

Identification of prognostic and disulfidptosis-related genes signature in breast cancer

Jun-Gang Wang and Yun Zhang*

Department of general surgery (thyroid and breast surgery), Haiyan People's Hospital, Haiyan country, Jiaxing City, Zhejiang Province, 314300, China.

Corresponding author: Prof. Yun Zhang

E-mail: 33304124@qq.com

Genet. Mol. Res. 23 (3): gmr2338

Received June 10, 2024

Accepted August 31, 2024

Published September 14, 2024

DOI <http://dx.doi.org/10.4238/gmr2338>

ABSTRACT. Purpose. To explore the association between disulfidptosis and breast cancer. **Methods.** The METABRIC cohort containing clinical data and gene expression data was downloaded. Based on 9 previously reported disulfidptosis-related genes, we performed correlation analysis, Cox regression analysis, and least absolute shrinkage and selection operator (LASSO) regression analysis to screen genes and then constructed a prognostic model, the effectiveness of which was validated in external data sets including METABRIC validation data and GSE20685. A risk score was calculated then and samples were divided into high- and low-risk group based on the median value. Enrichment analysis, immune microenvironmental analysis, and drug sensitivity analysis were also carried out. **Results.** A total of 7 genes was finally identified after LASSO regression analysis. We constructed a nomogram predict model. The area under curve (AUC) values for survival rates at most time points were larger than 0.650, no matter in the METABRIC training data, METABRIC validation data, or in the GSE20685 data. Dividing into two groups based on the median risk score can effectively distinguish the prognosis of patients ($P < 0.001$ in the internal group, $P = 0.064$ in the METABRIC training data, and $P < 0.001$ in the GSE20685 data). High-risk group was associated with a significantly lower infiltration level of cytotoxic lymphocytes compared with low-risk group ($P = 0.026$), while the infiltration levels of stromal cells were the opposite. **Conclusions.** The study provided valuable insights into

the association between breast cancer and disulfidptosis. The findings needs to be validated in further studies.

Key words: Disulfidptosis; Prognosis; LASSO; Breast cancer

INTRODUCTION

Breast cancer poses a major threat to women's health in the world. The latest report shows that breast cancer is now the most common type of cancer worldwide, accounting for about one eighth of the global new cancer cases (American Cancer Society, 2023). Breast cancer may occur at any age after puberty, and its incidence rate increases with age (Shah A et al., 2022). Although in the past decades, scientific researchers have achieved remarkable progress in the prevention and treatment of breast cancer, its mortality rate still ranks fifth among all the cancer types worldwide (Sung H et al., 2021). Therefore, it is necessary to carry out more clinical and translational research on breast cancer, so as to find more effective therapeutic targets and reduce the disease burden.

Metabolic reprogramming is plays an important role in cancer biology, which often leads to cancer cells being highly dependent on specific nutrients or metabolic pathways. Selectively and targeted killing of cancer cells has been widely used in today's era of precision oncology (Stine ZE et al., 2022). Programmed cell death plays an important role in cancer metabolic therapy. For example, there is a growing number of evidence that ferroptosis, one of the programmed cell death types that is dependent on iron metabolism and characterized by abnormal intracellular accumulation of lipid peroxides, is involved in many aspects of breast cancer, such as tumorigenesis, development, and metastasis (Sui S et al., 2022). Therefore, inducing ferroptosis in cancer cells can effectively inhibiting tumor growth, which has become a promising strategy for treating breast cancer (Li Z et al., 2020; Liu Y et al., 2022).

Recently, a novel programmed cell death form has been revealed, which is named disulfidptosis (Liu X et al., 2023). Liu et al. reported that under glucose starvation conditions, high expression of member 11 of the solute carrier family 7 (SLC7A11) accelerates the depletion of nicotinamide adenine dinucleotide phosphate (NADPH) in the renal cancer cell cytoplasm. The accumulation of irreducible disulfides induces disulfide stress, ultimately leads to disulfidptosis. SLC7A11 is responsible for the uptake of cysteine, and in renal cancer cells, as the concentration of SLC7A11 increases, the uptake rate of cysteine also accelerates. Under normal circumstances, NADPH can neutralize disulfide stress and so maintain intracellular homeostasis through its reducing ability. However, glucose starvation conditions limit the production of NADPH through the pentose phosphate pathway. Additionally, cystine accumulation mediated by SLC7A11 further reduces the NADPH level, and induces the disulfide accumulation. Disulfide bonds between actin cytoskeleton proteins was triggered then, which eventually leads to disulfidptosis.

