A Novel Ferroptosis-Related Gene Signature for Chemotherapy Resistance Prediction in Triple-negative Breast Cancer

Huan You (✉ yh981106@163.com)  
Affiliated Hospital of Nantong University, Nantong University

Hongyan Qian  
Affiliated Tumor Hospital of Nantong University

Shichen Miao  
Affiliated Hospital of Nantong University, Nantong University

Xuan Li  
Affiliated Hospital of Nantong University, Nantong University

Bingyi Liu  
Affiliated Hospital of Nantong University, Nantong University

Dan Zhang  
Affiliated Hospital of Nantong University, Nantong University

Yanpeng Chen  
Affiliated Hospital of Nantong University, Nantong University

Qichao Ni  
Affiliated Hospital of Nantong University, Nantong University

Chenyi Sheng  
Affiliated Hospital of Nantong University, Nantong University

Article

Keywords:

Posted Date: July 25th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3176896/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Taxanes are first-line chemotherapeutic agents for patients with triple-negative breast cancer (TNBC). However, resistance, the main cause of clinical treatment failure and poor prognosis, reduces their effectiveness and has become an increasingly important problem. Recently, a form of iron-dependent programmed cell death called ferroptosis was reported to play an important role in regulating tumor biological behavior. In this study, we revealed the prognostic significance of the ferroptosis-related gene (FERG) model and clarified that ferroptosis-related genes may be promising candidate biomarkers in cancer therapy. First, resistance-related FERGs were screened, and univariate Cox regression analysis was used to construct a prognostic model, including GRIK3, IDO1, and CLGN. Then, the patients with TNBC in the TCGA database were classified into high-risk and low-risk groups. The identification of TNBC in the TCGA database revealed that patients with high scores had a higher probability of dying earlier than those with low scores. Moreover, these three genes were associated with immune infiltrates and checkpoints in TNBC patients. In conclusion, this study suggested that FERGs are significantly associated with chemotherapy resistance in patients with TNBC and that these genes can be used as prognostic predictors in these patients and possibly for targeted therapy in the future.

Introduction

According to the latest global cancer data published by the International Agency for Research on Cancer, breast cancer in women is the most common cancer worldwide, exceeding lung cancer for the first time\(^1\). In addition, breast cancer is the most common malignancy and the leading cause of cancer-related death in women, and triple-negative breast cancer (TNBC) has one of the worst prognoses of all breast cancer variants\(^2\). TNBC is a subtype of breast tumor that lacks hormone receptor expression and human epidermal growth factor receptor 2 (HER2) gene amplification and exhibits high relapse risk, high metastasis, and low overall survival\(^3\). The current treatment options are very limited owing to the lack of effective therapeutic targets in TNBC\(^4\). At present, there are no clinical guidelines for TNBC treatment, and it is usually treated according to the conventional standard of care for breast cancer\(^5\). The conventional treatment modalities of breast cancer include surgery, radiotherapy and chemotherapy, endocrine therapy, and targeted therapy\(^6\). Among these options, the remission rate of TNBC after chemotherapy is higher than that after other types of treatment, making it the current preferred treatment for TNBC\(^7\). The efficacy of currently approved several chemotherapeutics, including cisplatin, anthracycline, paclitaxel, and tamoxifen, is limited owing to heterogeneity with oncogenic drivers and resistance development\(^8,9\). Accordingly, in recent years, efforts have been made to investigate the chemoresistance of TNBC to improve the chemotherapy effect and survival rate of patients with TNBC.

At present, taxane-based regimens represent the mainstay in TNBC therapy. Although paclitaxel is effective for some patients with TNBC, approximately 30–50% develop resistance, resulting in a lower overall survival rate\(^10\). The genomic and molecular basis of chemotherapy resistance in patients with TNBC remains poorly understood because the development of TNBC chemotherapy resistance is
multifaceted and based on complex interactions among the tumor microenvironment (TME), drug effector, cancer stem cells, and many tumor cells. Because drug resistance is a complex process driven by disrupted gene expression, it is critical to screen differentially expressed genes (DEGs) associated with drug resistance. Furthermore, it is of great significance to analyze the mechanism of DEGs on Taxol sensitivity for identifying new targets to reverse drug resistance and improve prognosis.

