Research Article

Construction and Validation of Immune-related LncRNAs Signature to Predict the Prognosis and Therapeutic Efficacy of Breast Cancer

🆔 **Shujing Wang,**[1] 🆔 **Jing Yang,**[2] 🆔 **Qin Tang,**[1,3] 🆔 **Qiang Wu**[3]

[1]Department of Pathology, The Second Affiliated Hospital of Anhui Medical University, Hefei, Anhui, China
[2]Department of Breast Surgery, Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, China
[3]Department of Pathology, School of Basic Medical Sciences, Anhui Medical University, Hefei, China

Abstract

Objectives: Breast cancer ranks first in morbidity and mortality among women worldwide. Herein, we constructed immune-related long non-coding RNAs (irlncRNAs) signature that could predict breast cancer prognosis.

Methods: The data of breast samples were downloaded from TCGA. Differentially expressed irlncRNAs (DEirlncRNAs) compared tumor with normal samples were obtained. Established a prognostic model after DEirlncRNAs were cyclically and separately paired. The model's accuracy was validated using ROC curve, survival analysis, clinicopathological features, tumor-infiltrating immune cells, immune checkpoints, and chemotherapeutic treatment. Quantitative real-time PCR was used to analyzed the risk score of breast cancer samples.

Results: Fifteen DEirlncRNAs pairs were selected to construct the prognostic model and distinguish high- or low-risk groups. Patients in high-risk group had poorer prognosis, more aggressive clinicopathologic features, lower expression of immune checkpoints, and higher drug sensitivity than those in low-risk group. For tumor-infiltrating immune cells, the high-risk group was positively related to cancer-promoting immune cells and negatively associated with anticancer immune cells. In clinical samples, risk score was positively correlated with patient age and KI67 index.

Conclusion: We constructed a prognostic model, which based on fifteen pairs of irlncRNAs, predicting both prognosis of breast cancer and efficacy of immunotherapy and chemotherapy as well.

Keywords: Breast cancer, the cancer genome atlas, immune-related gene, long non-coding RNA, prognostic

Cite This Article: Wang S, Yang J, Tang Q, Wu Q. Construction and Validation of Immune-related LncRNAs Signature to Predict the Prognosis and Therapeutic Efficacy of Breast Cancer. EJMO 2023;7(4):350–361.

According to the global report, 2.26 million new cases of breast cancer were diagnosed and 0.68 million died from the disease in 2020.[1] Breast cancer continues to threaten maternal health by ranking first in morbidity and mortality.[2] The disease is known for its complex clinical manifestations, morphological and molecular biological characteristics, and therapeutic resistance.[3] Therefore, finding a suitable detection method to predict the prognosis of breast cancer is imminent due to the disease's high mortality rate and heterogeneity.

Long non-coding RNAs (lncRNAs) are defined as RNAs longer than 200 nucleotides that do not encode proteins. LncRNAs, accounting for approximately 80% of the human transcriptome, perform many biological functions depend-

ing on their location. In the nucleus, lncRNAs perform diverse roles, including regulating gene expression in cis or trans, regulation of splicing, and nucleation of subnuclear domains. Their presence in the cytoplasm is associated with cytoplasmic functions such as miRNA sponging, interaction with signaling proteins, and modulation of translation of specific mRNAs.[4, 5]

The tumor microenvironment is a crucial variable to breast cancer progression.[6, 7] This results from the infiltration of the tumor microenvironment by several immune cells such as T cells, B cells, and lymphocytes infiltrating around the breast cancer and tumor stroma.[8] Emerging evidence has suggested that the dysregulation of these immune cells correlates to immunosuppression and progression in several malignant tumors. Hence, the utility of immunotherapy in the treatment of tumors has been very critical in recent times.[9, 10] Accordingly, some of the biomarkers associated with T cells, known as immune checkpoints like PD-1/PD-L1, have shown to be a novel strategy in breast cancer treatment.[11] To further understand the benefit and risks associated with these therapies, biomarkers that can predict the treatment response are urgently needed.

Over the past years, growing evidence shows that lncRNAs can regulate tumor immunity in immune cells and the immune microenvironment.[12] Accordingly, the relationship between lncRNAs and cancer immunity has received increasing attention. Some immune-related lncRNAs (irlncRNAs) are involved in tumor cell migration, invasion, epithelial-mesenchymal transformation, and metabolism. In breast cancer, lncRNA SNHG1 regulates the differentiation of Treg cells to promote the immune escape of breast cancer via regulating miR-448/IDO.[13] Furthermore, lncRNA HISLA from tumor-associated macrophages regulates aerobic glycolysis in breast cancer cells.[14] Moreover, lncRNA BCRT1 promotes breast cancer progression by targeting miR-1303/PTBP3 axis.[15] In addition, lncRNA GATA3-AS1 facilitates tumor progression and immune escape in triple-negative breast cancer through destabilization of GATA3 and stabilization of PD-L1.[16]

Signatures focusing on the tumor immune infiltration show promising predictive and prognostic value in the diagnosis, evaluation, and treatment of cancer. Individual genes are generally used to analyze immune-related signature to predict the prognosis of breast cancer. Here, we utilized a novel model of irlncRNAs pairing and iteration, which predicted breast cancer prognosis irrespective of the specific expression levels.

Methods

Data Source, Preprocessing and Differentially Expressed Analysis

The RNA-sequencing data consisting of 1164 female breast samples were downloaded from The Cancer Genome Atlas (TCGA) website (https://portal.gdc.cancer.gov). The corresponding clinical characteristics such as age, survival information, and clinical stage, were also downloaded from TCGA. R4.1.0 software was used to normalize, process, and analyze the data. Perl (https://www.perl.org) was used to convert the Ensembl ID of genes into a matrix of gene symbols and merge the RNA-sequencing data files into a matrix file. Gene transfer format (GTF) files were downloaded from Ensembl (http://asia.ensembl.org) to distinguish mRNAs from lncRNAs for further analysis. A list of recognized immune-related genes was downloaded from the ImmPort database (http://www.immport.org) and was used to screen irlncRNAs by a co-expression strategy. The immune-related genes with correlation coefficients of more than 0.4 and p-value less than 0.001 were considered as irlncRNAs. All methods were performed in accordance with the relevant guidelines and regulations.

To identify the differentially expressed irlncRNAs (DEirlncRNAs), we used the "limma" software package in R for differential expression analysis. The cutoff conditions were set as: |log2 fold change (log2FC)| >2.0, false discovery rate (FDR) <0.05.

Pairing DEirlncRNAs

The DEirlncRNAs were cyclically and separately paired, and a 0-or-1 matrix was constructed assuming Z is equal to lncRNA X plus lncRNA Y; Z is defined as 1 if the expression level of lncRNA X is higher than lncRNA Y, otherwise Z is defined as 0. Then, the constructed 0-or-1 matrix was further screened. When the expression quantity of lncRNA pairs was 0 or 1, there was no relationship between pairs and prognosis. An effective match was considered when the pairs of lncRNAs with 0 or 1 expression exceeded 20%.

Construction of the Prognostic Model

First, the survival information was combined with DEirlncRNAs pairs after which a univariate Cox regression analysis was conducted to select survival-related irlncRNAs pairs with p<0.01 as filter criteria. Then we divided the samples randomly and equally into a training set (n=518) and a validation set (n=517), combining the two sets was the total set. In the training set, the Lasso regression analysis was performed with cross validation to select the irlncRNAs pairs most correlated with prognosis. These final DEirln-

cRNAs pairs were used for the construction of the prognostic model. The following formula was used to calculate the risk score with the constructed prognostic model for all the cases:[17, 18] RiskScore = Exp1*Coef1 + Exp2*Coef2 + …… + Expi*Coefi (Expi represents the expression level of each irlncRNA pair, and Coefi represents the coefficient of each DEirlncRNA pair).