Although the role of earlier discovered programmed cell death forms, such as ferroptosis and pyroptosis in breast cancer have been studied in many studies, the role of disulfidptosis remains unclear. In the present study, through identifying disulfidptosis-related genes and performing the subsequent analyses using corresponding data, we explored the potential role of disulfidptosis in the prognosis, microenvironment, and drug sensitivity of breast cancer.

MATERIALS AND METHODS

Source of data

The METABRIC cohort containing clinical data and gene expression data was downloaded from <https://www.cbioportal.org/>. The dataset included a total of 1904 tumor samples with clinical information and mRNA expression data. We randomly selected 1700 samples as our training dataset, and the remaining samples were used as METABRIC validation data for external validation. To further confirm the robustness of our results, we also downloaded the GSE20685 dataset for external validation.

Identification of disulfidptosis-related genes

The present paper aimed to screen disulfidptosis-related genes with prognostic value. Firstly, we selected a total of 9 previously reported disulfidptosis-related genes. Secondly, we performed GSEA analysis on the 9 genes, and then we performed single-gene correlation analysis on the GSEA score. We chose genes with adjusted P value of the correlation results < 0.05 and correlation coefficient > 0.5 as disulfidptosis-related genes and subjected them to further analysis.

Prognostic model construction and validation

Cox regression analyses were performed and then genes with P value < 0.05 were selected for the subsequent least absolute shrinkage and selection operator (LASSO) regression analysis, by the use of “fit_lasso” function of the “hdnom” R package. Specifically, we set the following parameters as follows: nfolds = 10, seed = 1001, and rule = “lambda.1se”. Next, we used “as_nomogram” function to construct a nomogram predict model. Internal validation and external validation were then performed using “validate” function and “validate_external” function on METABRIC training data, and METABRIC validation data as well as GSE20685 dataset, respectively. Calibrate curves were plotted using “calibrate” function for internal calibration and “calibrate_external” function for external calibration. We used “kmplot” function to evaluate the prognostic predict ability of the model. The parameters “group,name” was set as “c(‘High risk’, ‘Low risk’)”

We used “predict” function to calculate a risk score for each sample. Then, based on the median risk score, the samples were divided into two groups: the high-risk group and the low-risk group. The grouping results were used for subsequent analysis.

Differential analysis based on grouping results

We used “limma” R package to perform differential analysis based on the above grouping results and identified genes with P value less than 0.05 and $\log_2(\text{fold change})$ larger than 0.5 as differentially expressed genes. We then used “pheatmap” function to plot the heatmap and used “ggplot” function to plot the volcano plot, respectively, based on the identified differentially expressed genes. Enrichment analyses were then performed using gene set enrichment analysis (GSEA) based on Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Hallmarks gene sets to investigate the biological functions of differentially expressed genes.

Analysis on tumor microenvironment

ESTIMATE and MCP-counter are two commonly used algorithms for analyzing the tumor microenvironment (Yoshihara K et al., 2013; Becht E et al., 2016). In the present study, we used ESTIMATE and MCP-counter methods to estimate the relative infiltration level of immune and stromal cell types and then compare them between the high-risk and low-risk groups.

Drug sensitivity analysis

We utilized the “oncoPredict” R package to predict the IC_{50} values of each sample to different compounds, based on the Genomics of Drug Sensitivity in Cancer (GDSC, version 2) data, which contains the IC_{50} value data of 198 compounds in 809 cell lines (Maeser D et al., 2021).

Statistical analysis

All the above analyses were performed using R 4.3.2, and figures were generated used the corresponding functions in R software and then edited using Adobe Illustrator 2020 (version 24.0.3). P value less than 0.05 was considered as statistically significant.