Ferroptosis, a new type of cell death discovered recently, is usually characterized by abundant iron accumulation and lipid peroxidation during cell death. Different from the necrosis modes of apoptosis and autophagy, ferroptosis has unique manifestations. In terms of morphology, mitochondrial atrophy, mitochondrial membrane shrinkage, and rupture, a normal nucleus with a lack of chromatin condensation can be observed during ferroptosis. Cellular metabolism in cells undergoing ferroptosis is characterized by the accumulation of intracellular ferrous ions and Reactive Oxygen Species (ROS), significant phospholipid peroxidation, and impaired lipid peroxide repair function. Studies have shown that regulating programmed cell death has a significant potential in the treatment of cancer drug resistance and that inducing ferroptosis can inhibit tumor cells resistant to conventional therapy and enhance the effect of immunotherapy.

Ferroptosis can induce or inhibit tumor cells with specific drugs and regulate the expression of related genes. Studies have revealed that traditional anti-breast cancer drugs can induce ferroptosis, which can subsequently induce apoptosis in breast cancer cells. The combination of the lysosomal promoter siramesine and the kinase inhibitor lapatinib was found to induce ferroptosis. This binding leads to an increase in iron levels by decreasing ferritin expression and increasing transferrin expression, leading to ferroptosis. Another study found that sulfasalazine caused iron ptosis in breast cancer cells, particularly in those with low Estrogen Receptor (ER) expression, suggesting that it could be used as a potential therapeutic agent for breast cancer. NR5A2 synergized with NCOA3 was used to prevent BET-induced ferroptosis in breast cancer cells by inducing NRF2 expression and enhancing the anticancer effect of BET inhibitors in breast cancer. Another study confirmed that ferroptosis was significantly inhibited in gastric cancer and that cisplatin and paclitaxel promoted miR-522 secretion by Cancer-Associated Fibroblasts, reduced lipid-ROS accumulation in cancer cells, and ultimately resulted in reduced chemotherapy sensitivity. However, to our knowledge, no study has evaluated the relationship between ferroptosis-related genes and paclitaxel resistance in breast cancer. It was hypothesized that ferroptosis is closely related to the occurrence and development of paclitaxel resistance and that exploring its potential mechanism is of great significance to identify therapeutic targets to overcome paclitaxel resistance in patients with breast cancer. Thus, the present study aimed to use bioinformatic methods to find new targets and biomarkers.

In this study, DEGs associated with Taxol resistance were screened by analyzing the gene chip expression data in the GSE dataset of BC cells with Taxol resistance. Moreover, cluster analysis and functional enrichment analysis were performed. The Gene Expression Omnibus (GEO) dataset related to paclitaxel resistance was used to screen drug-resistant genes and construct a prognostic model. Moreover, an
association between genes associated with paclitaxel resistance and immune infiltration and immune checkpoints in patients with TNBC was demonstrated in this study. Thus, our study may be provides potentially useful prognostic biomarker and predictor for individualized treatment in TNBC.

Materials and Methods

Data collection

The processed gene expression datasets of clinical samples collected from patients with breast cancer who did (n = 8) and did not (n = 20) achieve clinicopathologic response after paclitaxel treatment were retrieved from the comprehensive gene expression profile database GEO (https://www.ncbi.nlm.nih.gov/geo/) (ID: GSE22513). The RNA-sequencing raw data of 116 TNBC samples obtained from The Cancer Genome Atlas (https://portal.gdc.cancer.gov/) included therapeutic information, somatic mutation data, and CNV data files. These raw data were first standardized to fragments per kilobase million expression levels prior to evaluating the expression of FERGs using the limma program.

Identification of genes associated with paclitaxel resistance in patients with breast cancer

The paclitaxel resistance dataset of 20 patients with breast cancer who were resistant and 8 who were sensitive to paclitaxel therapy. The “Limma” R package was used to obtain DEGs between the paclitaxel-resistant and -nonresistant patients. The conditions for DEG screening were difference multiple (FC) ≥ 2 and P < 0.05. The “survival” analysis package was used for the univariate Cox regression analysis of 116 TNBC samples from TCGA database (P < 0.05) to obtain DEGs associated with prognosis. The data from TCGA database are freely available to the public, and this study strictly followed the access policies of the database and publication guidelines. Thus, this study did not require ethical review and approval from an ethics committee.

