


Insights into a Machine Learning-Based Palmitoylation-Related Gene Model for Predicting the Prognosis and Treatment Response of Breast Cancer Patients

Technology in Cancer Research & Treatment
Volume 23: 1-16
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/15330338241263434
journals.sagepub.com/home/tct



Hongxia Zhu, MS^{1,2,#}, Haihong Hu, MS^{1,2,#}, Bo Hao, BD²,
Wendi Zhan, MS^{1,2}, Ting Yan, MS³, Jingdi Zhang, MS^{1,2},
Siyu Wang, MS⁴, Hongjuan Hu, MS⁵, and Taolan Zhang, MD^{2,6} 

Abstract

Background: Breast cancer is a prevalent public health concern affecting numerous women globally and is associated with palmitoylation, a post-translational protein modification. Despite increasing focus on palmitoylation, its specific implications for breast cancer prognosis remain unclear. The work aimed to identify prognostic factors linked to palmitoylation in breast cancer and assess its effectiveness in predicting responses to chemotherapy and immunotherapy. **Methods:** We utilized the “limma” package to analyze the differential expression of palmitoylation-related genes between breast cancer and normal tissues. Hub genes were identified using the “WGCNA” package. Using the least absolute shrinkage and selection operator (LASSO) Cox regression analysis, we identified a prognostic feature associated with palmitoylation and developed a prognostic nomogram with the “regplot” package. The predictive values of the model for chemotherapy and immunotherapy responses were assessed using immunophenoscore (IPS) and the “pRophetic” package. **Results:** We identified 211 differentially expressed genes related to palmitoylation, among which 44 demonstrated prognostic potential. Subsequently, a predictive model comprising eleven palmitoylation-related genes was developed. Patients were classified into high-risk and low-risk groups based on the median risk score. The findings revealed that individuals in the high-risk group exhibited lower survival rates, while those in the low-risk group showed increased immune cell infiltration and improved responses to chemotherapy and immunotherapy. Moreover, the BC-Palmitoylation Tool website was established. **Conclusion:** This study developed the first machine learning-based predictive model for palmitoylation-related genes and created a corresponding website, providing clinicians with a valuable tool to improve patient outcomes.

Keywords

breast cancer, palmitoylation, prognostic model, nomogram, website application

Received: March 7, 2024; Revised: May 10, 2024; Accepted: May 30, 2024.

¹ School of Pharmacy, Hengyang Medical College, University of South China, Hengyang, China

² Department of Pharmacy, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, China

³ Department of Breast and Thyroid Surgery, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, China

⁴ Department of Medical Oncology, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, China

⁵ Department of Public Health Service, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, China

⁶ Phase I Clinical Trial Center, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, China

These authors contributed equally to this work.

Corresponding Author:

Taolan Zhang, Department of Pharmacy, The First Affiliated Hospital, Hengyang Medical School, University of South China, 69 Chuanshan Road, Hengyang, Hunan 421000, China.

Email: TaolZhan@usc.edu.cn



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Introduction

Breast cancer, surpassing lung cancer, has become the most prevalent cancer globally, with an estimated 2.3 million new cases each year.¹ It remains the leading cause of cancer-related morbidity and mortality among women worldwide. The etiology of breast cancer is complex, influenced by genetic mutations, hormonal factors, lifestyle choices, and environmental exposure.² Advances in molecular typing, diagnostics, and therapeutic strategies have significantly improved breast cancer management. Treatment primarily involves local therapies, including surgery and radiation, and systemic therapies, including chemotherapy, endocrine therapy, targeted therapy, and immunotherapy.³ Despite significant progress in immunotherapy for breast cancer, its application is currently restricted to patients with triple-negative breast cancer, and its effectiveness as a solitary treatment option remains limited. There is still a need for new prognostic assessment methods that can further predict the clinical prognosis and efficacy of immunotherapy for breast cancer patients.

Palmitoylation, a post-translational modification, plays a crucial role in numerous cellular processes, including protein transport, membrane localization, and signal transduction.⁴ It is increasingly recognized as a contributor to the incidence and progression of breast cancer,⁵ regulating essential proteins such as estrogen receptor alpha (ER α)⁶ and caveolin-1,⁷ which are essential for the dynamics of breast cancer.⁸ Moreover, palmitoylation is associated with drug resistance, metastasis, and recurrence in breast cancer.^{9,10} Specifically, the palmitoylation of mTOR by ZDHHC22 inhibits the malignant proliferation of breast cancer cells and increases their sensitivity to tamoxifen endocrine therapy.⁹ The inhibition of PD-L1 palmitoylation enhances the T cell immune response against breast cancer, thereby reducing tumor growth.¹¹ Therefore, targeting palmitoylation presents a promising approach for overcoming tumor immune evasion and addressing challenges such as drug resistance, recurrence, and metastasis in breast cancer. Despite its importance, reliable predictive biomarkers for palmitoylation in breast cancer remain scarce. The development of a model related to palmitoylation could hold clinical significance for predicting the prognosis of breast cancer patients undergoing chemotherapy and immunotherapy.

This novel investigation explored differential expression patterns of genes associated with palmitoylation in breast cancer and normal tissues. Weighted gene co-expression network analysis (WGCNA) was utilized to identify hub genes, followed by enrichment analysis to elucidate their biological functions. A prognostic model was developed using the least absolute shrinkage and selection operator (LASSO) Cox regression analysis based on 11 palmitoylation-related genes for the first time. This model shows strong predictive capacity for survival time, chemotherapy response, and immunotherapy efficacy in breast cancer patients. To enhance prognostic accuracy, the risk score was integrated with relevant clinical characteristics, constructing a prognostic nomogram capable of predicting the survival rate at 1, 3, and 5 years for

breast cancer patients. Notably, the BC-Palmitoylation Tool, a web application, has been developed to provide clinicians with prognostic information for breast cancer patients. Although still in its initial stages of deployment, this tool has the potential to enhance personalized treatment strategies for breast cancer significantly. The workflow of this study is illustrated in Figure 1.

Materials and Methods

Data Source and Processing

A total of 2239 genes associated with palmitoylation were retrieved from GeneCards (<https://www.genecards.org/Search/Keyword?queryString=Palmitoylation>). Molecular data relevant to breast cancer patients, such as somatic mutations and gene expression profiles, along with clinical characteristics including gender, age, stage, and survival times, were obtained from the GDC TCGA BRCA (<https://xenabrowser.net/datapages/>). Ensembl IDs in the gene expression matrix were converted to gene symbols using probe annotation information. The expression matrix was then subjected to quantile normalization through the `normalizeBetweenArrays` function of the “limma” package. Information for 1075 breast cancer patients was obtained after normal tissue and duplicate samples were excluded. Ultimately, breast cancer patients in the TCGA cohort were divided into a training cohort ($n = 755$) and a test cohort ($n = 320$) at a 7:3 ratio using stratified random sampling. The dataset GSE20685 was sourced from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE20685>, including 327 breast cancer tissue samples). The details of all cohorts are shown in Table 1.

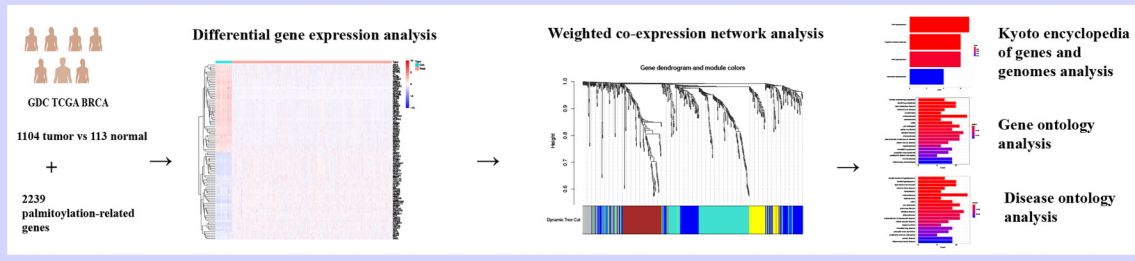
Identification of Palmitoylation-Related Hub Genes

Differential Gene Expression Analysis. Our analysis used the “limma” package in R to conduct differential gene expression analysis between breast cancer and normal samples from the TCGA dataset, applying stringent criteria of $|\log_{2}FC| > 1$ and $P < 0.05$.¹² The comprehensive approach enabled by “limma” included several essential functions: `lmFit` was utilized to model gene expression, `makeContrasts` defined group comparisons, `contrasts.fit` applied these contrasts, `eBayes` moderated statistical errors, and `topTable` ranked the differentially expressed genes (DEGs).

