) GSE93601 cohort and (C) GSE14999 cohort using the Wilcoxon test. (D) Comparison of the relative gene expression level in clinical samples. ns, not significant, * $P < .05$, **** $P < .0001$. BC = breast cancer, CCLE = cancer cell line encyclopedia, HSRGs = H2S-related genes.

kinase 14 (MAPK14) and subsequent modulation of cAMP responsive element binding protein 1 (CREB1)/KLF transcription factor 4 (KLF4) activity can reverse M2 polarization, enhance antitumor immune effects, and exogenous H_2S has been shown to possess inhibitory properties against MAPK14 phosphorylation.[34,35] Additionally, inhibiting nitric oxide synthase 1 (NOS1) reduces the growth of fibroblasts expressing C-X-C motif chemokine ligand 14 (CXCL14) and their ability to promote tumor formation when co-injected with BC cells,[36] while exogenous H_2S exhibits inhibitory effects on NOS1 activity.[37] These findings shed light on the potential significance of HSRGs in reshaping the tumor immune microenvironment. Gene set enrichment analysis reveals aberrant immune-related biological processes among BC molecular subtypes derived from HSRGs.

There is a burgeoning volume of research underscoring the participation of H_2S in modulating tumor drug responsiveness. Studies demonstrate that H_2S can act synergistically with doxorubicin to enhance the chemotherapeutic efficacy of the latter compound.[38] Moreover, reported findings implicate elevated expression of drug-metabolizing enzymes and increased endogenous H_2S production in the development of 5-fluorouracil resistance within HCT116 cells.[39] GYY4137, a water-soluble H_2S donor, has been shown to cooperate with simvastatin or metformin to induce acidification-mediated cell death in cancer cells.[40] These investigations partially elucidate the differential chemosensitivity observed among molecular subtypes of BC derived from HSRGs signatures in the present study. Moreover, these prognostic-related HSRGs offer additional targets for countering drug resistance in BC therapeutics.

Despite the promising findings, several limitations should be acknowledged. First, the assessment of endogenous H_2S status represented by BC molecular subtypes based on prognostic-related HSRGs remains unexplored. Future studies should investigate the relationship between HSRGs expression and actual H_2S levels in tumor tissues to validate the biological significance of these subtypes, Second, the evaluation of exogenous H_2S response models is still lacking. Experimental studies using exogenous H_2S donors in preclinical models would provide valuable insights into how H_2S affects tumor biology and therapeutic responses. Third, the validation of prognostic characteristics associated with the prognostic model and molecular subtypes requires a larger sample size encompassing

more clinical cases. Prospective studies with longitudinal follow-up are needed to confirm the robustness and generalizability of our findings.

This study opens new avenues for future research on the role of HSRGs in BC. Further investigations should focus on elucidating the molecular mechanisms underlying the differential expression of HSRGs in various BC subtypes and exploring the potential of H_2S-based therapies. Preclinical studies using animal models and patient-derived xenografts could help validate the therapeutic potential of targeting HSRGs in BC. Additionally, clinical trials evaluating the efficacy of H_2S donors in combination with standard treatments, such as chemotherapy or immunotherapy, could provide important insights into their clinical utility. Finally, the development of biomarkers for monitoring H_2S levels in patients could facilitate personalized treatment strategies and improve patient outcomes.

5. Conclusion

In summary, the present study elucidated the existence of 2 distinct molecular subtypes in BC utilizing a panel of 46 prognostic-related HSRGs. The investigation unveiled substantial heterogeneity within these subtypes concerning gene expression patterns, prognosis, immune microenvironment characteristics, and therapeutic responsiveness. Additionally, a multi-gene risk feature and nomogram model were established to facilitate prognostic evaluation in BC patients. These significant findings not only contribute to expanding the scope of H_2S-based therapeutic strategies for BC but also pave the way for personalized treatment modalities tailored to individual patients.

Author contributions

Conceptualization: Congyan Yu, Yong Shen, Xuen Li.
Data curation: Congyan Yu, Yong Shen, Xuen Li.
Formal analysis: Congyan Yu, Yong Shen, Xuen Li.
Investigation: Congyan Yu, Yong Shen, Xuen Li.
Methodology: Congyan Yu, Yong Shen, Xuen Li.
Project administration: Congyan Yu, Yong Shen, Xuen Li.
Resources: Congyan Yu, Yong Shen, Xuen Li.
Software: Congyan Yu, Yong Shen, Xuen Li.
Supervision: Congyan Yu, Yong Shen, Xuen Li.
Validation: Congyan Yu, Yong Shen, Xuen Li.
Visualization: Congyan Yu, Yong Shen, Xuen Li.
Writing – original draft: Congyan Yu, Yong Shen, Xuen Li.
Writing – review & editing: Congyan Yu, Yong Shen, Xuen Li.