Weighted Gene Co-expression Network Analysis (WGCNA)

A total of 711 DEGs were evaluated to test their availability, and the R package “WGCNA” was used to construct a gene coexpression network. First, Pearson's correlation matrices and average linkage method were both performed for all pairwise genes. Then, a weighted adjacency matrix was constructed using a power function $A_{mn} = |C_{mn}|^\beta$ ($C_{mn}$ = Pearson's correlation between gene_m and gene_n; $A_{mn}$ = adjacency between gene m and gene n). $\beta$ was a soft-thresholding parameter that emphasized strong correlations between genes and penalized weak correlations. After choosing a power of 9, the adjacency was transformed into a topological overlap matrix (TOM), which could measure the network connectivity of a gene defined as the sum of its adjacency with all other genes for network generation. The corresponding dissimilarity (1-TOM) was then calculated. To classify genes with similar expression profiles into gene modules, average linkage hierarchical clustering was performed according to the TOM-based dissimilarity measure with a minimum size (gene group) of 50 for the gene dendrogram. The
sensitivity was set to 4. To further analyze the module, the dissimilarity of module eigengenes was
calculated and a cut line for a module dendrogram was selected by merging certain modules. In addition,
modules with distance < 0.25 were merged to finally produce five coexpression modules. Of note, the grey
module was considered a gene set that could not be assigned to any module.

Identification of clinically significant modules and
screening of hub genes

The coexpression module is a collection of genes with high topological overlap similarity. Genes in the
same module often have a higher degree of coexpression. In the current study, two methods were used to
identify the important modules relevant to clinical traits. The module eigengene (ME) represents the first
principal component of the module that is used to describe the expression pattern of the module in each
sample. Module membership (MM) refers to the correlation coefficient between genes and module
eigengenes that is used to describe the reliability of a gene belonging to a module. The correlation
between the modules and clinical data was evaluated to identify significant clinical modules. Based the
cutoff criteria (|MM| > 0.7), 158 genes with high connectivity in the clinical significant module were
identified as hub genes.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and
Genomes (KEGG) enrichment analyses

For gene set functional enrichment analysis, the GO annotation of genes in the R package (version 3.1.0)
was used as the background, followed by mapping the genes into the background set and using the R
package cluster Profiler (version 3.14.3) for enrichment analysis. At the same time, the KEGG rest API
(https://www.kegg.jp/kegg/rest/keggapi.html) was used to obtain the latest KEGG pathway gene
annotation as the background.

Development of prognostic models

Based on all TNBC patients, the risk model was established using the lasso Cox regression analysis of
drug resistance-related DEGs. A risk score was calculated for each patient according to the formula, and
patients were classified into high or low risk subtypes based on the median risk score. Kaplan–Meier
analysis was used to compare survival differences between the high- and low-risk subtypes. A receiver
operating characteristic (ROC) curve and the corresponding area under the curve (AUC) values over time
were used to assess the prognostic value of the associated risk models.

Analysis of drug susceptibility

The pRRophetic package was used for drug susceptibility prediction. The input of the R package
algorithm was mainly the large cell line expression profiles and the corresponding IC$_{50}$ information. Ridge
regression was used to build a model, which was then used to predict the chemotherapeutic response in
clinical samples.
Real-time quantitative polymerase chain reaction (RT-qPCR) Total RNA was extracted from samples using the Trizol reagent (R1100, Solarbio) and purified using chloroform and ethanol. A cDNA kit (RR036A, TaKaRa) with reverse transcriptase was used to synthesize cDNA as the template for qPCR analysis according to the manufacturer’s protocol. The qPCR reaction mixture included 1.6 µL of a mixture of forward and reverse primers, 10 µL of TB Green mix (RR430A, TaKaRa), 2 µL of cDNA sample, and 6.4 µL of RNase-free water. According to the manufacturer’s protocol, reactions were performed on Light Cycler 96 Real-Time System (Roche, Switzerland). Each sample was run in triplicate. The Ct values and the relative expression levels were determined using the $2^{-\Delta \Delta Ct}$ method. The β-actin expression levels were used for normalization.