Application and Validation of the Prognostic Model

The receiver operating characteristic (ROC) curve was used to evaluate the predicted values of the model and the "survival ROC" package of R was used to calculate the area under the curve (AUC). The 1-, 5-, and 10-year ROC curves of the training set, the validation set, and the total set were plotted. According to the median risk score of the training set, the breast cancer patients were divided into the high- and low-risk groups within the three sets. The Kaplan-Meier log-rank analysis was used to compare the differences in survival between the two groups among the three sets using the "survival" package. The univariate and multivariate Cox regression analyses were performed to validate the relationship between the model and clinicopathological characteristics. Risk curves and point maps were used to observe the survival of patients. The Wilcoxon signed-rank test was used to analyze the differences in the risk score among groups with different clinicopathological characteristics. In addition to the above, R packages also included "survminer", "pHeatmap", "ggupbr", and "complexHeatmap".

Exploration of Tumor-Infiltrating Immune Cells on Risk Score

Some well-known methods including XCELL, TIMER, QUANTISEQ, MCPcounter, EPIC, CIBERSORT-ABS, and CIBERSORT were utilized to calculate the content of tumor-infiltrating immune cells of the cases downloaded from the TCGA website. Spearman correlation analysis was performed to evaluate the relationship between risk score and tumor-infiltrating immune cells. Wilcoxon signed-rank test was used to compare the differences in tumor-infiltrating immune cells content between high- and low-risk groups of the model. The results were shown as boxplots. These results were analyzed using the R "ggplot2" package, and p<0.05 was considered statistically significant.

Investigation of Clinical Performance on Risk Score

To investigate the clinical performance of the model, differential expression of immune checkpoints between high- and low-risk groups were compared. The results were shown in violin plots which were performed by "limma" and "ggpubr" packages of R. Furthermore, we calculated the half-maximal inhibitory concentration (IC50) of common chemotherapeutic drugs for breast cancer from the TCGA. According to the national comprehensive cancer network (NCCN), chemical drugs such as methotrexate, doxorubicin, gemcitabine, and paclitaxel are commonly used in the treatment of breast cancer. Wilcoxon signed-rank test was used to analyze the differences in the IC50 for the above-mentioned between the high-risk and low-risk groups. The results were shown as boxplots by using "pRRophetic" and "ggplot2" packages of R.

Collection of Clinical Samples

Fifteen female breast cancer patients from The First Affiliated Hospital of Anhui Medical University in the year of 2021 who had underwent modified radical mastectomy without preoperative therapy were collected in this study. Fresh breast cancer tissue samples were obtained and snap-frozen in liquid nitrogen immediately. The clinicopathological characteristics including patient age, grade, tumor size, lymph node status and the expression of ER, PR and KI67 were collected. All patients included in this study signed informed consent and this study was approved by the Institutional Review Board of The First Affiliated Hospital of Anhui Medical University, Hefei, China.

RNA Extraction and Quantitative Real-Time PCR Analysis (qRT-PCR)

Total RNA of the tissue samples was extracted using TRIzol Reagent (Thermo Fisher Scientific, China) and reverse-transcribed into complementary DNA (cDNA) using the Evo M-MLV RT Premix (Accurate Biology, China) according to the manufacturer's instructions. QRT-PCR was performed using a SYBR Green Premix Pro Taq HS qPCR Kit (Accurate Biology, China) and repeated 3 times. The detailed primer sequences of 18 lncRNAs in this study were shown in Table S1.

Statistical Analysis

For qRT-PCR, the mean cycle thresholds (CT) value was used to reflect lncRNA expression level and obtained the value of the lncRNA pairs. When assuming Z is equal to lncRNA X plus lncRNA Y; Z is defined as 1 if the expression level of lncRNA X is higher than lncRNA Y, otherwise Z is defined as 0. The risk score of each case was calculated by using the formula shown in "Construction of the prognostic model" section. All statistical analyses were performed using SPSS 22.0. Spearman's correlation coefficient was used to assess the correlation between clinicopathological parameters and risk score. Student t test was performed to compare values obtained from two groups. When p<0.05 were considered significant.

Results

Identification of DEirlncRNAs

The flow chart of the study was shown in Figure 1. Firstly, we downloaded transcriptome profiling consisting of 1053 breast cancer samples and 111 normal samples from TCGA website and then annotated according to GTF files from Ensembl. Next, we obtained the irlncRNAs by co-expression of immune-related genes from the ImmPort database. A total of 1041 irlncRNAs were identified (Table S2), and 55 were distinguished as DEirlncRNAs. Among the 55 distinguished DEirlncRNAs, 39 were upregulated while 16 were downregulated (Fig. 2a and b, Table S3).

Construction of DEirlncRNAs Pairs and Prognostic Model

According to the iteration loop and the 0 or 1 matrix, we obtained 1172 valid DEirlncRNAs pairs. Following, the DEirlncRNAs pairs were merged with the survival information. Through Cox regression analysis, we obtained 20 DEirlncRNAs pairs that were correlated with survival (Fig. 2c). Furthermore, 15 DEirlncRNAs pairs were included in the model after Lasso regression analysis (Fig. 2d and e, Table 1).

In order to verify the accuracy of the prognostic model, we drew the ROC curve and calculated AUC. We analyzed the model in the training set, and the 1-year AUC was 0.838, the 5-year AUC was 0.688, and the 10-year AUC was 0.788 (Fig. 3a). We validated the model in the validation set, and ob-
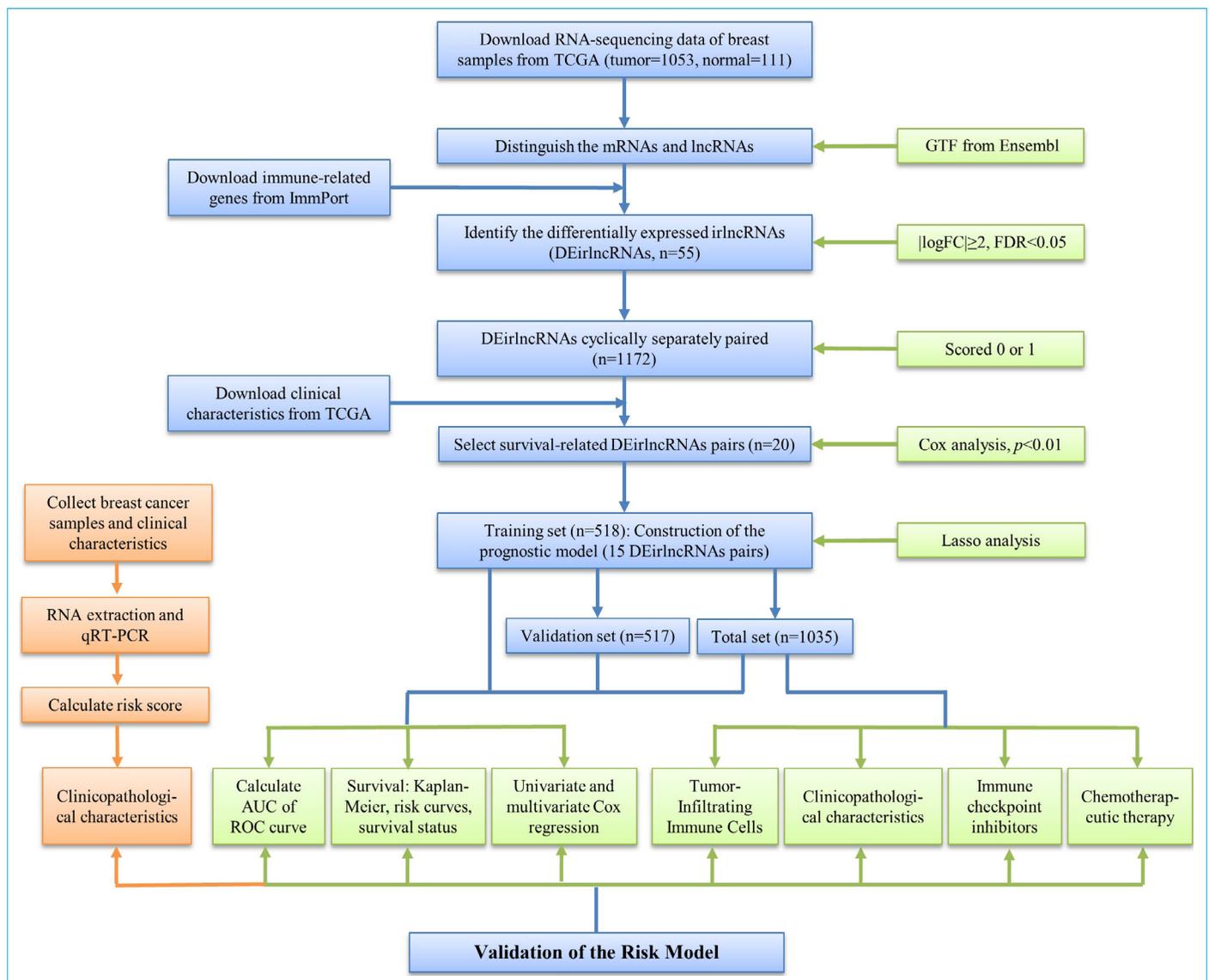