References

[1] Waks AG, Winer EP. Breast cancer treatment: a review. JAMA. 2019;321:288–300.
[2] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.
[3] Ferlay J, Colombet M, Soerjomataram I, et al. Cancer statistics for the year 2020: an overview. Int J Cancer. 2021;149:778–89.
[4] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71:209–49.
[5] Britt KL, Cuzick J, Phillips KA. Key steps for effective breast cancer prevention. Nat Rev Cancer. 2020;20:417–36.
[6] Metcalfe K, Lubinski J, Lynch HT, et al; Hereditary Breast Cancer Clinical Study Group. Family history of cancer and cancer risks in women with BRCA1 or BRCA2 mutations. J Natl Cancer Inst. 2010;102:1874–8.
[7] Khan NH, Duan S-F, Wu D-D, Ji X-Y. Better reporting and awareness campaigns needed for breast cancer in Pakistani women. Cancer Manag Res. 2021;13:2125–9.
[8] Crosby D, Bhatia S, Brindle KM, et al. Early detection of cancer. Science. 2022;375:eaay9040.
[9] Bhushan A, Gonsalves A, Menon JU. Current state of breast cancer diagnosis, treatment, and theranostics. Pharmaceutics. 2021;13:723.
[10] Sinha A, Naskar MNBJ, Pandey M, Rautaray SS. Challenges to the early diagnosis of breast cancer: current scenario and the challenges ahead. SN Comput Sci. 2024;5:170.
[11] Burguin A, Diorio C, Durocher F. Breast cancer treatments: updates and new challenges. J Pers Med. 2021;11:808.
[12] Khattak S, Rauf MA, Khan NH, et al. Hydrogen sulfide biology and its role in cancer. Molecules. 2022;27:3389.
[13] Distrutti E, Sediari L, Mencarelli A, et al. Evidence that hydrogen sulfide exerts antinociceptive effects in the gastrointestinal tract by activating KATP channels. J Pharmacol Exp Ther. 2006;316:325–35.
[14] Mani S, Untereiner A, Wu L, Wang R. Hydrogen sulfide and the pathogenesis of atherosclerosis. Antioxid Redox Signal. 2014;20:805–17.
[15] Carballal S, Trujillo M, Cuevasanta E, et al. Reactivity of hydrogen sulfide with peroxynitrite and other oxidants of biological interest. Free Radic Biol Med. 2011;50:196–205.
[16] Zhu XY, Liu S-J, Liu Y-J, Wang S, Ni X. Glucocorticoids suppress cystathionine gamma-lyase expression and H2S production in lipopolysaccharide-treated macrophages. Cell Mol Life Sci. 2010;67:1119–32.
[17] Shackelford RE, Mohammad IZ, Meram AT, et al. Molecular functions of hydrogen sulfide in cancer. Pathophysiology. 2021;28:437–56.
[18] Szabo C. Hydrogen sulfide, an endogenous stimulator of mitochondrial function in cancer cells. Cells. 2021;10:220.
[19] Chen S, Yue T, Huang Z, et al. Inhibition of hydrogen sulfide synthesis reverses acquired resistance to 5-FU through miR-215-5p-EREG/TYMS axis in colon cancer cells. Cancer Lett. 2019;466:49–60.
[20] Szabo C, Coletta C, Chao C, et al. Tumor-derived hydrogen sulfide, produced by cystathionine-β-synthase, stimulates bioenergetics, cell proliferation, and angiogenesis in colon cancer. Proc Natl Acad Sci U S A. 2013;110:12474–9.
[21] Wu D, Si W, Wang M, Lv S, Ji A, Li Y. Hydrogen sulfide in cancer: friend or foe? Nitric Oxide. 2015;50:38–45.
[22] Cai FF, Xu H-R, Yu S-H, et al. ADT-OH inhibits malignant melanoma metastasis in mice via suppressing CSE/CBS and FAK/Paxillin signaling pathway. Acta Pharmacol Sin. 2022;43:1829–42.
[23] Dong Q, Yang B, Han J-G, et al. A novel hydrogen sulfide-releasing donor, HA-ADT, suppresses the growth of human breast cancer cells through inhibiting the PI3K/AKT/mTOR and Ras/Raf/MEK/ERK signaling pathways. Cancer Lett. 2019;455:60–72.
[24] Khan NH, Wang D, Wang W, et al. Pharmacological inhibition of endogenous hydrogen sulfide attenuates breast cancer progression. Molecules. 2022;27:4049.
[25] Youness RA, Gad AZ, Sanber K, et al. Targeting hydrogen sulphide signaling in breast cancer. J Adv Res. 2021;27:177–90.
[26] Chen J, Shen X, Pardue S, et al. The Ataxia telangiectasia-mutated and Rad3-related protein kinase regulates cellular hydrogen sulfide concentrations. DNA Repair (Amst). 2019;73:55–63.
[27] Shackelford R, Ozluk E, Islam MZ, et al. Hydrogen sulfide and DNA repair. Redox Biol. 2021;38:101675.
[28] Hwang KT, Kim BH, Oh S, et al. Prognostic role of KRAS mRNA expression in breast cancer. J Breast Cancer. 2019;22:548–61.
[29] Fan K, Zhang S, Ni X, et al. KRAS G12D mutation eliminates reactive oxygen species through the Nrf2/CSE/H (2)S axis and contributes to pancreatic cancer growth. Acta Biochim Biophys Sin (Shanghai). 2022;54:1731–9.
[30] Chen S, Bu D, Zhu J, et al. Endogenous hydrogen sulfide regulates xCT stability through persulfidation of OTUB1 at cysteine 91 in colon cancer cells. Neoplasia. 2021;23:461–72.
[31] Parsanathan R, Jain SK. Hydrogen sulfide increases glutathione biosynthesis, and glucose uptake and utilisation in C(2)C(12) mouse myotubes. Free Radic Res. 2018;52:288–303.
[32] Cash H, Shah S, Moore E, et al. mTOR and MEK1/2 inhibition differentially modulate tumor growth and the immune microenvironment in syngeneic models of oral cavity cancer. Oncotarget. 2015;6:36400–17.
[33] Zhi L, Ang AD, Zhang H, Moore PK, Bhatia M. Hydrogen sulfide induces the synthesis of proinflammatory cytokines in human monocyte cell line U937 via the ERK-NF-kappaB pathway. J Leukoc Biol. 2007;81:1322–32.
[34] Hu HJ, Qiu J, Zhang C, Tang Z-H, Qu S-L, Jiang Z-S. Hydrogen sulfide improves ox-LDL-induced expression levels of Lp-PLA(2) in THP-1 monocytes via the p38MAPK pathway. Mol Med Rep. 2021;23:358.