Statistical analysis

All statistical analyses were performed using the R (R 4.1.0) software. Statistical analyses of RT-qPCR results was were performed using GraphPad Prism (version 8, GraphPad Software Inc., San Diego, CA, USA). All statistical p-values were two-sided, and $p < 0.05$ represented statistical significance (*$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$).

Results

Acquisition of DEGs associated with drug resistance

Figure 1 shows the flowchart of the construction and validation of data collection and analysis. Breast cancer paclitaxel resistance-related datasets (GSE22513) were retrieved from the GEO database. Limma differential analysis was performed on the dataset to identify DEGs between the drug resistance and the non-drug resistance groups. Limma differential expression analysis revealed 711 DEGs between the GSE22513 paclitaxel-resistant group and paclitaxel-sensitive group (257 upregulated and 454 downregulated genes; $P < 0.05$) (Fig. 2A). Figure 2B is the heatmap of the expression of relevant DEGs in different samples (Fig. 2B). To further elucidate the potential mechanisms of DEGs related to paclitaxel resistance, a functional enrichment analysis of these genes was performed. GO enrichment analysis showed that the DEGs related to drug resistance were mainly enriched in chemical carcinogenesis, cell differentiation, tissue differentiation, cell adhesion, metal ion homeostasis, and immune response (Fig. 2C). Moreover, KEGG enrichment analysis showed that the DEGs were closely related to cytokine receptor interaction, chemical carcinogenesis, primary immunodeficiency, platinum resistance, and ABC transporters (Fig. 2D).

Hub genes screened via WGCNA analysis

To screen genes associated with Taxol resistance, a WGCNA analysis of the differential gene expression matrices obtained in the previous step was performed. Five coexpressed modules were identified in GSE22513 (blue, brown, gray, turquoise, and yellow), of which the gray modules were considered as the collections of genes that could not be assigned to any module (Fig. 3A). As shown in the figure, the blue
module had the highest correlation with the turquoise module (Fig. 3B). The clinical information of each sample was used to calculate the correlation between gene modules and phenotypes to identify modules with resistant phenotypes. Figure 3C shows that there was a significant positive correlation between the turquoise module and the phenotype of drug resistance (Fig. 3C). Therefore, the turquoise module was selected for further analysis. To screen the hub genes, the expression correlation between the module feature vector and genes was calculated to obtain the MM value. A total of 154 genes with MM values > 0.7 for the yellow–green module were obtained as hub genes.

**DEGs related to drug resistance associated with ferroptosis**

The present study aimed to explore the interaction between ferroptosis and Taxol-resistant phenotype in breast cancer. Therefore, FERG datasets retrieved from the GeneCards website were intercrossed with hub genes identified in the turquoise module to obtain 48 genes (Fig. 4A). To investigate the impact of these genes on the prognosis of patients with TNBC, a batch survival analysis of these 48 genes was performed. According to the median value of gene expression as a grouping method, the prognostic differences of different groups were analyzed, identifying three genes with a significant effect on the prognosis of patients with TNBC patients (Fig. 4B). Among the three genes, patients with a high expression of GRIK3 exhibited a poor prognosis, suggesting that GRIK3 is a risk factor of tumor development (Fig. 4D). Patients with a low expression of IDO1 (Fig. 4E) and CLGN (Fig. 4C) exhibited poor prognosis, indicating that these two genes are protective against tumor development. GRIK3, IDO1, and CLGN expression were quantified in 10 pairs of TNBC and adjacent normal breast samples. Compared with the neighboring normal breast samples, TNBC samples had a significantly higher expression of GRIK3 and GRIK3 and significantly lower expression of IDO1 and CLGN, consistent with a previous prognostic model construction (Fig. 4F–H). These results indicated that a novel ferroptosis-related gene model can be used for prognostic prediction in TNBC.