RESULTS

Identification of disulfidptosis-related genes

In the present study, we used a previously reported genelist that are associated with disulfidptosis, which is composed of the following 9 genes: SLC7A11, SLC3A2, RPN1, NCKAP1, WASF2, CYFIP1, ABI2, BRK1, and RAC1. Firstly, we performed GSVA algorithm based on the genelist to generate disulfidptosis score. We obtained 452 genes using a correlation coefficient greater than 0.5 as the screening criterion, which were then shortened to 54 genes through univariate cox regression analysis with p value less than 0.2. Finally, a total of 7 genes was obtained and used for the later analyses after LASSO regression analysis, with penalty coefficients not equal to zero: ESRP1, ATXN3, MAD2L1, RPL6, PCDH17, ARFGAP1, and VANGL1.

Construction of Nomogram model

We used “as_nomogram” function of the “hdnom” R package to plot Nomogram considering the above 7 genes (**Figure 1A**). ROC analysis was then performed to assess the efficiency of the nomogram, the results showed that the AUC values for 24-month, 30-month, 36-month, 42-month, and 48-month survival rates were all larger than 0.650 (**Figure 1B**). We used calibration curves to evaluate the predictive accuracy of nomogram model. Calibration curves of 3-year and 5-year were relatively close to 45-degree diagonal lines, which indicates a relatively good agreement between the predicted and observed survival probabilities in the METABRIC breast cancer cohort. (**Figure 1C**).

We next performed external validation to further confirm the reliability of the nomogram prediction model in external datasets: METABRIC training data and GSE20685 data. As a result, the AUC values of 12-month and 24-month survival rates were both larger than 0.650 in the METABRIC validation data (**Figure 2A**) and the AUC values of 3-month, 6-month and 12-month survival rates were both larger than 0.690 in the GSE20685 data (**Figure 2B**). Calibration curves of

3-year and 5-year in the METABRIC data were showed in (Figure 2C), and Calibration curves of 1-year in the GSE20685 data were showed in (Figure 2D), respectively.

Then, we divided METABRIC breast cancer samples into high- and low-risk groups based on the median value of the risk score predicted by the model. The grouping results can effectively distinguish the prognosis, that is, high-risk group was associated with a significantly worse survival rates compared to the low-risk group ($P < 0.001$ in the internal group, $P = 0.064$ in the METABRIC training data, and $P < 0.001$ in the GSE20685 data).

Differential gene analysis between low-risk and high-risk groups

We used the “limma” R package to identify the differentially expressed genes between low- and high-risk groups. we identified a total of 1648 differentially expressed genes (736 up-regulated and 912 down-regulated) (Figure 3A and 3B). Next, we performed GO and KEGG enrichment analysis to explore the functions of these genes. The GO entries primarily enriched on the following pathways: chromosome segregation, nuclear division, microtubule binding and cytoskeletal motor activity, etc. (Figure 3C) The KEGG enrichment results showed that cell cycle, human T cell leukemia virus 1 infection and cellular senescence, etc. (Figure 3D).

We also performed GSEA on METABRIC breast cancer cohort using HALLMARKS gene sets. The three most activated pathways were UV response down, estrogen response, and epithelial-