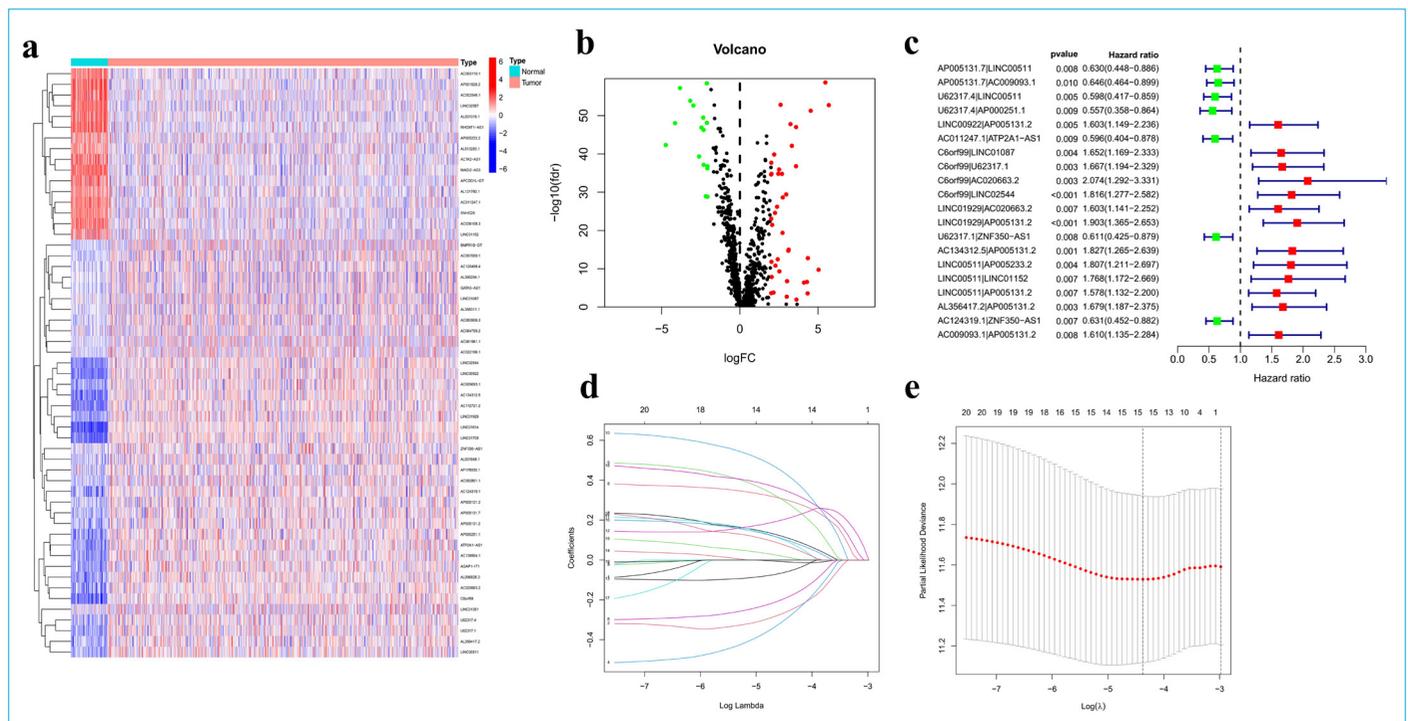**Figure 1.** The flow chart of this study.

Figure 2. Construction of a prognostic model using differentially expressed irlncRNAs (DEirlncRNAs) pairs. **(a)** The DEirlncRNAs showed by heatmap. **(b)** The DEirlncRNAs showed by volcano, red dots represented the upregulated while green dots represented the downregulated. **(c)** Forest plots showing HR of DEirlncRNAs pairs obtained by multivariate Cox regression analysis. **(d)** and **(e)** Lasso regression analysis to identify DEirlncRNAs most correlated with overall survival.

Table 1. The prognostic model of 15 differentially expressed irlncRNAs pairs for breast cancer

Pairs	Coef
AP005131.7|LINC00511	-0.01085
AP005131.7|AC009093.1	-0.25014
U62317.4|AP000251.1	-0.27626
AC011247.1|ATP2A1-AS1	-0.17886
C6orf99|LINC01087	0.104848
C6orf99|U62317.1	0.294303
C6orf99|AC020663.2	0.272957
C6orf99|LINC02544	0.411963
LINC01929|AC020663.2	0.083829
LINC01929|AP005131.2	0.213263
U62317.1|ZNF350-AS1	-0.05333
LINC00511|AP005233.2	0.010714
LINC00511|LINC01152	0.092366
AL356417.2|AP005131.2	0.337368
AC009093.1|AP005131.2	0.063

served that all the AUC values were more than 0.680 (1-year AUC = 0.760, 5-year AUC = 0.719, 10-year AUC = 0.683, Fig. 3b). Similar trend was obtained for the total set (1-year AUC = 0.797, 5-year AUC = 0.700, 10-year AUC = 0.736, Fig. 3c).

According to the median value of risk score in the training set, all the samples were divided into the high-risk group and the low-risk group. We then performed a survival analysis of the training, the validation, and the total sets using Kaplan-Meier curves. As shown in figure 3d-f, patients in the high-risk group had worse overall survival than those in the low-risk group (p<0.001).

Assessment of the Correlation Between the Prognostic Model and Clinicopathological Characteristics

The risk curves and scatter plots were used to display the risk score and the survival outcome of each breast cancer patient in the training, the validation, and the total sets. The results showed that the mortality in the low-risk group was lower than in the high-risk group (Fig. 4a-c).