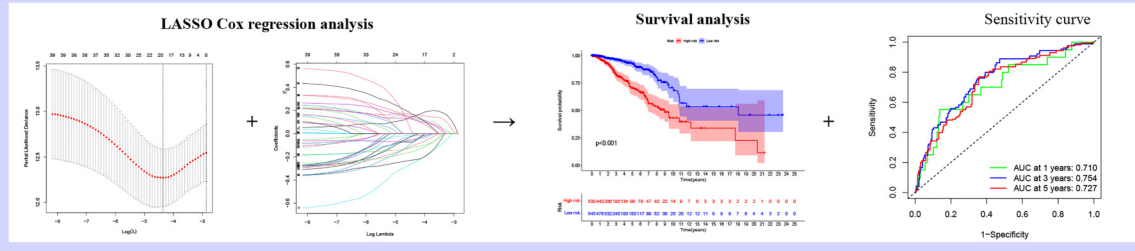
Applying WGCNA to Uncover the Hub Palmitoylation Module. The WGCNA algorithm was utilized to establish a co-expression network that identifies palmitoylation-related genes in breast cancer samples.¹³ Hierarchical clustering analysis was conducted using the `Hclust` function from the “WGCNA” package. The `pickSoftThreshold` function was employed to select an appropriate soft threshold, followed by using the `adjacency` function to create the neighbor-joining matrix. The `TOMsimilarity` function computed the Topological Overlap Matrix (TOM) from this matrix, transforming the adjacency

Workflow

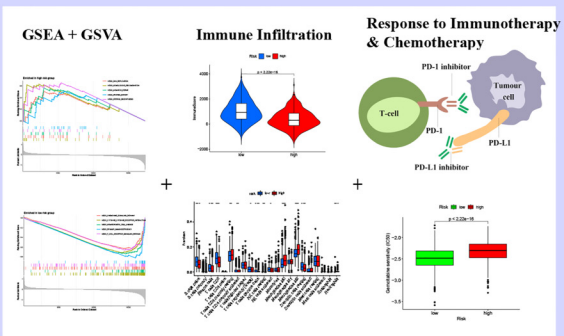
A Identification of Palmitoylation-related Hub Genes



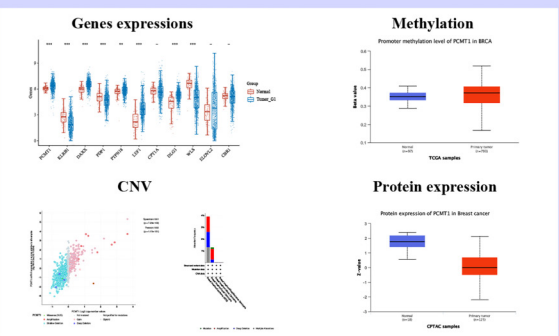
B Construct Prognostic Signature



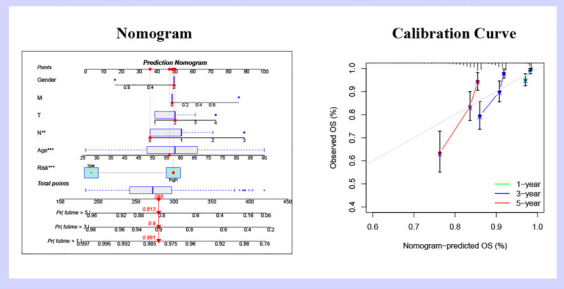
C Further Analysis of Risk Groups



D Analysis of the 11 Genes Regulation



E Developing a Prognostic Risk Nomogram



F Web Application

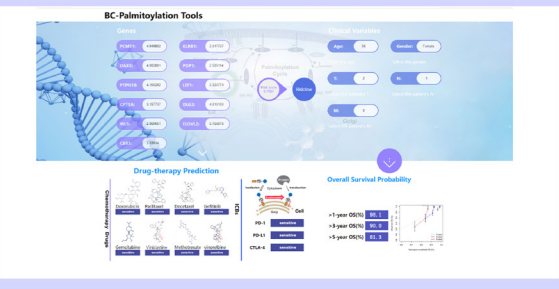


Figure 1. Workflow of the study.

Table 1. Clinical Information of the Breast Cancer Patients in the Three Cohorts.

Variables	TCGA train cohort (N = 755)	TCGA test cohort (N = 320)	GSE cohort (N = 327)
Incomplete	N = 116	N = 62	0
Age			
≥65 years	219	83	22
<65 years	536	237	305
Gender			
Female	747	316	327
Male	8	4	0
M classification			
M0	636	258	244
M1	15	7	83
N classification			
N0	348	158	122
N1	252	107	102
N2	88	32	63
N3	55	18	40
T classification			
T1	189	91	101
T2	442	179	188
T3	92	41	30
T4	31	7	8
Stage classification			
Stage I	121	61	NA
Stage II	433	177	NA
Stage III	174	67	NA
Stage IV	14	6	NA

matrix into a topological overlap matrix. Gene modules were detected and visualized within the network using the *cutreeDynamic* function. Ultimately, the correlation between module feature genes and breast cancer was calculated, identifying the most pertinent modules for subsequent analysis.

Enrichment Analyses of Biological Functions. Intersections were made between selected module genes and DEGs. Enrichment analyses were performed on these genes across Kyoto encyclopedia of genes and genomes (KEGG),¹⁴ gene ontology (GO),¹⁵ and disease ontology (DO)¹⁶ domains. The function and pathway terms were retrieved for further visualization in R software.

Construction of a Palmitoylation-Related Prognostic Signature for Breast Cancer

A palmitoylation-related prognostic signature for breast cancer was developed and validated through LASSO Cox regression.¹⁷ Univariate Cox regression analysis using the “*survival*” package was conducted to determine the prognostic value of these palmitoylation-related genes for overall survival (OS). Prognosis-related genes were identified using a significance threshold of $P < 0.01$. Cross-validation was performed using the *cv.glmnet* function to determine the optimal penalty parameter (λ) for the LASSO model, ensuring robustness and reducing overfitting. A Cox proportional hazards model was

subsequently developed using the selected genes to evaluate their impact on patient survival times. Forward and backward stepwise selection methods were employed to refine the model, retaining only the most significant predictors. Risk scores for predicting the risk of adverse outcomes were calculated for each patient using the final Cox model. The risk scores were calculated using the following formula:

$$\text{Risk score} = \sum_{i=1}^n (\text{Gene Expression}_i \times \text{coef}_i)$$

Patients were stratified into high- and low-risk groups based on the median risk score from the training set. This risk stratification method was further applied to the TCGA and GEO test cohorts to assess the performance of the model. Survival analysis was conducted using the “*survival*” and “*survminer*” packages to generate Kaplan–Meier curves, assessing the prognostic significance of the model. Receiver operating characteristic (ROC) curves were generated using the “*timeROC*” package to assess the accuracy of the model in predicting survival time. Area under the curve (AUC) values for 1 year, 3 years, and 5 years were calculated to quantify the performance of the model. Additionally, the “*pheatmap*” package was utilized to analyze and visualize the risk and survival status of patients across various cohorts.

Analysis of the Tumor Microenvironment

Gene set enrichment analysis (GSEA)¹⁸ and gene set variation analysis (GSVA)¹⁹ were conducted utilizing gene sets from the “*c2.cp.kegg.v7.4.symbols.gmt*” available through the MSigDB database (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>). GSEA was performed to compare biological functions between high-risk and low-risk groups using the R packages “*limma*,” “*GSEABase*,” “*GSVA*,” and “*pheatmap*.” Concurrently, GSVA was employed to quantify variations in gene set activity among different risk groups, utilizing the packages “*limma*,” “*org.Hs.eg.db*,” “*clusterProfiler*,” and “*enrichplot*.”

We utilized the “*estimate*” package to assess the stromal and immune cell infiltrates in the tumor microenvironment (TME) by calculating ESTIMATE scores from the normalized expression data.²⁰ We performed single-sample gene set enrichment analysis (ssGSEA) using the “*GSVA*” package to derive immune activity scores from the expression data against predefined immune gene sets. These scores were then merged with TME scores and cancer risk data, and visualized using a heatmap to depict variations in immune activity across different risk groups. Immune cell infiltration analysis was conducted using the CIBERSORT tool,²¹ which quantified the relative abundances of various immune cell types within the TME across high- and low-risk groups. We performed correlation analyses to understand further the relationship between immune cell infiltration and cancer risk scores.

Prediction of Response to Immunosuppressants and Chemotherapy

From the existing literature, 40 genes linked to immune checkpoints were identified and subsequently analyzed for differential expression across distinct risk groups using the R packages “limma” and “reshape2.” The efficacy of immune checkpoint inhibitors was assessed by the immunophenotypescore (IPS), which serves as a dependable predictor of patient responses to CTLA-4 and PD-1 inhibitors,²² obtained from the TCIA database (<https://tcia.at/>). Chemosensitivity was predicted by calculating the 50% inhibitory concentrations (IC50) of each sample employing the “pRophetic” package.²³ The Wilcoxon rank-sum test assessed the statistical significance of the differences in IC50 values between the two groups.