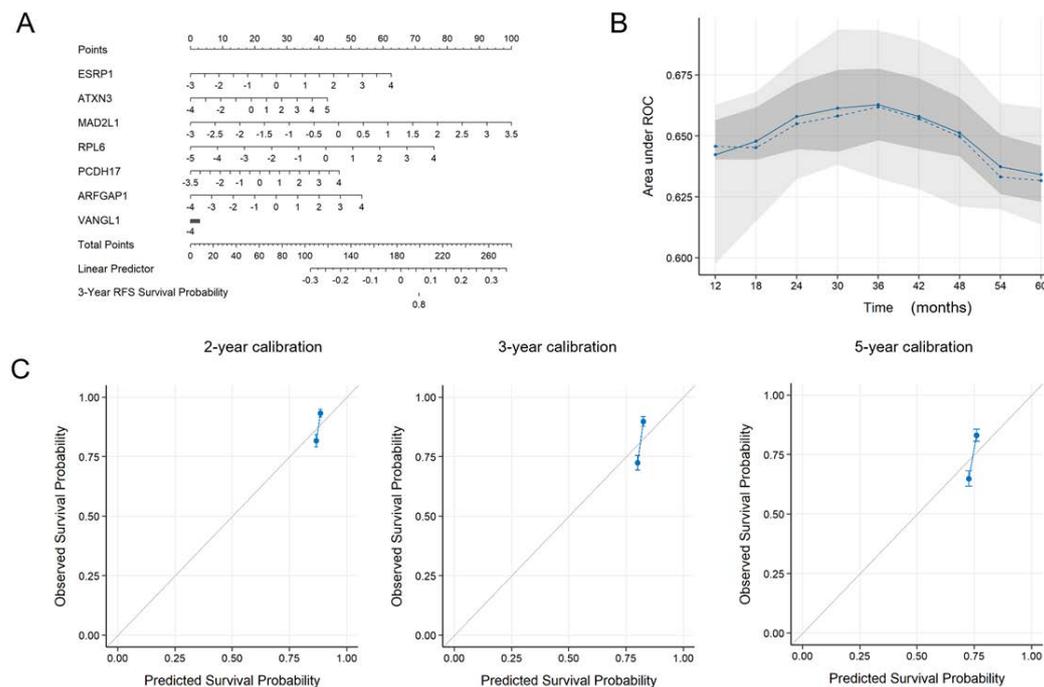


Figure 1. Prognostic model construction.

(A) Nomogram constructed in the METABRIC training data.

(B) AUC values for predicting survival probability at different times in the METABRIC training data.

(C) Calibration curves for 2-year, 3-year, and 5-year survival probability in the METABRIC training data.

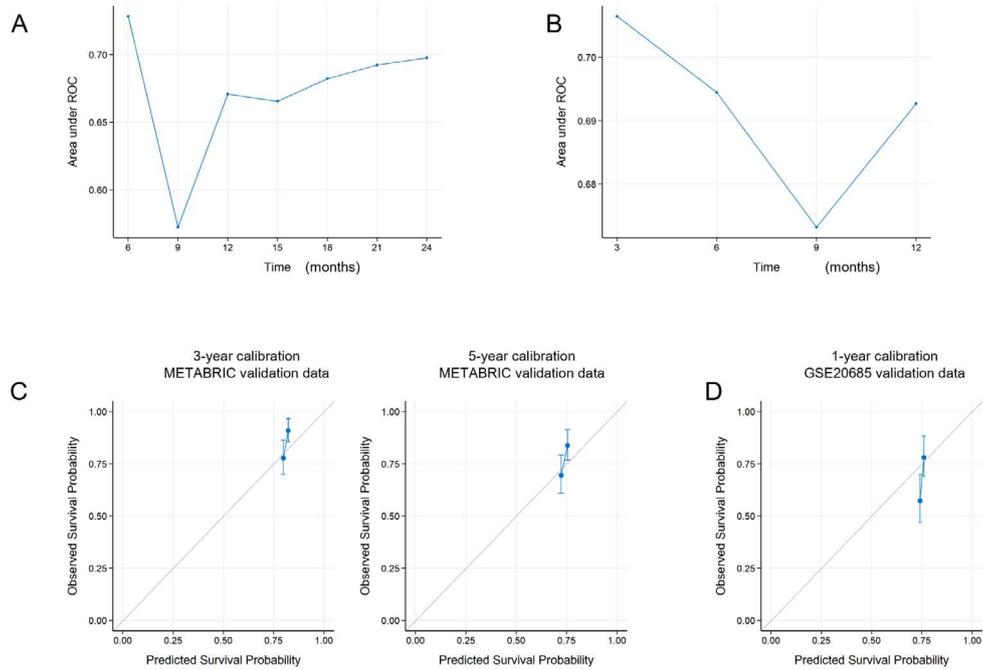


Figure 2. External validation.