To evaluate whether the prognostic model of the DEirln-cRNAs pairs was an independent prognostic factor for breast cancer, univariate and multivariate Cox regression analyses were conducted. In the training set, the hazard ratio (HR) of risk score and 95% CI were 1.733and 1.419-2.116 in univariate Cox regression analysis (p<0.001), and 1.714 and 1.338-2.197 in multivariate Cox regression analysis (p<0.001) respectively (Fig. 4d). This suggested that the prognostic model of DEirlncRNAs pairs were independent prognostic factors in patients with breast cancer. In the validation set, the prognostic model also showed statistical
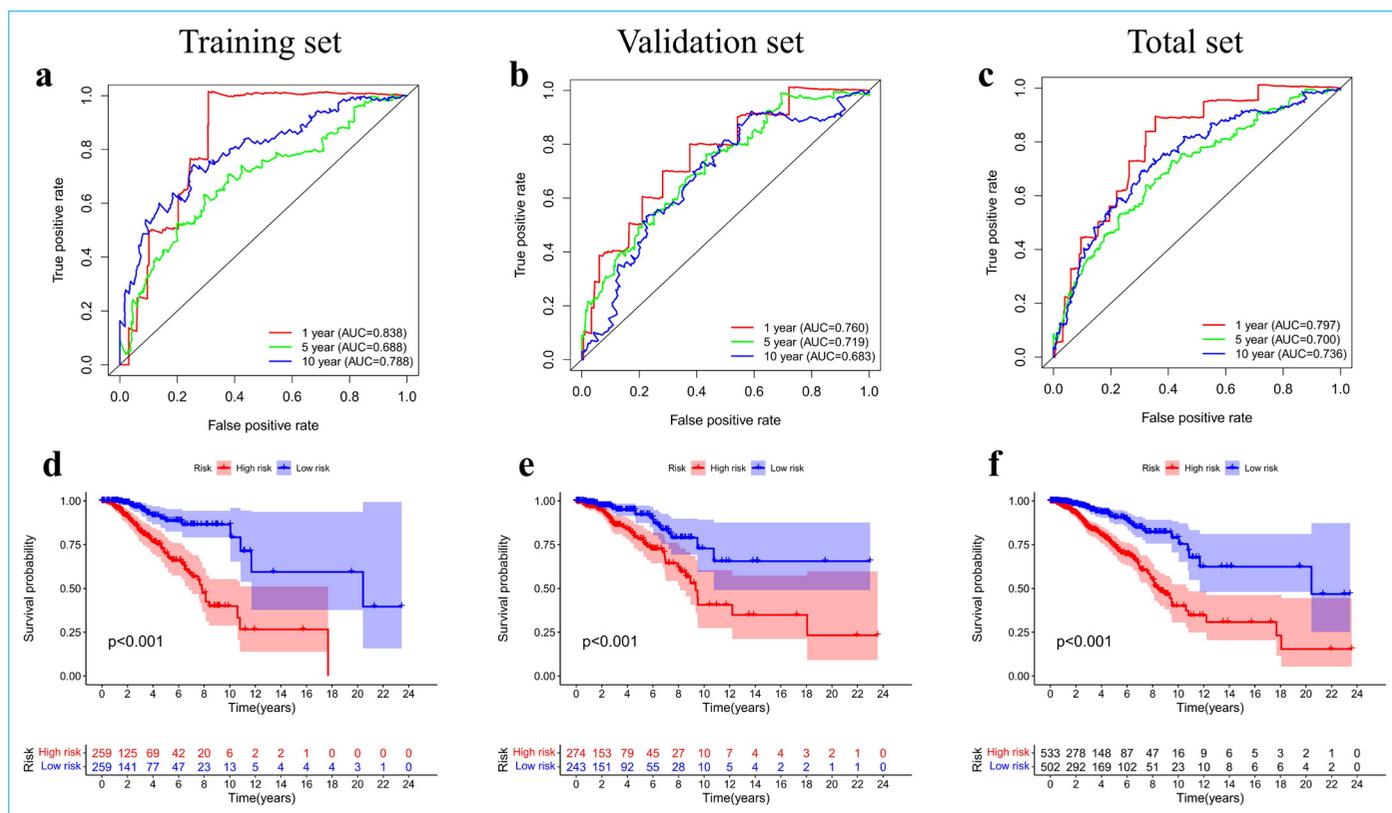
Figure 3. Validation of the prognostic model. **(a-c)** The 1-, 5-, and 10-year survival receiver operating characteristic (ROC) curves of the training, validation, and total sets. **(d-f)** Survival curves showed that patients in the high-risk group had worse overall survival than those in the low-risk group.

differences by univariate Cox regression analysis (p<0.001, HR = 1.728, 95% CI [1.334-2.239]) and multivariate Cox regression analysis (p<0.001, HR = 1.816, 95% CI [1.369-2.408], Fig. 4e). Similar results were obtained in the total set by univariate Cox regression analysis (p<0.001, HR = 1.722, 95% CI [1.471-2.016]) and multivariate Cox regression analysis (p<0.001, HR = 1.697, 95% CI [1.416 -2.035], Fig. 4f).

Furthermore, we compared the differences in clinicopathological characteristics between the low- and high-risk group of all the samples (Fig. 5a). As shown by the strip chart and scatter plots, ER, PR, age, TNM stage, T stage, M stage, and N stage were significantly related to the risk score (Fig. 5b-h).

Estimation of Tumor-Infiltrating Immune Cells and Immune Checkpoints with the Prognostic Model

Because lncRNAs play an important role in tumor immunity, we investigated the relationship between the model and the tumor immune microenvironment using all the breast cancer samples. As shown in Figure 6a and b, the immune score and stroma score were higher in the low-risk group than in the high-risk group (p<0.05). We compared various tumor-infiltrating immune cells between the high-risk and low-risk groups, the results are listed in Figure S1. In addition, we integrated the relationships between various immune cells within the risk group performed by Spearman

correlation analysis into a bubble chart. The results showed that the risk score was more positively associated with M2 macrophages, cancer-associated fibroblasts and neutrophils, whereas they were negatively associated with T cells, CD8+ T cells, NK cells, B cells, and M1 macrophages (Fig. 6c, Table S4).

Immunotherapy is a novel and effective treatment method for breast cancer. We analyzed differences in immune checkpoints common in breast cancer between high- and low-risk groups. The expression of PCDC1 (PD-1), CD274 (PD-L1), CTLA4, and CDK4 were significantly lower in the high-risk group (Fig. 7a-d).

Analysis of the Correlation Between the Prognostic Model and Chemotherapeutic Drugs