[35] Rinaldi L, Gobbi G, Pambianco M, Micheloni C, Mirandola P, Vitale M. Hydrogen sulfide prevents apoptosis of human PMN via inhibition of p38 and caspase 3. Lab Invest. 2006;86:391–7.
[36] Augsten M, Sjöberg E, Frings O, et al. Cancer-associated fibroblasts expressing CXCL14 rely upon NOS1-derived nitric oxide signaling for their tumor-supporting properties. Cancer Res. 2014;74: 2999–3010.
[37] Kubo S, Kurokawa Y, Doe I, Masuko T, Sekiguchi F, Kawabata A. Hydrogen sulfide inhibits activity of three isoforms of recombinant nitric oxide synthase. Toxicology. 2007;241:92–7.
[38] Tian L, Pei R, Zhang X, et al. Tumor cell-specific and lipase-responsive delivery of hydrogen sulfide for sensitizing chemotherapy of pancreatic cancer. Front Bioeng Biotechnol. 2022;10:934151.
[39] Untereiner AA, Pavlidou A, Druzhyna N, Papapetropoulos A, Hellmich MR, Szabo C. Drug resistance induces the upregulation of H(2)S-producing enzymes in HCT116 colon cancer cells. Biochem Pharmacol. 2018;149:174–85.
[40] Lee ZW, Teo X-Y, Song ZJ, et al. Intracellular hyper-acidification potentiated by hydrogen sulfide mediates invasive and therapy resistant cancer cell death. Front Pharmacol. 2017;8:763.