- (A) AUC values for predicting survival probability at different times in the METABRIC validation data.
- (B) AUC values for predicting survival probability at different times in the GSE20685 data.
- (C) Calibration curves for 2-year and 3-year survival probability in the METABRIC validation data.
- (D) Calibration curves for 1-year survival probability in the GSE20685 data.

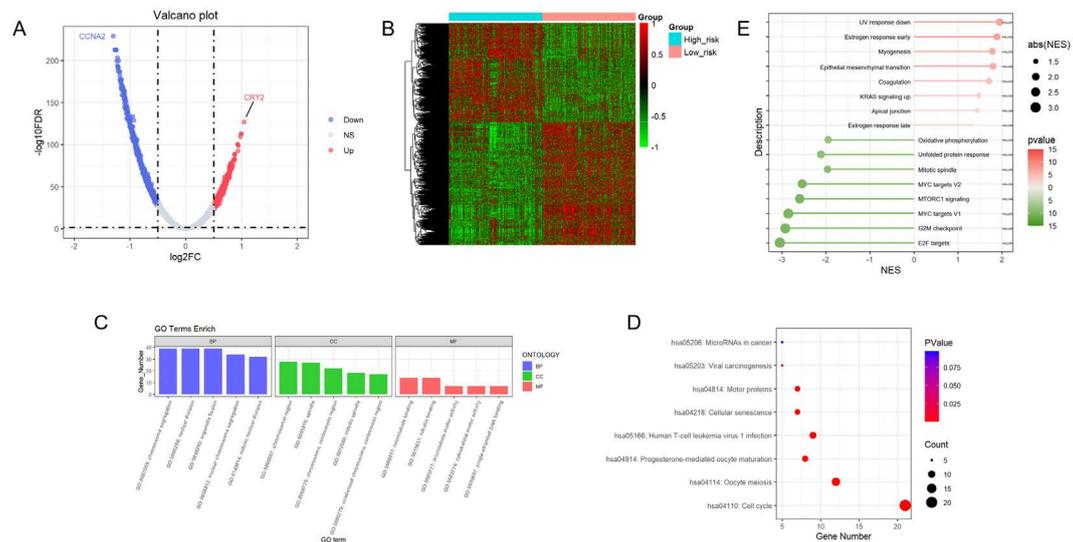


Figure 3. Differential analysis and enrichment analysis.

- (A-B) Volcano plot (A) and heatmap plot (B) of differentially expressed genes.
- (C-E) Enrichment analysis results based on GO terms (C), KEGG terms (D), and Hallmarks gene sets (E).

mesenchymal transition (**Figure 3E**), while the three most inhibited pathways were E2F targets, G2M checkpoint and MYC targets V1. (**Figure 3F**) Hallmark G2M checkpoint and HALLMARK E2F targets are involved in regulating the cell cycle, while the Hallmark epithelial-mesenchymal transition was related tumor migration and invasion.

Tumor micro-environment analysis

In the present study, we explored the impact of disulfidptosis-related genes on tumor micro-environment by the use of ESTIMATE and MCP-Counter algorithms. High-risk group was found to be associated with a significantly lower tumor purity ($P < 0.001$, **Figure 4A**), higher Stromal score ($P < 0.001$, **Figure 4B**), and ESTIMATE score ($P < 0.001$, **Figure 4C**) compared with low-risk group. The immune score was not significantly different between the two groups ($P = 0.15$) (**Figure 4D**). However, according to the MCP-counter results, high-risk group was associated with a significantly lower infiltration level of cytotoxic lymphocytes compared with the low-risk group ($P = 0.026$), while the infiltration level of stromal cells such as myeloid dendritic cells ($P < 0.001$), neutrophils ($P = 0.004$), endothelial cells ($P < 0.001$), and fibroblasts ($P < 0.001$) was opposite (**Figure 4E**). These