The prediction of how effective a chemotherapy drug can guide the selection of clinical drugs. We assessed the correlation between the prognostic model and the efficacy of common chemotherapeutics used in the treatment of breast cancer. A higher IC50 of chemotherapeutic agents, such as methotrexate, doxorubicin, and gemcitabine were associated with the high-risk group (p<0.05, Fig. 7e-g). While paclitaxel had no significance with the model (p>0.05, Fig. 7h). The data indicated that the model might predict the treatment response to chemotherapy agents.
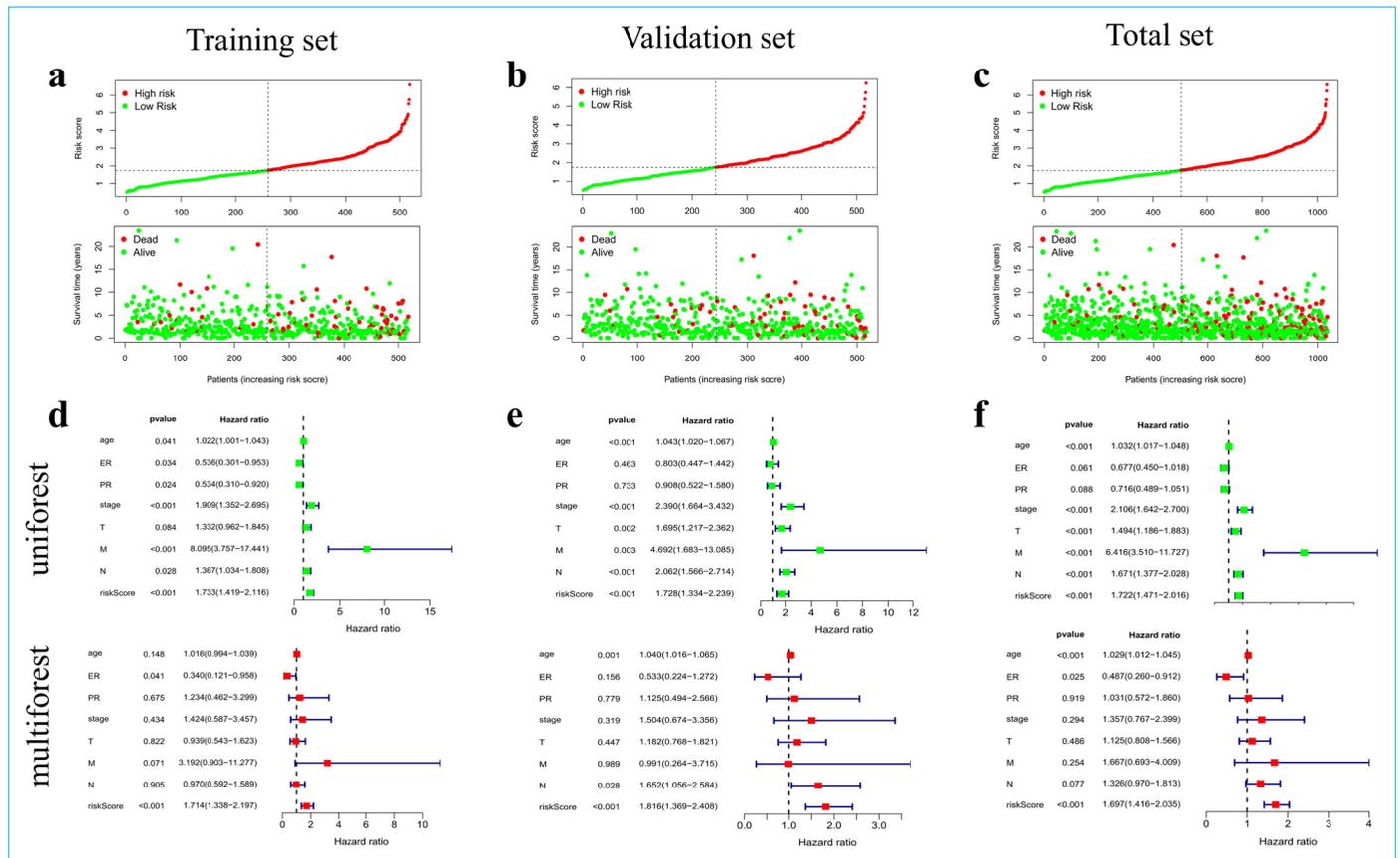
Figure 4. The survival outcome of breast cancer patient. Risk curves and scatter plot showed the survival outcome of each breast cancer patient in the training **(a)**, validation **(b)** and total sets **(c)**. The green and red dots respectively represent survival and death. The univariate and multivariate Cox regression analyses for evaluating the independent prognostic value of the prognostic model and clinicopathologic parameters in the training **(d)**, validation **(e)** and total sets **(f)**. P<0.05 was considered statistically significant.
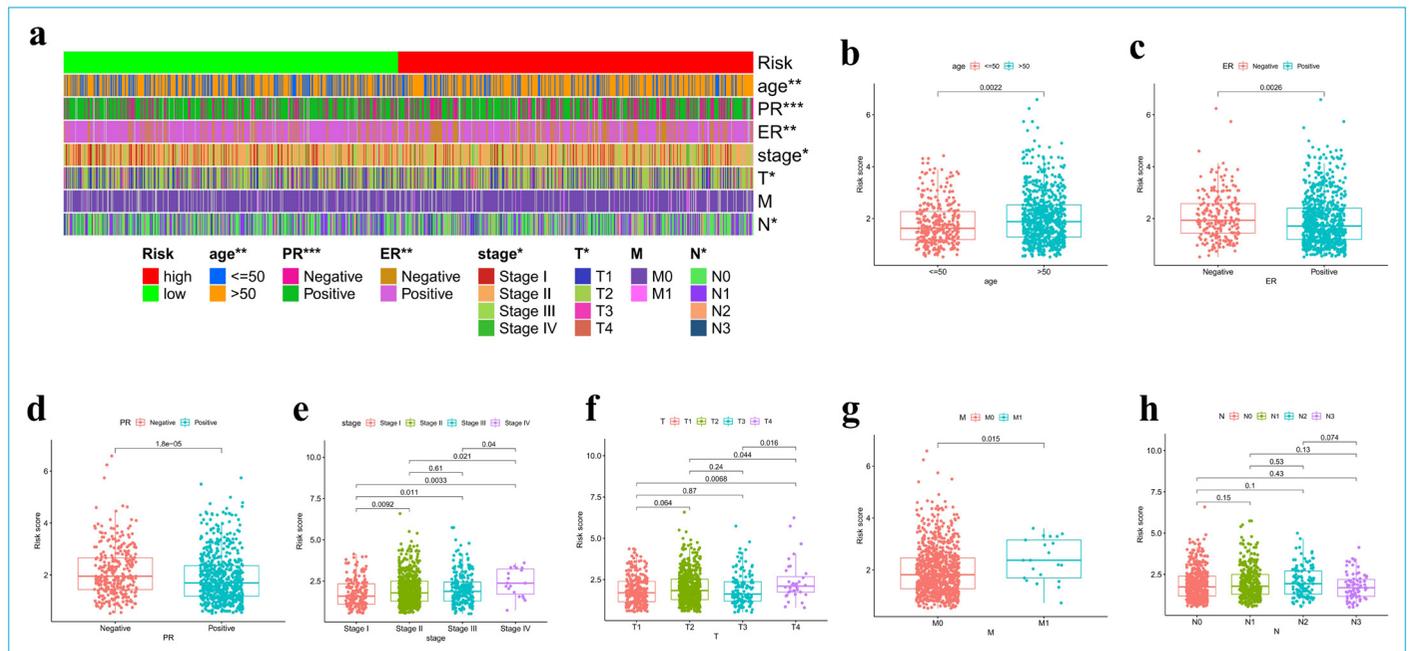