Analysis of Gene Regulation

Expression levels of eleven genes were analyzed for differences between breast cancer and normal tissues within the prognostic model, utilizing R version 4.0.3. The differences in promoter methylation and protein expression levels of these genes were further explored via the UALCAN database (<http://ualcan.path.uab.edu/index.html>). Moreover, the cBioPortal database was employed to evaluate the copy number variations (CNVs) and their correlation with mRNA expression in these genes.

Developing a Prognostic Risk Nomogram

Clinical features, including age, gender, TNM stage, and stage, were integrated with risk score to construct a prognostic nomogram using univariate and multivariate Cox regression analyses.²⁴ The “regplot” package was used to visualize this nomogram. To evaluate the predictive accuracy of the nomogram relative to other prognostic factors, ROC curve analysis and decision curve analysis (DCA) were performed using the “timeROC” and “ggDCA” packages, respectively.

Statistical Analysis

Statistical analyses were conducted using R software (versions 4.0.3, 4.1.3 and 4.2.2). The Wilcoxon test was utilized to assess differences between the two groups. Correlation analysis was performed using the Pearson test. Survival curves were created using the Kaplan–Meier method and compared via the log-rank test. A significance threshold was set at $P < 0.05$.

Results

Hub Genes Associated with Palmitoylation

Differential gene expression analysis identified 211 palmitoylation-related DEGs (Figure 2A). WGCNA was conducted on palmitoylation-associated genes, which categorized 663 genes into five modules (Figure 2B and C). Intersection analysis between the hub genes from the turquoise module, as defined by Pearson correlation coefficients, and the DEGs

resulted in a set of 58 significant genes (Figure 2D). KEGG, GO, and DO enrichment analyses were performed to investigate the biological functions of these intersecting genes. KEGG enrichment analysis indicated a pronounced enrichment of these genes in biological pathways related to metabolism and cellular regulation, particularly within the PPAR signaling, regulation of lipolysis in adipocytes, AMPK signaling, and adipocytokine signaling pathways (Figure 2E). GO enrichment analysis further revealed these genes’ potential roles in biological processes such as lipid metabolism and storage, as well as their localization to cellular structures like lipid droplets and the sarcolemma, and activity in molecular functions involving the catalysis of lipid catabolism (Figure 2F). DO enrichment analysis associated these genes with metabolic disorder diseases, such as familial hyperlipidemia, obesity, and type 2 diabetes (Figure 2G).

Construction of a Palmitoylation-Related Prognostic Signature for Breast Cancer

Univariate Cox regression analysis was conducted with a P -value threshold of < 0.01 , identifying 44 palmitoylation-related genes associated with clinical outcomes. Subsequently, genes possessing the highest prognostic value were selected through a LASSO Cox regression model, resulting in the choice of 11 candidate genes (PCMT1, KLRB1, DAXX, PDP1, PTPN18, LEF1, CPT1A, DLG3, WLS, ELOVL2, and CBR1) for constructing the model (Figure 3A–D). Risk scores were calculated using coefficients in the formula below:

$$\begin{aligned} \text{Risk score} = & \text{PCMT1} \times 0.3576917 + \text{KLRB1} \times (-0.43603077) \\ & + \text{DAXX} \times (-0.362550914) + \text{PDP1} \times 0.254808915 + \text{PTPN18} \\ & \times (-0.319773106) + \text{LEF1} \times (-0.177420729) + \text{CPT1A} \times \\ & 0.214187141 + \text{DLG3} \times 0.380288463 + \text{WLS} \times (-0.189086763) \\ & + \text{ELOVL2} \times (-0.118885307) + \text{CBR1} \times (-0.220853695). \end{aligned}$$

In the TCGA train cohort, the Kaplan–Meier survival curve demonstrated that patients in the low-risk group exhibited significantly higher survival rates and longer survival times than those in the high-risk group. Additionally, ROC curve analysis of this cohort revealed that the predicted AUC values for 1, 3, and 5 years were 0.697, 0.761, and 0.764, respectively, indicating that the model accurately predicts patient survival outcomes (Figure 3E). Consistency in survival curves was observed across the TCGA test cohort, the TCGA all cohort, and the GEO test cohort, with most datasets exhibiting AUC values exceeding 0.7 in ROC curve analyses (Figure 3F–H). The accuracy of this model in predicting patient prognoses was further validated through internal and external cohorts.

Characteristics of the Tumor Microenvironment in the High- and Low-Risk Groups

GSEA and GSVA were employed to identify enriched gene sets across different risk groups. Enrichment in gene sets related to DNA replication, homologous recombination, mismatch repair, protein export, and steroid biosynthesis was observed in the high-risk group. In contrast, the low-risk group was characterized by

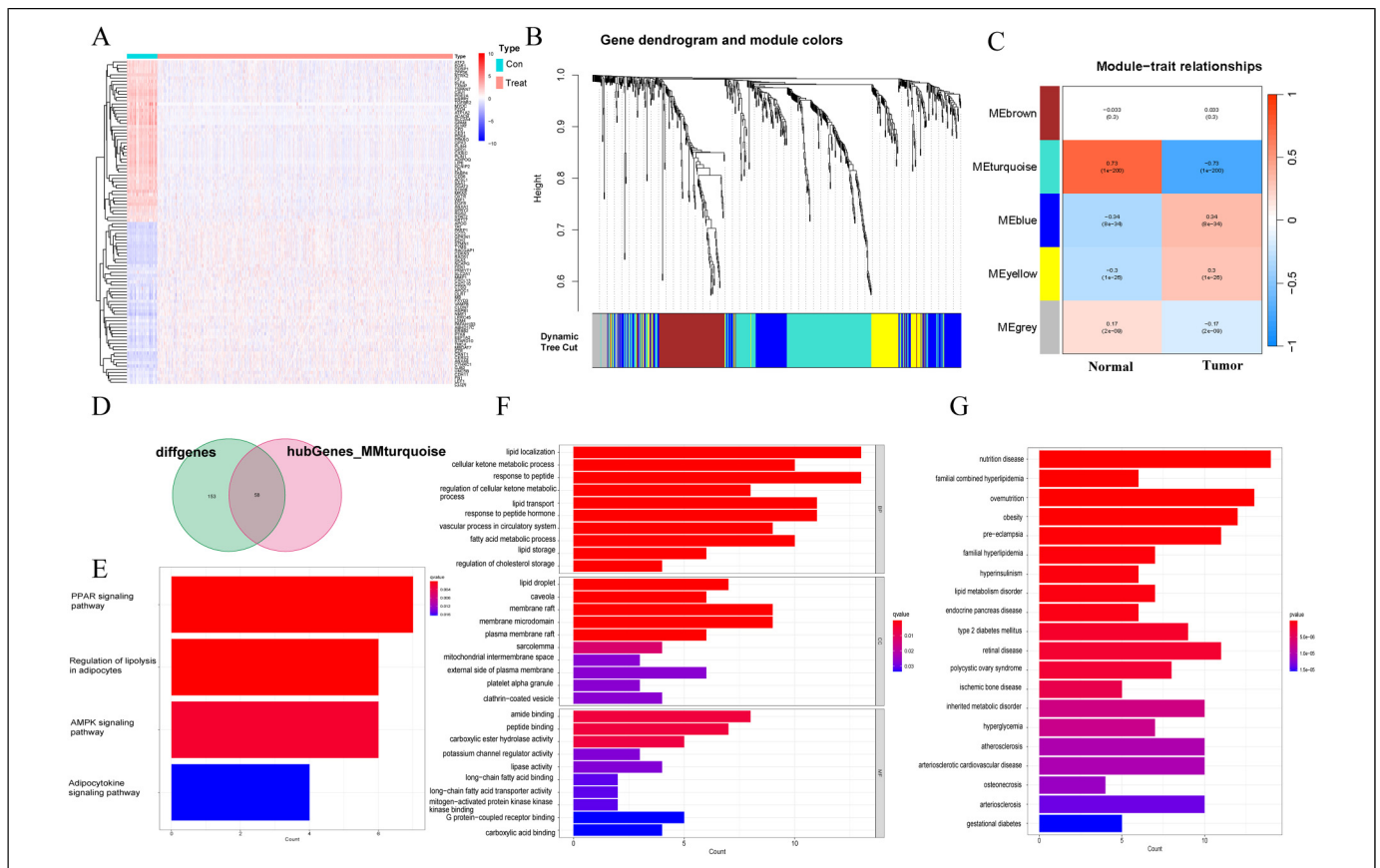


Figure 2. Identification of palmitoylation-related hub genes. (A) Heatmap of the palmitoylation-related differentially expressed genes (DEGs) between breast cancer and normal tissues. (B) Clustering dendrogram of genes; various colors represent different modules. (C) Relationships between the five modules and sample tissue type. (D) Venn diagram of DEGs and hub genes intersections. (E) Kyoto encyclopedia of genes and genomes (KEGG) enrichment analyses. (F) Gene ontology (GO) enrichment analyses. (G) Disease ontology (DO) enrichment analyses.

enriched pathways associated with primary immunodeficiency, T-cell receptor signaling, hematopoietic cell lineage, cytokine-cytokine receptor interaction, and chemokine signaling (Figure 4A). GSEA indicated significant enrichment of pathways related to protein and lipid anabolism in the high-risk group, whereas the low-risk group demonstrated notable enrichment in immune response-related signaling pathways (Figure 4B).