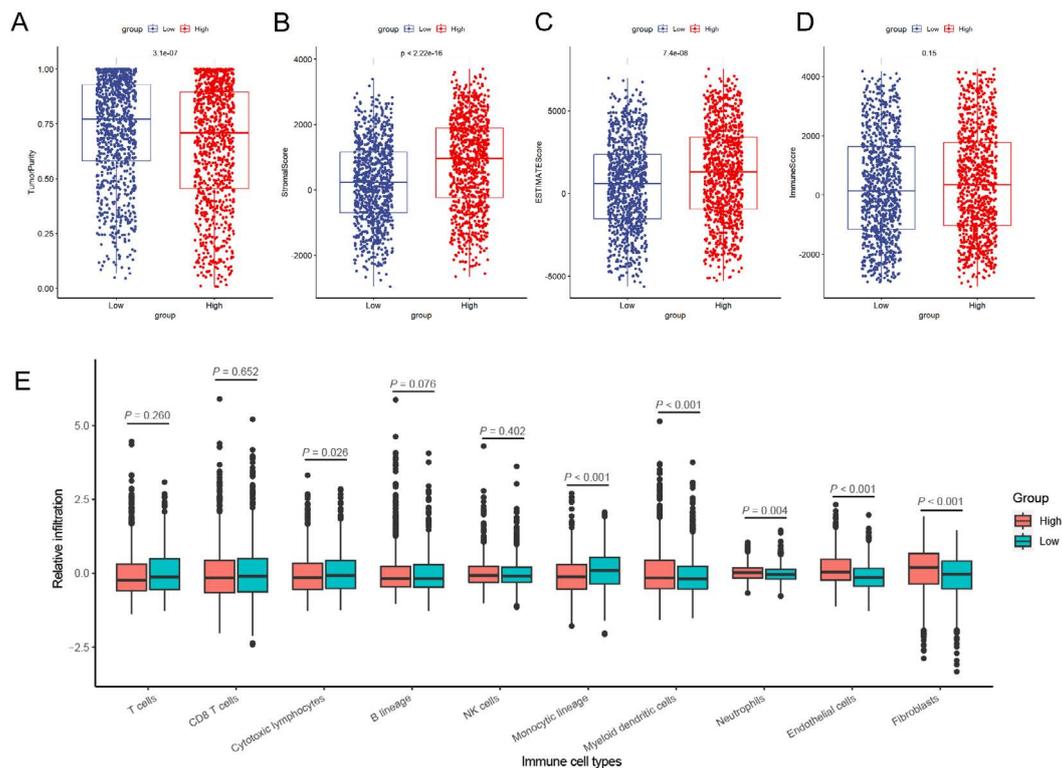


Figure 4. Immune environment analysis.

(A-D) The comparison of tumor purity score (A), stromal score (B), ESTIMATE score (C), and immune score (D) calculated by ESTIMATE algorithm between high-risk and low-risk groups.

(E) The comparison of infiltration levels of different immune cell types calculated by MCP-counter algorithm between high-risk and low-risk groups.