Figure 5. Evaluation of clinicopathological characteristics by the prognostic model. **(a)** The strip chart showed the differences in clinicopathological characteristics between high- and low-risk groups. **(b-h)** Scatter plots showed the differences in risk score among different clinical feature groupings. *p<0.05, **p<0.01, ***p<0.001.
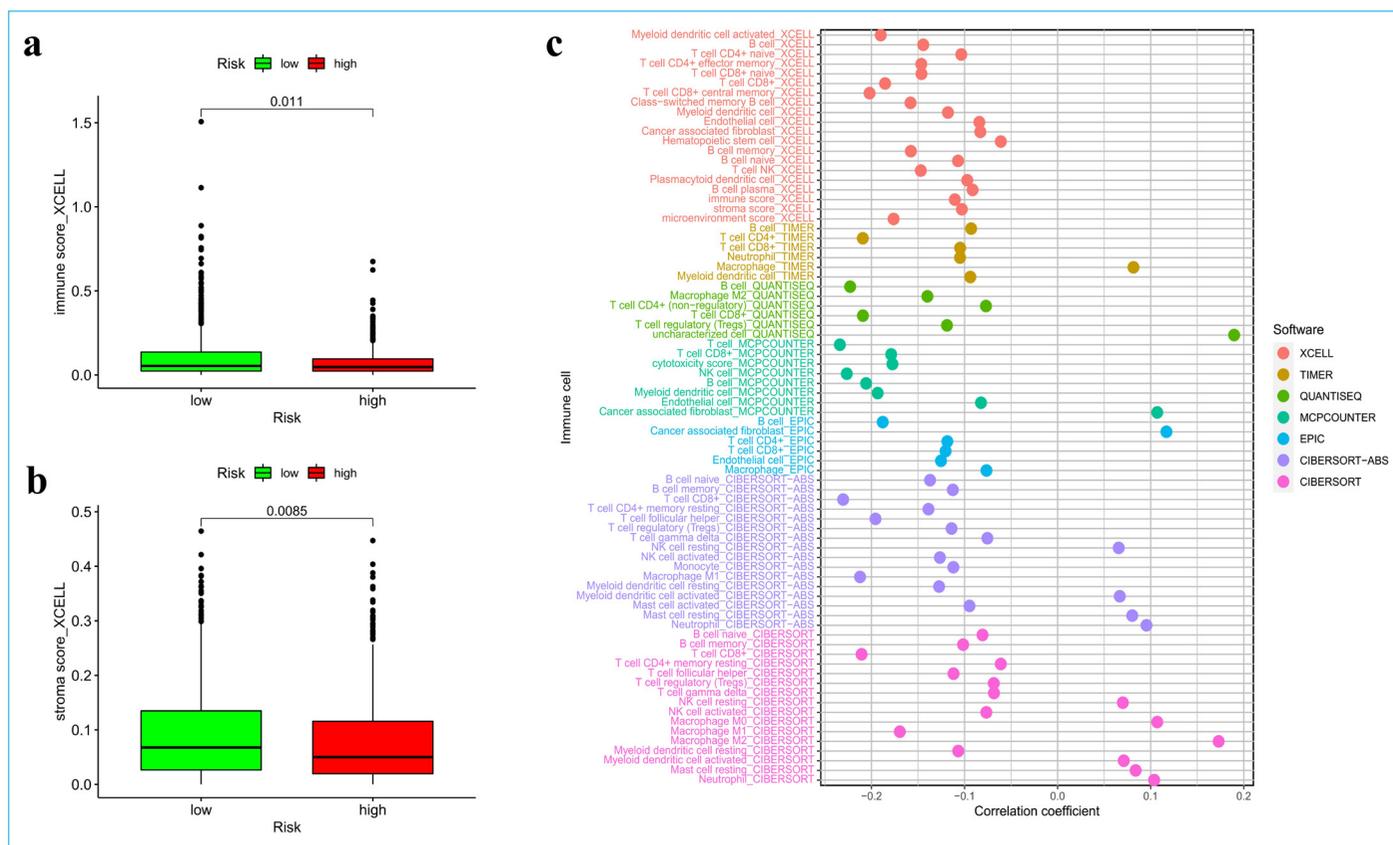
Figure 6. Estimation of tumor-infiltrating immune cells by the prognostic model. Patients in the high-risk group were more negatively associated with tumor-infiltrating immune score **(a)** and stroma score **(b)**. The bubble chart **(c)** showed the detailed correlation between different tumor-infiltrating immune cells.

Validation of the Prognostic Model Using Clinical Samples by qRT-PCR

To further determined the accuracy of this model, we collected 15 breast cancer fresh samples and carried out qRT-PCR experiments, finally obtained the risk score for each patient based on the coefficients of DEirlncRNAs pairs (Table 2). We grouped patient age, grade, T stage, N stage, ER/PR and KI67 index, then compared the differences in risk score between different groups. There was no statistical significance between the different groups (Fig. 8a-f). Afterward, we analyzed the correlation between risk score with either age or KI67 index, the results showed both statistical significance (p<0.05, Fig. 8g and h). The results were almost in line with our previous data.

Discussion

Breast cancer is the most commonly diagnosed cancer among women worldwide. Although the 5-year survival rate of stage II breast cancer is about 93% and 72% for stage III,[3] breast cancer is still the leading cause of cancer death. [1] Therefore, besides the traditional clinical risk factors, additional biomarkers to predict the prognosis and treatment

of breast cancer are needed as well.

In recent years, the utility of immunotherapy for the treatment of breast cancer patients has widely gained attention. Although previous studies have used immune-related lncRNAs to predict breast cancer prognosis, the signature in those studies was constructed using a single gene.[19, 20] Herein, our study presents the first report using a differentially expressed irlncRNAs pairs model to predict the treatment and prognosis of breast cancer. The two-lncRNA pairs were superior to a single gene, that it is not dependent on the expression levels of each gene, but the relative expression of the two genes. For different detection systems, data correction can be done without.

To construct the model, we downloaded breast cancer transcription data from TCGA database and obtained DEirlncRNAs. The DEirlncRNAs pairs were identified using an improved method of cyclically and singly pairing along with a 0 or 1 matrix. We further performed univariate regression analysis to select survival-related DEirlncRNAs pairs. To better verify the accuracy of the model, we randomly and equally divided the samples into the training set and the validation set. The Lasso regression analysis was used to screen the DEirlncRNAs pairs for the prog-
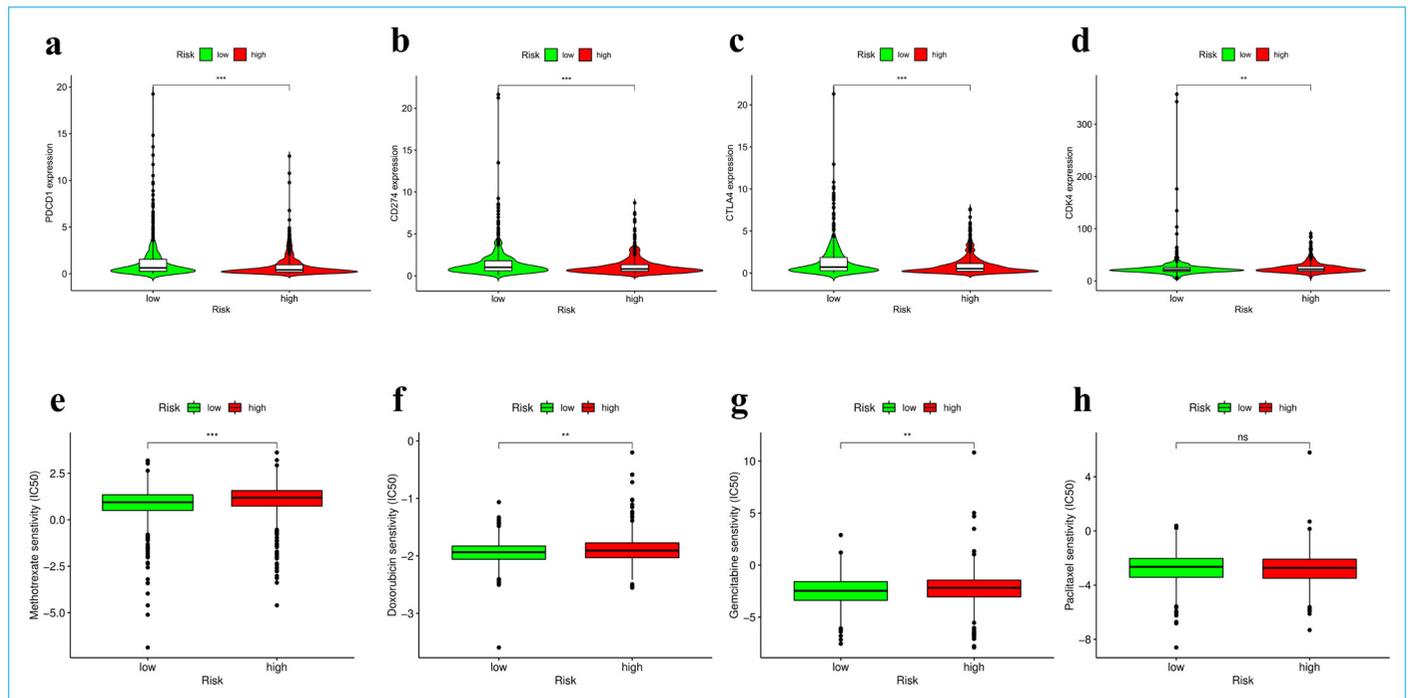
Figure 7. Verification of the correlation between the prognostic model and therapeutic targets in breast cancer. Violin plots showed that the expression of immune checkpoints biomarkers including PCDC1 **(a)**, CD274 **(b)**, CTLA4 **(c)**, and CDK4 **(d)** were significantly lower in the high-risk group. Box plots showed that chemotherapeutic agents, such as methotrexate **(e)**, doxorubicin **(f)** and gemcitabine **(g)** were associated with the high-risk group while paclitaxel **(h)** showed no significant difference between the two groups. *p<0.05, **p<0.01, ***p<0.001, ns represents no significance.