The ESTIMATE analysis revealed that the stromal score, immune score, and estimate score were higher in the low-risk group compared to the high-risk group, while the tumor purity showed the opposite trend (Figure 4C). The immune infiltration levels in the high-risk and low-risk groups were assessed, and a heatmap was generated to depict the abundance levels of 29 immune-related cells and molecules (Figure 4D). Further analysis of immune cell infiltration indicated that the low-risk group exhibited more extensive infiltration, including naive B cells, plasma cells, CD8+ T cells, activated memory CD4+ T cells, regulatory T cells (Tregs), $\gamma\delta$ T cells, resting natural killer (NK) cells, activated NK cells, and resting dendritic cells; conversely, a higher infiltration of resting memory CD4+ T cells, as well as M0 and M2 macrophages, was observed in the high-risk group (Figure 4E). Correlation analysis demonstrated that risk scores were negatively associated with naive B cells, resting dendritic cells, activated NK

cells, plasma cells, CD8+ T cells, $\gamma\delta$ T cells, activated memory CD4+ T cells, and regulatory T cells (Figure 4F).

Assessing the Model's Ability to Predict Immunotherapy and Chemotherapy Effectiveness

The expression levels of 36 immune checkpoint-related genes were compared between different risk groups, revealing that all genes were significantly upregulated in the low-risk group (Figure 5A). To evaluate the potential clinical efficacy of immune checkpoint inhibitors (ICIs), we utilized IPS scores to analyze immune infiltration in the tumors of both groups. The results demonstrated that IPS scores based on CTLA-4 or PD-1 status were higher in the low-risk group, suggesting that patients in this group may exhibit better therapeutic responses to ICIs (Figure 5B). Furthermore, differences in sensitivity to common chemotherapeutic agents and targeted drugs between the two groups were investigated. Compared to patients in the high-risk group, those in the low-risk group exhibited relatively lower IC50 values for drugs such as gemcitabine, doxorubicin, docetaxel, vinblastine, vinorelbine, paclitaxel, methotrexate, and gefitinib (Figure 5C). These findings indicate greater sensitivity to chemotherapeutic drugs in the low-risk group.

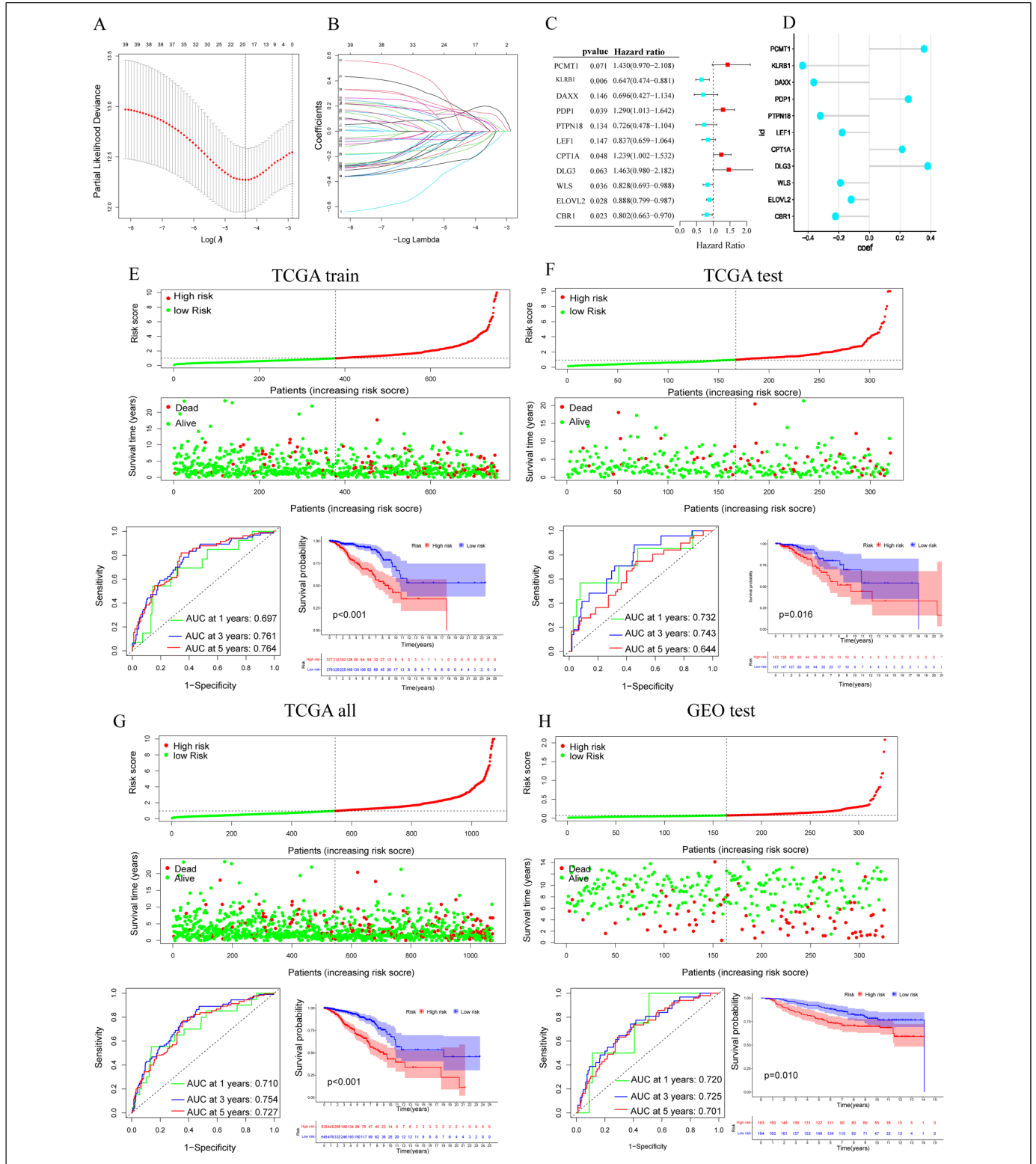


Figure 3. Construction of a palmitoylation-related prognostic signature for breast cancer. (A) Partial likelihood deviance of the least absolute shrinkage and selection operator (LASSO) regression. (B) LASSO coefficient profiles based on logarithmic (lambda) sequences. (C) Univariate Cox analysis of 11 survival-related genes. (D) Lollipop chart showing the LASSO coefficient of the 11 genes in the model. The distribution of the LASSO coefficients of the genes in the signature. (E–H) In the TCGA train, TCGA test, TCGA all and GEO test datasets, distribution of the patient’s risk score, patients’ overall survival time along with their risk score, Kaplan–Meier OS curves, and time-dependent ROC curves for breast cancer patients at 1, 3, and 5 years.