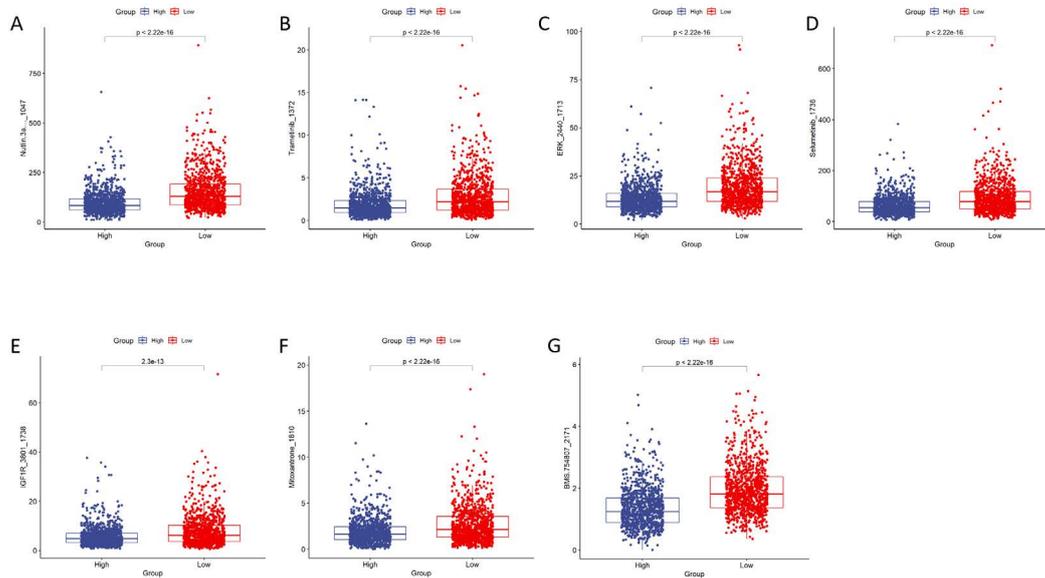


Figure 5. Drug sensitivity analysis.

(A-G) The comparison of the sensitivity to nutlin-3a (-) (A), trametinib (B), ERK_2440 (C), selumetinib (D), IGF1R_3801 (E), Mitoxantrone (F), and BMS-754807 (G).

findings indicates the stromal cells infiltrated in the high-risk group might attenuate the infiltrated lymphocytes' cytotoxic effect, which then promoted the immune escape and development of tumors.

Drug sensitivity prediction

Based on the GDSC v2 data, we used “oncoPredict” R package to predict the sensitivity of each sample. Then we compared the IC_{50} value between the two groups. We found that compared with the low-risk group, the high-risk group were associated with a significantly lower IC_{50} value to the following compounds: nutlin-3a (-) (Figure 5A), which targets MDM2; trametinib (Figure 5B) and selumetinib (Figure 5D), which target MEK1/2; ERK_2440 (Figure 5C), which targets ERK1/2; IGF1R_3801 (Figure 5E), which targets IGFR1; Mitoxantrone (Figure 5F), which targets Top2; and BMS-754807 (Figure 5G), which targets IGF1R and IR, which suggests the potential use of these drugs in the high-risk patients.

DISCUSSION

Tumor cells can undergo different pathways of death under different conditions. These pathways included programmed cell death like necrosis, apoptosis, and autophagy, which was important in regulating tumor and progression (Su Z et al., 2015; D'Arcy MS et al., 2019). It has been reported that inducing cell death pathways can effectively inhibit tumor growth. For example, it has been shown that combining ferroptosis inducers with other therapies can significantly improve the therapeutic effect on breast cancer (Desterke C et al., 2023; Gong G et al., 2023). Disulfidptosis, as a newly discovered programmed cell death pathway, is different from ferroptosis and cuproptosis. Unlike ferroptosis, whose mechanisms and application value in various cancers have been fully

studied, research on disulfidptosis in cancers has just begun, and its effect in breast cancer is unclear, which needs to be studied. In the present paper, we explored the potential role of disulfidptosis based on disulfidptosis-related genes.

In this study, we constructed a model with prognostic value and were disulfidptosis-related. The accuracy of the prediction model was validated in both the METABRIC validation data and external dataset. Based on the median value of risk score predicted by this model, samples were divided into high-risk group and low-risk group. The results of GO enrichment analyses indicated the importance of pathways related to the reproduction and inheritance of genes, such as chromosome segregation and nuclear division, etc. in high-risk group patients. The enrichment of these pathways suggests that high-risk populations may have experienced more cell division and proliferation, indicating the potential for the application of relevant target agents. This finding was also confirmed in KEGG enrichment analysis results, which indicated that cell cycle pathways were enriched in high-risk group patients. Additionally, the UV response pathway enriched by KALLMARKS gene sets and the human T cell leukemia virus 1 infection pathway in the KEGG enrichment analysis results indicated that high-risk patients' tumors might also have undergone immune microenvironment remodeling. We subsequently analyzed the tumor microenvironment using corresponding algorithms and found that high-risk patients' tumors were associated with a significantly higher infiltration level of stromal cells such as endothelial cells and fibroblasts, etc., which might attenuate the cytotoxic effect of infiltrated lymphocytes, as revealed by the significantly decreased infiltration of cytotoxic lymphocytes. Finally, drug sensitivity analyses suggested several agents that may be effective for high-risk populations.