**Prognostic model construction**

To investigate the impact of these three genes on the prognosis of patients with breast cancer, the lasso regression method was used to reduce dimension and construct a prognostic model \((-0.282) \times \text{IDO1} + (-0.1804) \times \text{CLGN} + (0.2349) \times \text{GRIK3}\) (Fig. 5A–B). In Fig. 5C, the scatter plots of the risk scores from low to high are shown from left to right and the different colors represent the heatmap of expression of the genes included in this label in different risk groups. Information of 116 patients with TNBC obtained from TCGA database was used to calculate a risk score for each TNBC sample based on the prognostic model. The average risk score of all samples was calculated, and all patients with TNBC were categorized into high- and low-risk groups according to the average value. The KM survival curve distribution of the risk model in the TNBC dataset showed that the overall survival time of the high-risk group was significantly lower than that of the low-risk group, indicating that the risk model strongly predicts the prognosis of patients (Fig. 5D). The ROC curves and AUCs plotted with regard to different time risk models revealed that higher AUC values indicated a better predictive power of the model. The 1-, 3-, and 5-year AUC values of the model were 0.84, 0.84, and 0.78, respectively, suggesting that the model has a strong
predictive ability (Fig. 5E). To evaluate the clinical usefulness of the prognostic model, decision curves were computed to calculate the net benefit. Decision curve analysis, a novel method to assess diagnostic tests and prediction models, showed that the prognostic model had a higher overall net benefit, indicating that the model was clinically useful (Fig. 5F–H).

**DEGs screened for enrichment analysis according to the risk score**

Patients with TNBC were grouped according to the risk model score, and the R package limma was used for differential analysis to identify DEGs between the high- and low-risk groups (Fig. 6A). All 1.5-fold DEGs were selected, and GO and KEGG enrichment analyses were performed for these genes. In the KEGG enrichment analysis, the top enriched entries were mainly immune-related pathways, including antigen processing and presentation, cell adhesion molecules, Th1/Th2 cell differentiation, and PD-L1 expression and PD-1 checkpoint pathway in cancer (Fig. 6B). GO enrichment analysis suggested that these DEGs were related to certain factors of the immune system, such as the immune response, defensive responses, leukocyte activation, and lymphocyte activation (Fig. 6C).

**Immunoinfiltration**

TME is a complex network composed of various cell types and factors that play important roles in the occurrence and development of tumors. Tumor-infiltrating lymphocytes (TILs) and other tumor-infiltrating immune cells (TIICs) are key to the understanding of tumor immune surveillance based on tumor immunogenicity, which refers to the density and location of TILs and TIICs in TME. Their functional programs, including the “immune score,” play important roles in the prognosis and prediction of several cancers. First, an in-depth study was conducted on the relationship between the risk score and degree of immune cell infiltration. Based on the TIMER database, the risk score was negatively correlated with the infiltration levels of B cells, CD4+ T cells, CD8+ T cells, neutrophils, and myeloid dendritic cells (Fig. 7A). All patients with TNBC were classified into two groups according to the risk model score, and different algorithms were used to compare the degree of immune cell infiltration between the two groups. Immune infiltration analysis performed on the corrected TCGA-TNBC dataset using mcpcounter showed that there were significant differences in the expression of B cells, CD8+ T cells, monocytes, myeloid dendritic cells, natural killer (NK) cells, and T cells between the two groups. Moreover, their expression was significantly lower in the high-risk group than in the low-risk group, whereas no significant difference was found between the groups with respect to endothelial cells, fibroblasts, and neutrophils (Fig. 7B). This suggests, to some extent, that there is a significant correlation between the risk model and the degree of immune cell infiltration. The epic algorithm was used to verify the above results, which revealed that the differential expression of certain immune cells, such as B cells, CD4+ T cells, CD8+ T cells, and macrophages, was statistically significant and negatively correlated with the risk score (Fig. 7C). Finally, the infiltration abundance of TIICs, fibroblasts, and epithelial cells in the two groups was assessed via the xcell algorithm, and the results were generally consistent with those of the other two algorithms (Fig. 7D).
These results indicate that this risk model can be used to predict the degree of immune cell infiltration in tumors. A higher level of immune cell infiltration represents a better prognosis.