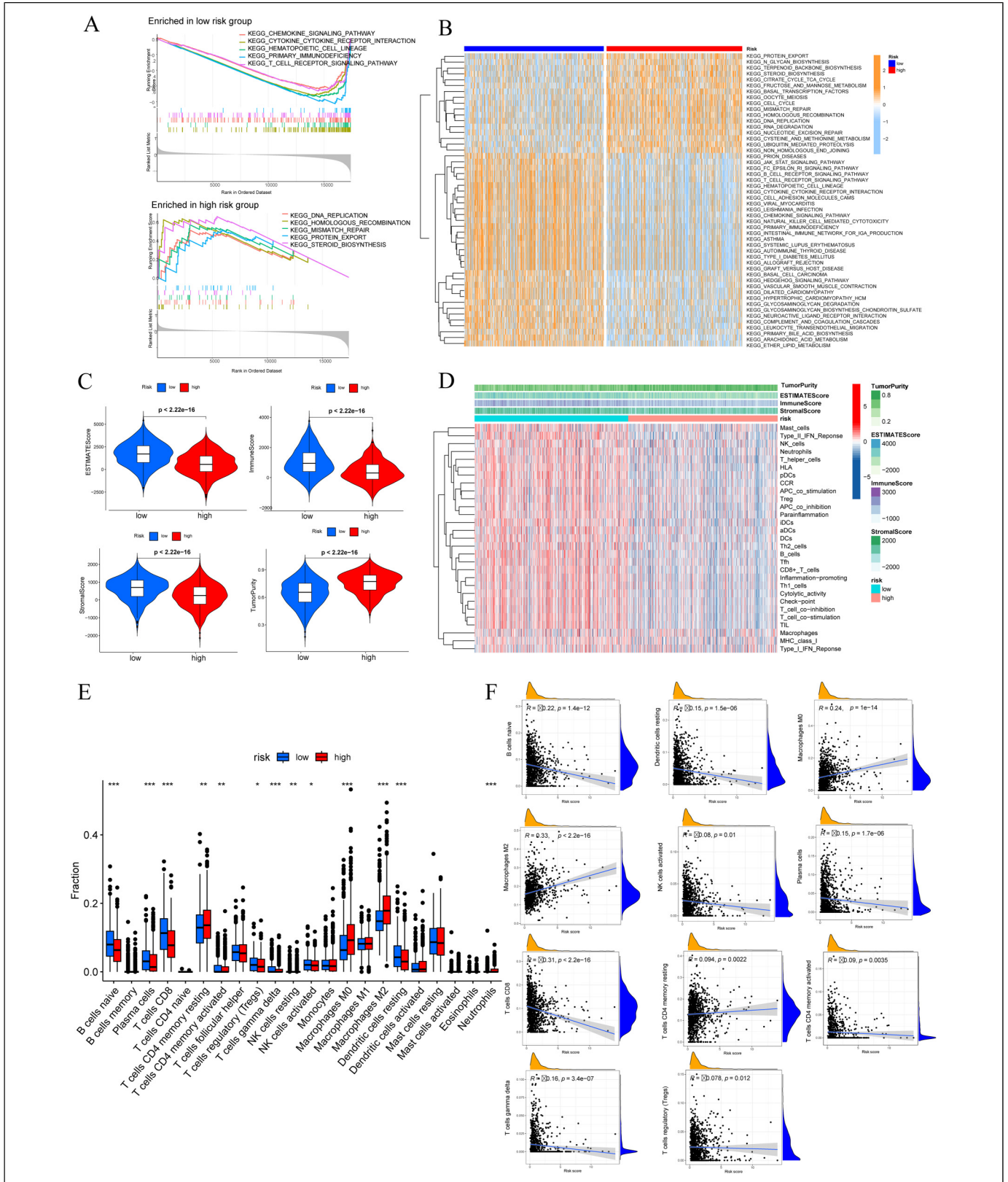


Figure 4. The landscape of immune infiltration. (A) Plots of gene set enrichment analysis for the high- and low-risk groups. (B) Heatmap of gene set variation analysis results for the high- and low-risk groups. (C) Violin plots of the estimate score, immune score, stromal score, and tumor purity in the two groups. (D) Heatmap to evaluate the levels of immune infiltration in the high- and low-risk groups. (E) The proportions of immune cells in the two groups (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$). (F) Relevance analysis between the risk score and the level of 11 infiltrating immune cells.

between breast cancer tissues and normal tissues (Figure 6B). In breast cancer, only the promoter methylation level of DLG3 was significantly lower, whereas those of WLS, ELOVL2, PTPN18, LEF1, and CBR1 were notably higher, with no significant differences observed in the remaining five genes. Subsequently, the protein expressions of these 11 genes were analyzed, revealing that in breast cancer, the protein expression of PCMT1, PDP1, PTPN18, WLS, and CBR1 was significantly reduced, whereas that of DLG3 was markedly increased (Figure 6C).

The mutation frequencies of 11 genes in breast cancer were analyzed, uncovering that PDP1 and CPT1A exhibited the highest mutation frequencies, followed by ELOVL2, KLRB1, WLS, PCMT1, DAXX, CBR1, LEF1, PTPN18, and DLG3 (Supplementary Figure S1A and S1B). Apart from PTPN18 and DLG3, which primarily presented mutations, most mutations were identified as amplifications. A correlation analysis was conducted between CNVs and gene expression to explore their association. As shown in Supplementary Figure S1C, the expression levels of PCMT1, DAXX, PDP1, PTPN18, CPT1A, WLS, and CBR1 were significantly influenced by copy number changes, whereas the expression of KLRB1, LEF1, DLG3, and ELOVL2 showed minimal correlation with copy number changes (no statistical difference in *P*-values).

Building a Predictive Probability Nomogram

To investigate the independence of the model's effect, a Cox analysis was performed on the TCGA dataset, integrating the risk score with clinical characteristics such as age, gender, and stage. Univariate Cox analysis revealed significant associations between M, N, T, age, stage, and the risk score with OS in breast cancer patients (Figure 7A). However, in the multivariate Cox analysis, only age and the risk score emerged as independent prognostic factors (Figure 7B, all $P < 0.001$). Subsequently, a nomogram was developed to predict the 1-year, 3-year, and 5-year survival rates of breast cancer patients (Figure 7C). Calibration plots demonstrated the excellent predictive accuracy of the nomogram, as it closely approximated the ideal curve and exhibited high stability (Figure 7D). Furthermore, according to ROC analysis, the nomogram showed greater accuracy in predicting OS compared to other features (Figures 7E–G). DCA also indicated that the net benefit of the nomogram surpassed that of any single clinical feature alone (Figures 7H–J). Therefore, our nomogram demonstrates significant predictive value and may have practical implications for clinical decision-making.

Website Application

The final prognostic nomogram model was implemented within a web-based application. The primary purpose of this application is to provide a visual representation and prognostication of the OS rates of breast cancer patients over 1-, 3-, and 5-year periods. To achieve this, the application incorporates

various input parameters, including the patient's gender, age, stage, TNM stage, and gene expression levels associated with prognostic characteristics. Furthermore, the tool can predict the responses of individuals to chemotherapy and immunotherapy. The application offers valuable insights into the expected treatment responses by leveraging these parameters. For convenient access, the application is available online at the following URL: <http://demo.dlhwg.com/cox-h/cox-h.html> (Figure 8). Users can readily access and utilize its features through this platform.

Discussion

According to data from the International Agency for Research on Cancer (IARC) of the World Health Organization (WHO), breast cancer became the most common cancer worldwide for the first time in 2020. Breast cancer also remains the leading cause of cancer-related deaths among women.¹ Despite advances in medical technology, the treatment of breast cancer presents a significant challenge in the field of oncology due to its phenotypic and genotypic heterogeneity, which often leads to drug resistance, metastasis, and recurrence. Palmitoylation is a post-translational modification that plays a crucial role in various cellular processes, including membrane localization, protein transport, and signal transduction. There is growing evidence that palmitoylation significantly impacts breast cancer development and progression, as well as drug resistance, metastasis, and recurrence. For instance, Huang et al reported that ectopic expression of ZDHHC22 inhibits the malignant proliferation of breast cancer cells and restores their sensitivity to tamoxifen treatment by palmitoylating mTOR.⁹ PD-L1, known for its role in the immune evasion strategy of cancer cells, plays a key role in suppressing the immune response. Yao et al reported that targeted inhibition of PD-L1 palmitoylation enhances the immune response of T cells to breast cancer, thereby inhibiting tumor growth.¹¹ Palmitoylation modifications exert regulatory control over T-cell activation and tumor immunity through their impact on the stability, cellular localization, and function of PD-L1. Furthermore, the retention of CD44 palmitoylation within lipid rafts, which decreases its binding to the migration binding partner ezrin, was found to exert an inhibitory effect on the migration of breast cancer cells.¹⁰ Despite recognizing the crucial role of palmitoylation in breast cancer, the molecular markers associated with palmitoylation are still not fully understood. In this study, we aimed to address this gap by developing a predictive model for the survival of breast cancer patients in the TCGA cohort. The predictive accuracy of the model was internally validated using TCGA data and externally validated using GEO data (GSE20685). We generated an alignment diagram that included age, gender, TNM stage, stage, and risk score. The results consistently demonstrate the model's high performance and potential clinical utility.