Table 2. Risk score and clinicopathological characteristics of breast cancer patients

Samples	Risk score	Age(y)*	Grade	T	N	ER	PR	KI67(%)#
No.1	-0.40194	43	2	1	-	+	+	10
No.2	-0.40194	47	2	2	-	+	+	25
No.3	-0.12568	42	3	1	\	+	+	12
No.4	-0.12568	48	2	2	+	+	+	40
No.5	-0.10764	67	2	2	-	+	+	35
No.6	-0.06502	56	2	2	+	+	+	30
No.7	-0.05431	60	3	2	-	+	+	30
No.8	0.060509	58	2	1	+	+	+	30
No.9	0.12446	57	2	1	-	+	+	20
No.10	0.152875	46	2	1	+	+	+	55
No.11	0.158956	71	3	1	+	-	-	40
No.12	0.211239	47	2	2	-	-	-	75
No.13	0.348666	57	3	2	+	+	+	40
No.14	0.348666	59	2	2	+	+	+	15
No.15	0.365837	69	2	2	-	-	-	50

*: Age was correlated to risk score (R = 0.517, p<0.05). #: KI67 was correlated to risk score (R = 0.532, p<0.05). -: negative; +: positive.

nostic model construction in the training set. According to the median value of risk score in the training set, all the patients, whether in the training set or the validation set, were divided into high-risk and low-risk groups. Next, we calculated the AUC value from the ROC curve to validate the prognostic model. Finally, we evaluated the accuracy of the model by analyzing the differences in various clinical factors including survival, clinicopathological features, tumor-infiltrating immune cells, and immune checkpoints. In our signature, 15 DEirlncRNAs pairs consisting of 18
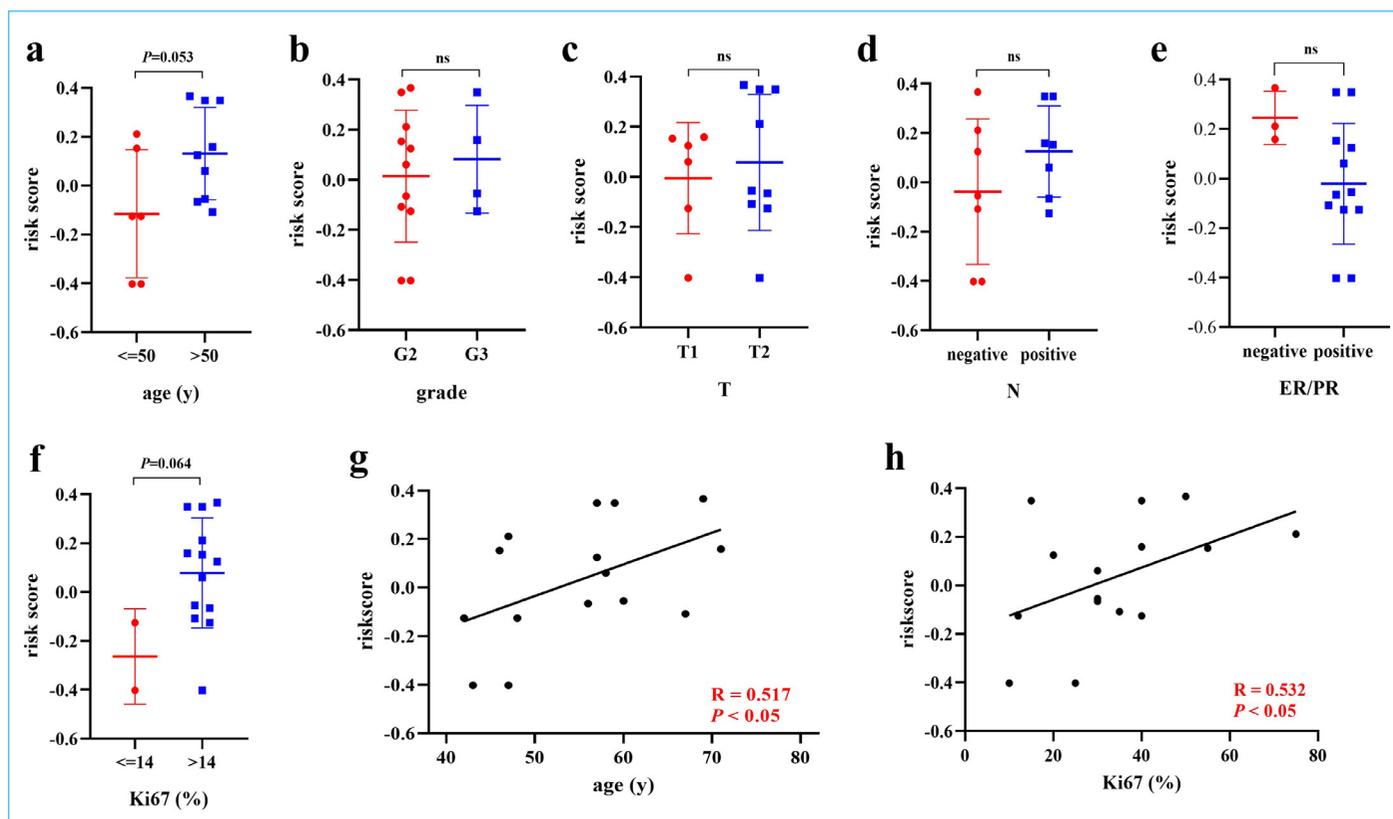
Figure 8. Validation of the prognostic model using breast cancer samples. Risk score in different groups of patient age **(a)**, grade **(b)**, T stage **(c)**, N stage **(d)**, ER/PR expression **(e)** and KI67 index **(f)** had no statistical significance. Risk score was correlated to patient age **(g)** and KI67 index **(h)**.

DEirlncRNAs were used to construct the model. Some of the DEirlncRNAs identified in our study have been reported to play an important role in malignant tumors. For example, LINC00511 may contribute to breast cancer tumorigenesis, proliferation, migration and invasion, and stemness.[21, 22] The tumor-promoting functions of LINC00511 have also been reported in gastric cancer,[23, 24] hepatocellular carcinoma,[25, 26] colorectal cancer[27, 28] and bladder cancer.[29, 30] LINC01087 could represent a novel, specific and promising biomarker not only for the diagnosis and prognosis of luminal subtypes and triple-negative breast cancers but also as a predictive biomarker of pharmacological interventions.[31] It has also been reported that the overexpression of LINC01087 in breast cancer can promote the invasion and migration of breast cancer cells.[32] LINC02544 may promote proliferation, invasion, and migration of breast cancer cells after neoadjuvant chemotherapy.[33] LINC01152 could induce tumorigenesis in glioblastoma via the Notch signaling pathway[34] and promotes cell proliferation and survival in hepatocellular carcinoma.[35] Amelia et al. found that there were many lncRNAs dysregulation in non-small cell lung carcinoma, in which LINC01929 was upregulated. [36] LINC01929 functioned as a tumor-promoting lncRNA in oral squamous cell carcinoma via the miR-137-3p/FOXC1

axis.[37] In addition, some DEirlncRNAs such as AL356417.2, AP005233.2, ATP2A1-AS1, C6orf99, U62317.1, and U62317.4 appear only in bioassay and have not been experimentally confirmed. Other DEirlncRNAs, AC009093.1, AC011247.1, AC020663.2, AP000251.1, AP005131.2, AP005131.7, and ZNF350-AS1 were revealed for the first time.