**Immune checkpoints**

The relationship between the risk score and expression level of the common immune checkpoints was assessed. Figure 10A shows the distribution of common immune checkpoints in different risk patients. As the risk score increased (left to right on the X-axis), the survival rate of patients significantly decreased (see the middle panel) and the expression levels of common immune checkpoints showed a downward trend (Fig. 8A). Patients with TNBC were grouped according to the risk score, and differences in immune checkpoint expression between the two groups were compared. The results revealed that the expression of some common immune checkpoints, such as PDCD1, ICOS, and CTLA4, in the high-risk group was significantly lower than that in the low-risk group (Fig. 8B). Figure 8C demonstrates the relatively low expression of immune checkpoints in the high-risk group (Fig. 8C). The tumor inflammation signature (TIS) is a set of 18 genes, including the IFN-γ signaling pathway as well as T cell and NK cell abundance, that are highly associated with clinical response to immune checkpoint inhibitors. Thus, TIS can better reflect the degree of immune cell infiltration in TME. At present, TIS has been used in a number of clinical trials on tumor immunotherapy. In addition to early melanoma, head and neck cancer, and gastric cancer, TIS has begun to be used in difficulties recognized by clinical frontline workers in other cancers, such as TNBC, offering strong clinical benefits. The TIS score has been significantly associated with the survival benefit of patients treated with ICIS combined with chemotherapy. Figure 8D shows that the TIS score was significantly lower in the high-risk group than in the low-risk group (Fig. 8D). These results suggest that the risk model is significant in predicting the efficacy of immunotherapy in patients with TNBC. For patients with low score, in addition to conventional Taxol chemotherapy drugs that are prone to resistance, the effect of immunotherapy was not optimistic.

The maftools package was used to analyze any difference in the distribution of somatic mutations in low and high clinical risk group (CRG) scores in the TCGA-TNBC cohort. As shown in Figs. 8E and F, TMB was more extensive in the low score group. In theory, the higher the TMB, the more neoantigens are recognized by T cells and the better the response to immune checkpoint inhibitor treatment. The most obvious somatic mutations in the low score group were TP53 (91%) and TTN (29%), whereas those in the high score group were TP53 (85%) and PTEN (18%) (Fig. 8E–F). Evidence suggests that patients with L-TNBC may benefit from immunotherapy. Further survival analysis demonstrated that the H-TMB subgroup had significant survival benefits.

Sensitivity analysis was performed between the two groups with regard to a few medications that are currently used to treat breast cancer. The results showed that the paclitaxel IC\textsubscript{50} was significantly higher in the high score group than in the low score group, indicating the importance of the risk model in predicting paclitaxel sensitivity in patients with TNBC. In patients with high CRG scores, the IC\textsubscript{50} values of doxorubicin, tamoxifen, vinblastine, bleomycin, and AUY922, among others, were considerably higher.
This may provide new options for the treatment of patients with paclitaxel-resistant triple-negative breast cancer.

Discussion

The most common chemotherapy drug for TNBC is paclitaxel. In recent years, several studies have focused on the efficacy of paclitaxel to enhance the final therapeutic effect. However, its therapeutic effect is limited, and the most important disadvantage of paclitaxel is the acquired drug resistance. The mechanism of paclitaxel action includes not only mitosis but also angiogenesis, apoptosis, inflammation, and ROS generation. A prognostic model was established by analyzing the intersection of DEGs related to paclitaxel resistance and ferroptosis. To improve treatment outcomes in breast cancer, it is critical to first identify patients who are more susceptible to drug resistance and then find measures to reduce this risk.

The discovery of regulatory cell death has led to great progress in the field of cancer treatment. In recent years, scientists have identified ferroptosis as a new form of iron-dependent regulatory cell death caused by excessive lipid peroxidation. Ferroptosis is associated with the occurrence and treatment response of various cancer types, attracting widespread attention. Previous studies have shown that resistance to ferroptosis can affect cell proliferation, migration, and drug resistance in different cancer types. In addition, ferroptosis has been shown to play a crucial role in tumor progression and cancer therapy.