GSEA was performed between the high- and low-risk groups, revealing significant enrichment of DNA replication, homologous recombination, mismatch repair, protein export,

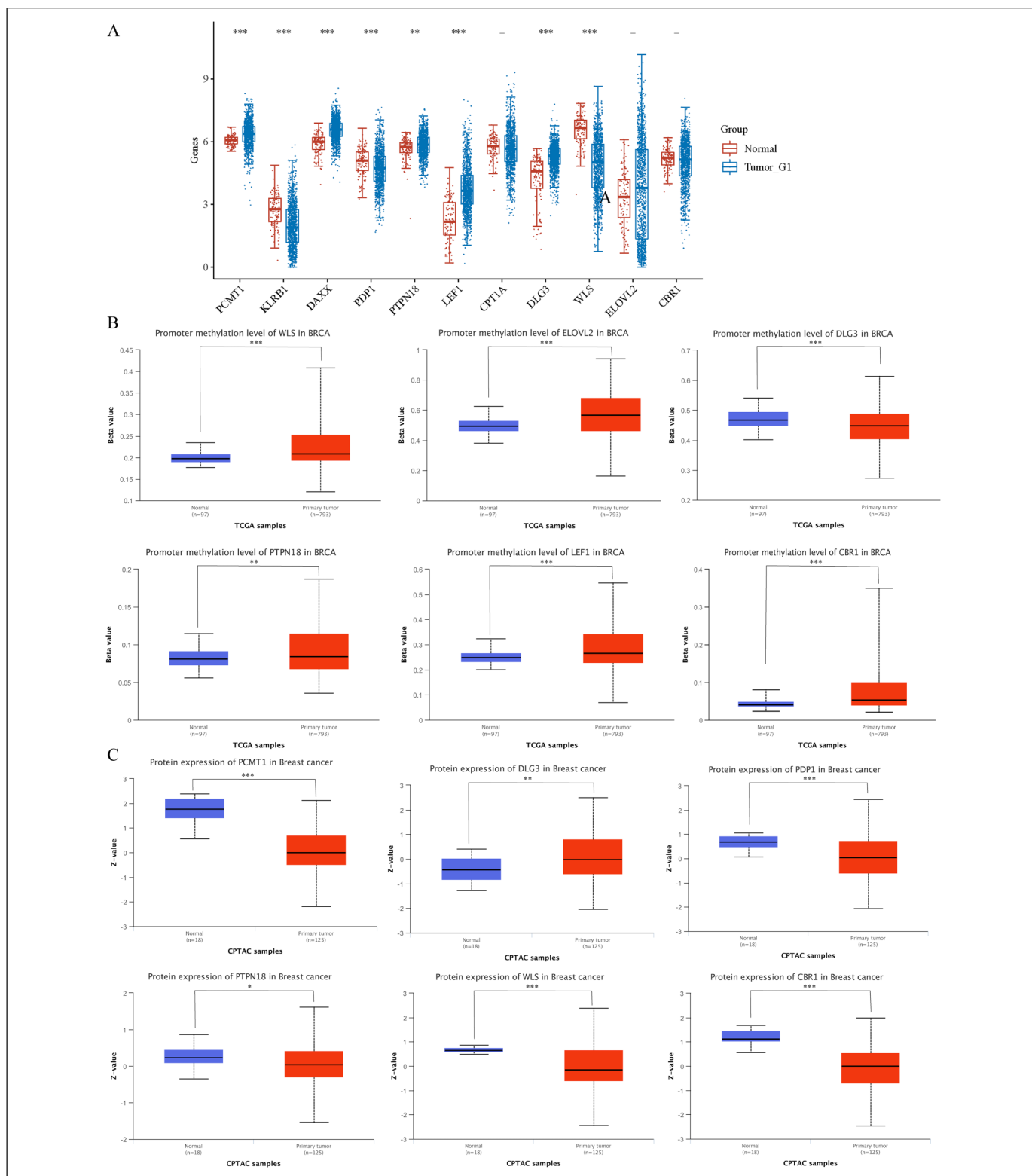


Figure 6. Expression and methylation analysis of prognostic model genes. (A) The expression distribution of 11 genes in breast cancer tissue and normal tissue. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, – P no statistically significant. (B) Promoter methylation levels of WLS, ELOVL2, DLG3, PTPN18, LEF1, and CBR1 in breast cancer. (C) The protein levels of PCMT1, DLG3, PDP1, PTPN18, WLS, and CBR1 in breast cancer.

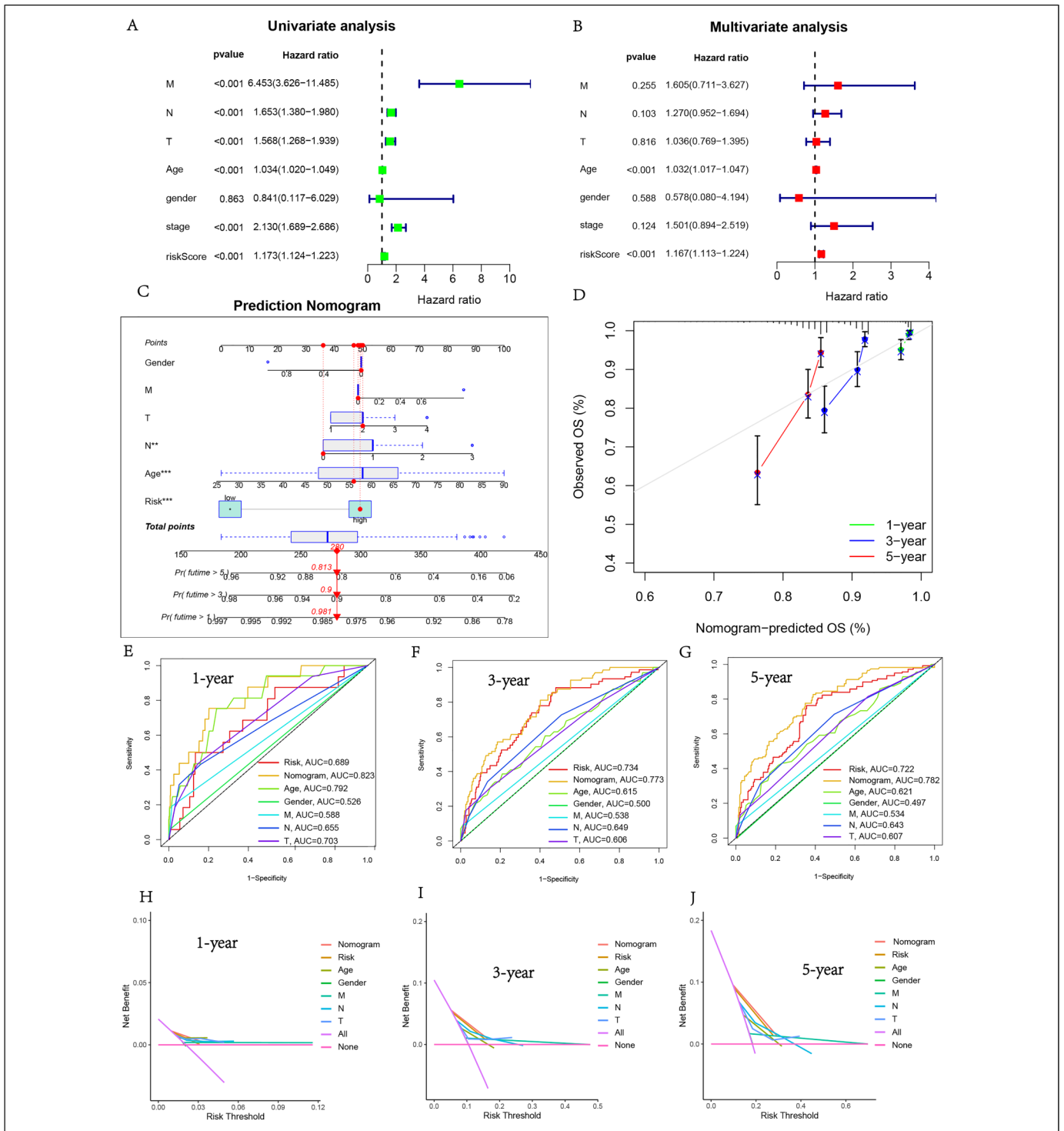


Figure 7. Construction of the nomogram. (A) Univariate forest plot for clinical features and the risk score. (B) Multivariate forest plot for clinical features and the risk score. (C) Establishment of a nomogram to predict prognostic risk. (D) The nomogram calibration curve. (E–G) Receiver operating characteristic (ROC) analysis of the nomogram, clinical features, and risk score at 1, 3, and 5 years. (H–J) Decision curve analysis (DCA) of the nomogram for predicting 1-, 3-, and 5-year overall survival.