It was reported that disulfidptosis remodel the tumor immune microenvironment in some other cancer types such as hepatocellular carcinoma (Tang J et al., 2024; Mulati Y et al., 2024) and prostate cancer (Mulati Y et al., 2024) etc. The advantage of this study lies in the use of external datasets to validate the accuracy of the prediction model and we predicted drug sensitivity. Further basic and translational research are warranted to confirm our findings and investigate the detailed role of disulfidptosis in breast cancer and its corresponding mechanism.

CONCLUSIONS

The study provided valuable insights into the association between breast cancer and disulfidptosis. The findings needs to be validated in further studies.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

COMPETING INTERESTS

The authors have no relevant financial or non- financial interests to disclose.

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Jun-Gang Wang. The first draft of the manuscript was written by Yun Zhang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

DATA AVAILABILITY

All the data analyzed in the present study are publicly available, with the source stated in the main text.

ETHICS APPROVAL

Not applicable.

CONSENT TO PARTICIPATE

Not applicable.

CONSENT TO PUBLISH

Not applicable.

REFERENCE

- American Cancer Society (2023). Cancer Facts & Figures for African American 2022–2024. American Cancer Society.
- Becht E, Giraldo NA, Lacroix L, Buttard B, et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome biology*. 17: 218.
- D'Arcy MS (2019) Cell death: a review of the major forms of apoptosis, necrosis and autophagy. *Cell Biol Int*. 43: 582-592.
- Desterke C, Xiang Y, Elhage R, Duruel C, et al. (2023). Ferroptosis Inducers Upregulate PD-L1 in Recurrent Triple-Negative Breast Cancer. *Cancers*. 16.
- Gong G, Ganesan K, Liu Y, Huang Y, et al. (2023). Danggui Buxue Tang improves therapeutic efficacy of doxorubicin in triple negative breast cancer via ferroptosis. *J Ethnopharmacol*. 323: 117655.
- Liu X, Nie L, Zhang Y, Yan Y, et al. (2023). Actin cytoskeleton vulnerability to disulfide stress mediates disulfidptosis. *Nature cell biology*. 25: 404-414.
- Liu Y, Hu Y, Jiang Y, Bu J, et al. (2022). Targeting ferroptosis, the achilles' heel of breast cancer: A review. *Frontiers in pharmacology*. 13: 1036140.
- Li Z, Chen L, Chen C, Zhou Y, et al. (2020). Targeting ferroptosis in breast cancer. *Biomarker research*. 8: 58.
- Maeser D, Gruener RF, Huang RS. (2021) oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Briefings in bioinformatics*. 22.
- Mulati Y, Lai C, Luo J, Hu J, et al. (2024). Establishment of a prognostic risk prediction model incorporating disulfidptosis-related lncRNA for patients with prostate cancer. *BMC cancer*. 24: 44.
- Shah A, Haider G, Abro N, Bhutto S, et al. (2022). Correlation Between Age and Hormone Receptor Status in Women With Breast Cancer. *Cureus*. 14: e21652.
- Stine ZE, Schug ZT, Salvino JM, Dang CV. (2022). Targeting cancer metabolism in the era of precision oncology. *Nature reviews Drug discovery*. 21: 141-62.
- Sui S, Xu S, Pang D. (2022). Emerging role of ferroptosis in breast cancer: New dawn for overcoming tumor progression. *Pharmacology & therapeutics*. 232: 107992.
- Sung H, Ferlay J, Siegel RL, Laversanne M, et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians*. 71: 209-249.

- Su Z, Yang Z, Xu Y, Chen Y, et al. (2015). Apoptosis, autophagy, necroptosis, and cancer metastasis. *Molecular cancer*. 14: 48.
- Tang J, Peng X, Xiao D, Liu S, et al. (2024). Disulfidptosis-related signature predicts prognosis and characterizes the immune microenvironment in hepatocellular carcinoma. *Cancer cell international*. 24: 19.
- Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature communications*. 4: 2612.