Chemotherapy resistance is a complex process involving multiple genes and signaling pathways. Tumor cells can develop drug resistance through various mechanisms, such as reducing drug intake, increasing drug pump, inhibiting apoptosis, increasing DNA repair capacity, and inhibiting apoptosis. Studies have shown that when tumor cells are exposed to chemotherapeutic drugs, these drugs can induce significant ROS production, in turn leading to tumor cell death. However, once tumor cells initiate a mechanism to change their metabolic microenvironment, inhibit ROS formation, and enhance oxidative stress defense or tolerance ability, drug resistance is induced.

From DEGs identified in two clusters, three significant genes (IDO1, CLGN, and GRIK3) were selected, which were used to construct a prognostic model through univariate and lasso Cox analysis. IDO1 is an intracellular heme-dependent oxidase that causes local tryptophan depletion and ferroptosis inhibition, and the regulation of immune cells via tryptophan metabolism regulation is associated with chemotherapy tolerance promotion and poor cancer prognosis. IDO1 is strongly expressed in tamoxifen-resistant breast cancer cells and mediates the proliferation, metastasis, and T resistance of tamoxifen-resistant breast cancer cells in vitro and in vivo via STAT1 and IL-6/STAT3 activation. GRIK3, an important excitatory neurotransmitter receptor, has been positively correlated with the prognosis of patients with breast cancer as it affects various signaling pathways and key signal transduction pathways, including two epithelial–mesenchymal transition regulators, SPDEF and CDH1. The low expression of CLGN was confirmed to be associated with tamoxifen sensitivity in premenopausal
patients with lumina A subtype breast cancer\textsuperscript{47}. However, the biological function of GRIK3 and CLGN in malignant tumors is largely unknown.

Several prognostic models based on FERGs have been reported to explore prognostic biomarkers and predict the prognosis of various cancer types\textsuperscript{48}. In one study, a 13-gene prognostic model of acute myeloid leukemia was developed. According to this model, patients were classified into high- and low-risk groups, with higher risk scores indicating shorter survival and association with tumor-associated immune abnormalities, mutational patterns and pathway dysregulation, and clinical outcomes\textsuperscript{49}. Another study used the GEO and TCGA databases to collect thyroid cancer gene expression data and clinical outcomes to evaluate the prognostic value of 75 FERG expression. Five FERG prognostic models and nomograms were developed to provide unique insights for predicting the prognosis of thyroid cancer\textsuperscript{50}. In a head and neck squamous cell carcinoma study, three stable molecular subtypes with different prognostic, mutational, and immunologic profiles were identified using consensus clustering with ferroptosis marker genes. WGCNA was then used to identify the gene modules related to molecular subtypes. After screening and lasso regression analysis, eight genes were determined to be related to prognosis. A prognostic model with score related to ferroptosis was finally constructed, which reflected the risk and positive prognostic factors of patients with head and neck squamous cell carcinoma\textsuperscript{51}. The role of ferroptosis related to drug resistance in TNBC, however, has not been fully elucidated\textsuperscript{52,53}.

It is well known that TME is composed of tumor cells and their surrounding cells, such as lymphocytes, TIICs, and tumor vasculature\textsuperscript{54}. There is strong evidence to support the hypothesis that TME is essential for tumor formation, progression, and treatment resistance\textsuperscript{55,56}. In the present study results, B cells, CD4\textsuperscript{+} T cells, CD8\textsuperscript{+} T cells, neutrophils, and myeloid dendritic cells were all significantly negatively correlated with the risk score. In the TNBC low-risk group, the expression of these TIICs was significantly higher than that in the high-risk group. In view of the in-depth research that has been conducted on breast cancer immunotherapy, the study of TME and immune cell infiltration may be helpful to discover new directions and mechanisms of breast cancer immunotherapy.

First, DEGs were identified from the GEO database to establish a FERG-based breast cancer prediction model. Because ferroptosis is different from the other accepted modes of cell death, it may offer new therapeutic possibilities for treating cancer. The model in the present study was validated by comparing the included data derived from the TCGA database with those from different database sources, which improved the effectiveness of the model. A prognostic model with score related to ferroptosis developed in this study can appropriately reflect the risk and positive prognostic factors of patients with TNBC. Thus, the model can be used to guide individualized adjuvant therapy and chemotherapy for patients with TNBC.