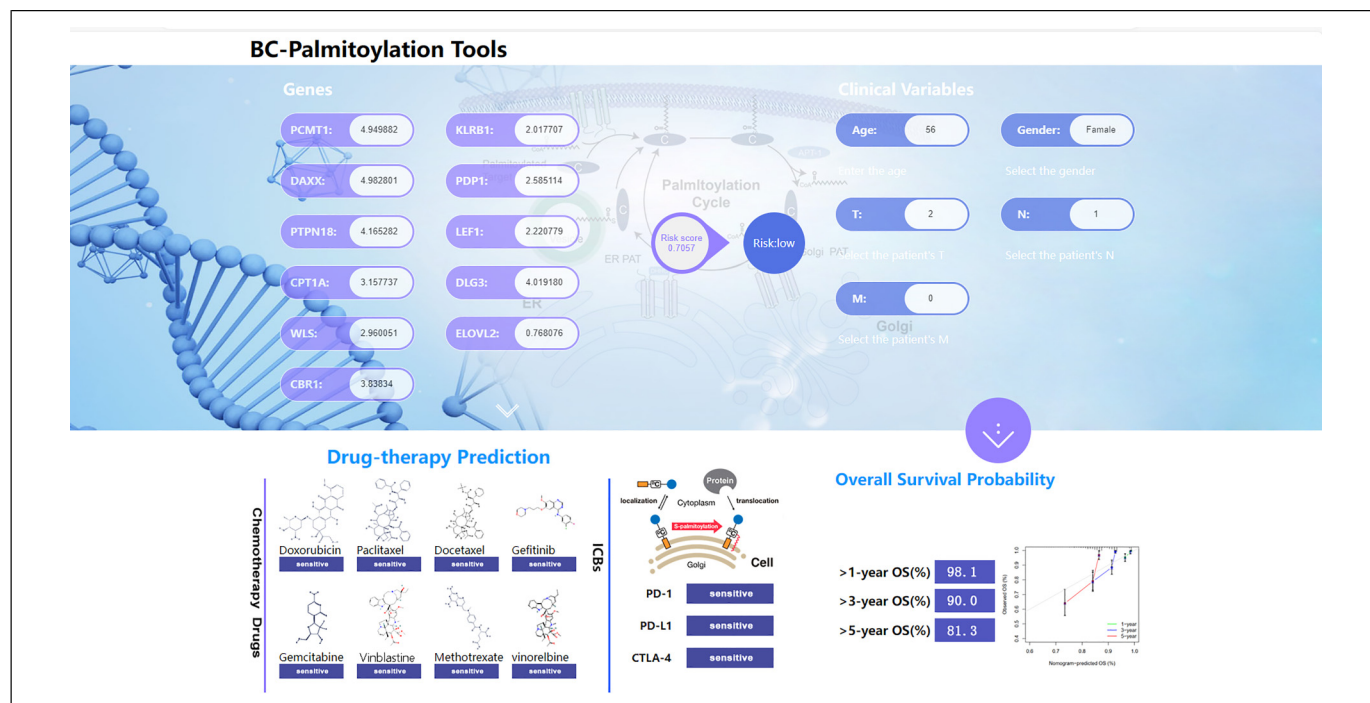


Figure 8. Website application of the BC-palmitoylation tool.

and steroid biosynthesis in the high-risk group. In contrast, the chemokine signaling pathway, cytokine–cytokine receptor interaction, hematopoietic cell lineage, primary immunodeficiency, and T-cell receptor signaling pathway were enriched in the low-risk group. Notably, the chemokine signaling pathway plays a crucial role in immune system function, as it regulates the recruitment and activation of immune cells in response to infections or inflammation.²⁵ Similarly, the T-cell receptor signaling pathway is pivotal in the adaptive immune response.²⁶ Compared to the high-risk group, the low-risk group exhibited higher immune scores, stromal scores, and estimate scores, as well as lower tumor purity. The infiltration of immune cells such as B cells, CD8+ T cells, CD4+ T cells, $\gamma\delta$ T cells, NK cells, and dendritic cells increased in the low-risk group, and the infiltration of these cells was negatively correlated with the risk score. Immune cells, critical components of the TME, can either inhibit or promote tumor growth.²⁷ NK cells and CD8+ T cells are the primary effector cells of the antitumor immune response and can kill tumor cells.²⁸ Nevertheless, the depletion of CD8+ T cells in tumors leads to the loss of killing effects.²⁹ CD4+ T cells, previously thought to primarily assist in regulating the function of other immune cells such as CD8+ T cells, have now been shown to possess direct antitumor activity.³⁰ Although a small subgroup, $\gamma\delta$ T cells exhibit the strongest tumor-killing effect among T cells and can recognize and destroy tumors earlier.³¹ B cells can exert their antitumor effect through antibody-mediated cytotoxicity and complement activation.³² Dendritic cells are the most powerful antigen-presenting cells and can stimulate T-cell proliferation and activation to kill tumor cells.³³ M2

macrophages, a major component of the TME, promote tumor cell growth, invasion, metastasis, and angiogenesis through various pathways, participate in tumor immune escape, and inhibit the antitumor effect of T cells.³⁴ In this research, the infiltration of M0 and M2 macrophages was greater in the high-risk group than in the low-risk group. Furthermore, we investigated the correlation between the risk score and the response to immune checkpoint inhibitors and chemotherapy drugs. Our findings indicate that the low-risk group is more responsive to immunotherapy and chemotherapy than the high-risk group, potentially due to their enhanced antitumor immune activity.

Through this study, a comparative analysis was performed on the mRNA and protein expression levels of 11 palmitoylation-related genes (PDP1, CPT1A, ELOVL2, KLRB1, WLS, PCMT1, DAXX, CBR1, LEF1, PTPN18, and DLG3) in breast cancer tissues compared to normal tissues. In addition, the promoter methylation levels and CNVs of these genes were investigated. The objective was to identify potential biomarkers and therapeutic targets for breast cancer. Furthermore, the analyses of promoter methylation and CNVs provided deeper insights into the molecular mechanisms underlying the development and progression of breast cancer.

CPT1A, essential for fatty acid metabolism, is implicated in the progression of several malignancies, including breast cancer, prostate cancer, and lymphoma.³⁵ Elevated CPT1A levels, as demonstrated in this study, correlate with poorer breast cancer outcomes, substantiating its role in cancer prognosis. DLG3, encoding a protein in the membrane-associated guanylate kinase (MAGUK) family,³⁶ is posited as a tumor suppressor. Its overexpression has been shown to inhibit

malignant cell proliferation, invasion, and migration.³⁷ PDP1 is critical for regulating cellular metabolism through the dephosphorylation and activation of the pyruvate dehydrogenase complex. PCMT1, an essential enzyme for protein repair, is significantly more expressed in breast cancer cells than in normal cells, and its overexpression correlates with enhanced growth and migration.³⁸ ELOVL2, essential for synthesizing very long-chain fatty acids, acts as a tumor suppressor by reducing breast cancer metastasis and endocrine therapy resistance.^{39,40} Our analysis revealed that methylation levels of the ELOVL2 promoter are higher in breast cancer tissues compared to normal tissues. Nevertheless, no significant differences in mRNA and protein expression levels were detected between the two groups. This suggests that the tumor-suppressive function of ELOVL2 is epigenetically regulated rather than via transcriptional or translational mechanisms. KLRB1, also known as CD161, encodes the protein killer cell lectin-like receptor subfamily B member 1, which is expressed on NK cells and some T cells and plays a role in immune recognition and activation.⁴¹ In our analysis, KLRB1 expression was significantly reduced in breast cancer tissues compared to normal breast tissue. Notably, higher KLRB1 expression was associated with improved prognosis in breast cancer patients. WLS, encoding a protein crucial for the Wnt signaling pathway, which governs cell growth, differentiation, and development, is dysregulated in various diseases, including cancer.^{42,43} This investigation demonstrated that both mRNA and protein levels of WLS were significantly lower in breast cancer tissues compared to adjacent non-cancerous tissues, potentially linked to promoter hypermethylation. DAXX, a multifunctional protein, participates in DNA repair, apoptosis, and transcription regulation.⁴⁴ Our findings indicate significantly higher mRNA levels in cancerous tissues, with no notable differences in protein levels or promoter methylation between the tissues. PTPN18, a gene involved in cell growth and differentiation, has been shown to inhibit breast cancer metastasis.⁴⁵ We assessed mRNA expression and promoter methylation of PTPN18 in breast versus normal tissues, finding significantly higher levels in cancerous tissues with lower protein levels. LEF1, a protein in the Wnt signaling pathway, influences cell proliferation, differentiation, and survival.⁴⁶ Associated with breast cancer, LEF1 may enhance tumor growth and metastasis. In this investigation, increased mRNA expression and promoter methylation levels of LEF1 were observed in breast cancer tissues compared to normal tissues, while protein levels showed no significant difference. CBR1 encodes an enzyme critical for metabolizing various drugs and exogenous substances.⁴⁷ This study found higher CBR1 promoter methylation and lower protein levels in breast cancer tissues, with no significant differences in mRNA expression compared to normal tissues.

These insights collectively enhance our understanding of the molecular mechanisms involved in breast cancer, offering potential targets for therapeutic intervention and improving prognostic assessments. However, this study has some limitations, mainly the retrospective nature of the analysis and

reliance on public datasets such as TCGA and GEO. The retrospective design limits our ability to establish causality and might affect the extrapolation of the results to broader clinical practices. Furthermore, the public datasets predominantly feature more prevalent or better-documented subtypes, potentially introducing biases that skew our analysis and affect the generalizability of our model across less common subtypes. While we have strived to mitigate these issues by conducting rigorous statistical analyses, the inherent limitations of the data source cannot be completely overcome. Recognizing these factors is crucial for interpreting the results accurately and applying our findings in clinical settings.

Conclusion

In summary, this study significantly advances our understanding of the prognostic implications of palmitoylation-related genes in breast cancer, highlighting the potential of these molecular markers to predict clinical outcomes and guide personalized treatment strategies. A robust prognostic model has been developed and integrated into a nomogram, effectively predicting survival and treatment responses. This model is accessible through the BC-Palmitoylation Tool, facilitating its clinical application. Despite its reliance on retrospective data, the findings offer a promising direction for improving clinical management and patient outcomes in breast cancer. Prospective studies and more diverse datasets are required to validate these findings further and ensure their applicability across various demographic and genetic backgrounds.

Author Contributions

All authors participated in data acquisition. T.-L.Z., H.-J.H., H.-X.Z., and H.-H.H. contributed to the conception and design of the study. T.-L.Z., H.-J.H., H.-X.Z., H.-H.H., B.H., W.-D.Z., T.Y., S.-Y.W. and J.-D.Z performed the data analysis and interpretation. T.-L.Z., H.-J.H., H.-X.Z. and H.-H.H. contributed to the drafting and revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

All raw data used in this work can be obtained from the GEO with accession number GSE20685 (<https://www.ncbi.nlm.nih.gov/geo/>). Additional data and materials are available from the University of California, Santa Cruz (UCSC) Xena browser (<https://xenabrowser.net/>).

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The present study was supported by the National Natural Science Foundation of China (grant number 82003802), the Natural Science Foundation of Hunan Province (grant number 2019JJ50542 and 2023JJ50156), the

Science and Technology Program of Hunan Health Commission (grant number 20201978), the China Scholarship Council (grant number 201808430085), Clinical Research 4310 Program of the First Affiliated Hospital of the University of South China (grant number 20224310NHYCG04), the Science and technology innovation Program of Hengyang City (grant number 202250045223).


Informed Consent Statement

Given that our research exclusively entailed the analysis of de-identified patient data, the requirement for individual patient consent was waived.

Institutional Review Board Statement

The need for ethical review and approval was waived due to the nature of our research, which solely encompassed the analysis of anonymized patient data without employing any identifiable individual information.

ORCID iD

Taolan Zhang MD  <https://orcid.org/0000-0002-0049-128X>

Supplemental Material

Supplemental material for this article is available online.

References

- 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(3):209–249.
- Mahoney MC, Bevers T, Linos E, Willett WC. Opportunities and strategies for breast cancer prevention through risk reduction. *CA Cancer J Clin.* 2008;58(6):347–371.
- Allemani C, Matsuda T, Carlo VD, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet.* 2018;391(10125):1023–1075.
- Jiang H, Zhang X, Chen X, Aramsangtienchai P, Tong Z, Lin H. Protein lipidation: occurrence, mechanisms, biological functions, and enabling technologies. *Chem Rev.* 2018;118(3):919–988.
- Anderson AM, Ragan MA. Palmitoylation: a protein S-acylation with implications for breast cancer. *NPJ Breast Cancer.* 2016;2:16028.
- Adlanmerini M, Solinhac R, Abot A, et al. Mutation of the palmitoylation site of estrogen receptor alpha in vivo reveals tissue-specific roles for membrane versus nuclear actions. *Proc Natl Acad Sci U S A.* 2014;111(2):E283–E290.
- Tonn Eisinger KR, Woolfrey KM, Swanson SP, et al. Palmitoylation of caveolin-1 is regulated by the same DHHC acyltransferases that modify steroid hormone receptors. *J Biol Chem.* 2018;293(41):15901–15911.
- Liu Z, Xiao M, Mo Y, et al. Emerging roles of protein palmitoylation and its modifying enzymes in cancer cell signal transduction and cancer therapy. *Int J Biol Sci.* 2022;18(8):3447–3457.
- Huang J, Li J, Tang J, et al. ZDHHC22-mediated mTOR palmitoylation restrains breast cancer growth and endocrine therapy resistance. *Int J Biol Sci.* 2022;18(7):2833–2850.
- Babina IS, McSherry EA, Donatello S, Hill AD, Hopkins AM. A novel mechanism of regulating breast cancer cell migration via palmitoylation-dependent alterations in the lipid raft affiliation of CD44. *Breast Cancer Res.* 2014;16(1):R19.
- Yang Y, Hsu JM, Sun L, et al. Palmitoylation stabilizes PD-L1 to promote breast tumor growth. *Cell Res.* 2019;29(1):83–86.
- Ritchie ME, Phipson B, Wu D, et al. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
- Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
- Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* 2009;25(8):1091–1093.
- Schriml LM, Munro JB, Schor M, et al. The human disease ontology 2022 update. *Nucleic Acids Res.* 2022;50(D1):D1255–D1261.
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 2010;33(1):1–22.
- Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–15550.
- Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics.* 2013;14:7.
- Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013;4:2612.
- Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12(5):453–457.
- Charoentong P, Finotello F, Angelova M, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 2017;18(1):248–262.
- Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One.* 2014;9(9):e107468.
- Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* 2008;26(8):1364–1370.
- Wang J, Knaut H. Chemokine signaling in development and disease. *Development.* 2014;141(22):4199–4205.
- Shah K, Al-Haidari A, Sun J, Kazi JU. T cell receptor (TCR) signaling in health and disease. *Signal Transduct Target Ther.* 2021;6(1):412.
- Anderson NM, Simon MC. The tumor microenvironment. *Curr Biol.* 2020;30(6):R921–R925.

28. de Visser KE, Joyce JA. The evolving tumor microenvironment: from cancer initiation to metastatic outgrowth. *Cancer Cell*. 2023;41(3):374–403.
29. Philip M, Schietinger A. CD8(+) T cell differentiation and dysfunction in cancer. *Nat Rev Immunol*. 2022;22(4):209–223.
30. Speiser DE, Chijioke O, Schaeuble K, Munz C. CD4(+) T cells in cancer. *Nat Cancer*. 2023;4(3):317–329.
31. Mensurado S, Blanco-Dominguez R, Silva-Santos B. The emerging roles of gammadelta T cells in cancer immunotherapy. *Nat Rev Clin Oncol*. 2023;20(3):178–191.
32. Engelhard V, Conejo-Garcia JR, Ahmed R, et al. B cells and cancer. *Cancer Cell*. 2021;39(10):1293–1296.
33. Del Prete A, Salvi V, Soriani A, et al. Dendritic cell subsets in cancer immunity and tumor antigen sensing. *Cell Mol Immunol*. 2023;20(5):432–447.
34. Chen S, Saeed A, Liu Q, et al. Macrophages in immunoregulation and therapeutics. *Signal Transduct Target Ther*. 2023;8(1):207.
35. Das M, Giannoudis A, Sharma V. The role of CPT1A as a biomarker of breast cancer progression: a bioinformatic approach. *Sci Rep*. 2022;12(1):16441.
36. Roberts S, Delury C, Marsh E. The PDZ protein discs-large (DLG): the ‘Jekyll and Hyde’ of the epithelial polarity proteins. *FEBS J*. 2012;279(19):3549–3558.
37. Li D, Hu X, Yu S, et al. Silence of lncRNA MIAT-mediated inhibition of DLG3 promoter methylation suppresses breast cancer progression via the hippo signaling pathway. *Cell Signal*. 2020;73:109697.
38. Guo J, Du X, Li C. PCMT1 is a potential prognostic biomarker and is correlated with immune infiltrates in breast cancer. *Biomed Res Int*. 2022;2022:4434887.
39. Zhu M, Zhang J, Li G, Liu Z. ELOVL2-AS1 inhibits migration of triple negative breast cancer. *PeerJ*. 2022;10:e13264.
40. Jeong D, Ham J, Kim HW, et al. ELOVL2: a novel tumor suppressor attenuating tamoxifen resistance in breast cancer. *Am J Cancer Res*. 2021;11(6):2568–2589.
41. Cheng X, Cao Y, Wang X, et al. Systematic pan-cancer analysis of KLRB1 with prognostic value and immunological activity across human tumors. *J Immunol Res*. 2022;2022:5254911.
42. Lu D, Li Y, Liu QR, et al. Wls promotes the proliferation of breast cancer cells via Wnt signaling. *Med Oncol*. 2015;32(5):140.
43. Zheng D, Jiang C, Yan N, et al. Wntless (Wls): a prognostic index for progression and patient survival of breast cancer. *Onco Targets Ther*. 2020;13:12649–12659.
44. Mahmud I, Tian G, Wang J, et al. DAXX drives de novo lipogenesis and contributes to tumorigenesis. *Nat Commun*. 2023;14(1):1927.
45. Wang T, Ba X, Zhang X, et al. Nuclear import of PTPN18 inhibits breast cancer metastasis mediated by MVP and importin beta2. *Cell Death Dis*. 2022;13(8):720.
46. Prieto-Vila M, Shimomura I, Kogure A, et al. Quercetin inhibits Lef1 and resensitizes docetaxel-resistant breast cancer cells. *Molecules*. 2020;25(11):2576.
47. Lal S, Sandanaraj E, Wong ZW, et al. CBR1 and CBR3 pharmacogenetics and their influence on doxorubicin disposition in Asian breast cancer patients. *Cancer Sci*. 2008;99(10):2045–2054.