The current study has a few limitations. Only data from public sources were used in this study, necessitating further validation using more accurate clinical data. Because the prognostic signature was developed and validated using publicly sourced data, experimental studies and extensive prospective studies are needed to confirm these results.
In conclusion, a novel ferroptosis-related gene model can be used for prognostic prediction in TNBC. New ferroptosis-related genes might be used for TNBC targeting therapy in the future.

**Declarations**

**Data availability statement**

The datasets supporting the results and conclusions of this study were downloaded from the TCGA (https://portal.gdc.cancer.gov/) and GEO (accession no. GSE22513, http://www.ncbi.nlm.nih.gov/geo/). The data used for the prediction of potential drugs were obtained from the CellMiner and DrugBank databases.

**Sources of funding**

This work was supported by grants from The National Natural Science Foundation of China (No. 81672596 to Qichao Ni).

**Author contributions**

H.Y. designed the study, analyzed data and wrote the manuscript. HY.Q. and SC.M. collected and analyzed data. X.L., BY.L., D.Z., and YPC. collected samples and performed experiments. CY.S. and QC.N. conceptualized and supervised the study, provided funding, and edited the manuscript.

**Declaration of competing interest**

The authors have nothing to disclose.

**References**


**Figures**
Figure 1

The flowchart of construction and validation of data collection and analysis.
Figure 2

Acquisition and enrichment analysis of differentially expressed genes related to drug resistance. (A) Volcano plot shows related differentially expressed genes. (B) Heat map of differential gene expression. (C) GO enrichment analysis of differential genes related to paclitaxel resistance. (D) KEGG enrichment analysis of differential genes related to paclitaxel resistance.
Figure 3

The co-expression modules analysis. (A) Clustering dendrogram of genes, various colors represent different modules. (B) Correlation between modules. (C) Relationship between the five modules and clinical features.
Figure 4

Univariate regression was used to screen prognostic genes. (A) Venn diagram showing the intersection of differential genes related to paclitaxel resistance and ferroptosis-related genes. (B) Single factor regression analysis. (C) CLGN Kaplan-Meier survival analysis. (D) GRIK3 Kaplan-Meier survival analysis. (E) IDO1 Kaplan-Meier survival analysis. (G) RT-qPCR to analyze the expression of GRIK3, IDO1, and CLGN
in TNBC and adjacent normal breast samples (n=10). Expression levels were normalized against the geometric mean of β-actin.

Figure 5

A risk model was constructed in the TNBC cohort. (A) Cross-validation of LASSO regression parameter selection. (B) LASSO was used for regression analysis of differentially expressed genes. (C) Risk plot
distribution, survival status of patients, and heat map of expression of included genes in the whole TCGA-TNBC dataset. (D) Kaplan–Meier survival curves for the risk model based on the TCGA-TNBC dataset. (E) Receiver operating characteristic (ROC) curves for the risk model in the TNBC. (F-H) 1 year, 3 years and 5 years Decision curve analysis (DCA).

Figure 6

Potential biological pathways affected by Prognostic models. (A) The different expression genes (DEGs) between the high-risk and low-risk groups. (B) The KEGG enrichment of high and low risk groups. (C) The Gene Ontology (GO) enrichment of DEGs.
Figure 7

Correlations between the risk model and infiltration abundances of immune cells. (A) Correlations between the risk score and six types of tumor-infiltrating immune cells. (B-D) Comparison of compositional fractions of immune cells between the high-risk and low-risk groups evaluated using the mcpcounter formula (B), epic formula (C) and xcell formula (D).
Figure 8

Association of risk models with immunotherapy. (A) The relationship between different risk scores and patient follow-up time and changes in the expression of each immune checkpoint gene. (B) Immune checkpoint expression differences between high and low risk groups. (C) Correlation of risk score with Immune checkpoint expression. (D) The TIS score was higher in the low-risk score group. (E-F) The
waterfall plot of tumor somatic mutation created by groups with high-risk score(E) and low risk score(F).
(G) Differential chemotherapeutic drug responses responses in high- and low-risk patients.