Over the years, the role of immune cells in breast cancer has been increasingly discovered. To investigate the relationship between risk score and tumor-infiltrating immune cells, we used seven common acceptable methods including XCELL, TIMER, QUANTISEQ, MCPcounter, EPIC, CIBERSORT-ABS, and CIBERSORT to estimate the infiltration of immune cells in breast cancer samples. Due to the differences and complexity among the various algorithms, the results were not compared with each other. Our results showed that the risk score was more positively related to cancer-promoting immune cells like M2 macrophages and cancer-associated fibroblast, while negatively related to anti-cancer immune cells like T cells, CD8+ T cells, NK cells, B cells, and M1 macrophages. In addition to immune cells, we also analyzed the relationship between risk score and immune checkpoints common to breast cancer. Immune checkpoints, especially those associated with T cells, have

been used in clinical treatment. The first immunotherapies to the immunomodulatory receptor CTLA4 and blockade of the immunoinhibitory receptor PD-1 in cancer immunotherapy have created a paradigm of cancer therapy.[11, 38] Based on our results, PCDC1 (PD-1), CD274 (PD-L1), and CTLA4 were highly expressed in the low-risk group. Our findings suggested that the prognostic model can predict the efficacy of immunotherapy in clinical settings.

Chemotherapy is a common treatment for breast cancer and hence we analyzed the relationship between risk score and IC50 of 4 common chemotherapy drugs used in breast cancer treatment. It was observed that methotrexate, doxorubicin, and gemcitabine were associated with the high-risk group. This finding showed that the prognostic model could indicate the chemotherapeutics sensitivity.

In addition, we collected breast cancer samples and detected the relative expression levels of 18 lncRNAs by qRT-PCR. Based on the contrasting CT values, 1 or 0 was obtained for each irlncRNA pair, then 1 or 0 was multiplied by the coefficients to obtain the risk score for each patient. The risk score were then used for subsequent analysis. According to the model results, the risk score was statistically significant with patient age, TNM stage, T stage , M stage, N stage, ER and PR expression. Although in our samples, the risk values were not statistically significant in the age, T stage, N stage, ER, PR and KI67 index subgroups, the trend was the same as in the model. Small sample covering only 15 cases may be the reason. Additionally, we analyzed the correlation of risk values with either patient age or KI67 index, and the results showed statistically significant. These results indicated that the established risk-predicting model may have potential value for predicting patient outcomes of breast cancer. Further investigation to validate the model needs to be carried with large sample set and patient follow up.

Conclusion

We constructed a prognostic model, which based on fifteen pairs of irlncRNAs, predicting both prognosis of breast cancer and efficacy of immunotherapy and chemotherapy as well.

Disclosures

Ethics Committee Approval: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of The First Affiliated Hospital of Anhui Medical University, Hefei, China (No. 5101389).

Peer-review: Externally peer-reviewed.

Conflict of Interest: None declared.

Authorship Contributions: Concept – S.-J.W.; Design – Q.W.; Supervision– Q.W.; Materials – S.-J.W., J.Y.; Data collection &/or processing – J.Y., Q.T.; Analysis and/or interpretation – S.-J.W., Z.-Z.F.; Literature search – S.-J.W., Z.-Z.F.; Writing – S.-J.W.; Critical review – Q.W.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–49.

2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin 2022;72:7–33.

3. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. Nat Rev Dis Primers 2019;5:66.

4. Chachoua I, Tzelepis I, Dai H, Lim J, Lewandowska-Ronnegren A, Casagrande F, et al. Canonical WNT signaling-dependent gating of MYC requires a noncanonical CTCF function at a distal binding site. Nat Commun 2022;13:204.

5. Liu Y, Shi M, He X, Cao Y, Liu P, Li F, et al. LncRNA-PACERR induces pro-tumour macrophages via interacting with miR-671-3p and m6A-reader IGF2BP2 in pancreatic ductal adenocarcinoma. J Hematol Oncol 2022;15:52.

6. Zheng H, Siddharth S, Parida S, Wu X, Sharma D. Tumor microenvironment: Key players in triple negative breast cancer immunomodulation. Cancers 2021;13:3357.

7. Eftekhari R, Esmaeili R, Mirzaei R, Bidad K, de Lima S, Ajami M, et al. Study of the tumor microenvironment during breast cancer progression. Cancer Cell Int 2017;17:123.

8. Goff SL, Danforth DN. The role of immune cells in breast tissue and immunotherapy for the treatment of breast cancer. Clin Breast Cancer 2021;21:e63–73.

9. Giles J, Manne S, Freilich E, Oldridge D, Baxter A, George S, et al. Human epigenetic and transcriptional T cell differentiation atlas for identifying functional T cell-specific enhancers. Immunity 2022;55:557–74.

10. Evgin L, Kottke T, Tonne J, Thompson J, Huff A, van Vloten J, et al. Oncolytic virus-mediated expansion of dual-specific CAR T cells improves efficacy against solid tumors in mice. Sci Transl Med 2022;14:eabn2231.

11. Pilipow K, Darwich A, Losurdo A. T-cell-based breast cancer immunotherapy. Semin Cancer Biol 2021;72:90–101.

12. Denaro N, Merlano MC, Lo Nigro C. Long noncoding RNAs as regulators of cancer immunity. Mol Oncol 2019;13:61–73.

13. Pei X, Wang X, Li H. LncRNA SNHG1 regulates the differentiation of Treg cells and affects the immune escape of breast cancer via regulating miR-448/IDO. Int J Biol Macromol 2018;118:24–30.

14. Chen F, Chen J, Yang L, Liu J, Zhang X, Zhang Y, et al. Extracellular vesicle-packaged HIF-1alpha-stabilizing lncRNA from

tumour-associated macrophages regulates aerobic glycolysis of breast cancer cells. Nat Cell Biol 2019;21:498–510.

15. Liang Y, Song X, Li Y, Chen B, Zhao W, Wang L, et al. LncRNA BCRT1 promotes breast cancer progression by targeting miR-1303/PTBP3 axis. Mol Cancer 2020;19:85.

16. Zhang M, Wang N, Song P, Fu Y, Ren Y, Li Z, et al. LncRNA GA-TA3-AS1 facilitates tumour progression and immune escape in triple-negative breast cancer through destabilization of GATA3 but stabilization of PD-L1. Cell Prolif 2020;53:e12855.

17. Hong W, Liang L, Gu Y, Qi Z, Qiu H, Yang X, et al. Immune-related lncRNA to construct novel signature and predict the immune landscape of human hepatocellular carcinoma. Mol Ther Nucleic Acids 2020;22:937–47.

18. Ping S, Wang S, He J, Chen J. Identification and validation of immune-related lncRNA signature as a prognostic model for skin cutaneous melanoma. Pharmgenomics Pers Med 2021;14:667–81.

19. Shen Y, Peng X, Shen C. Identification and validation of immune-related lncRNA prognostic signature for breast cancer. Genomics 2020;112:2640–6.

20. Zhang D, Zheng Y, Yang S, Li Y, Wang M, Yao J, et al. Identification of a novel glycolysis-related gene signature for predicting breast cancer survival. Front Oncol 2021;10:596087.

21. Lu G, Li Y, Ma Y, Lu J, Chen Y, Jiang Q, et al. Long noncoding RNA LINC00511 contributes to breast cancer tumourigenesis and stemness by inducing the miR-185-3p/E2F1/Nanog axis. J Exp Clin Cancer Res 2018;37:289.

22. Shi G, Cheng Y, Zhang Y, Guo R, Li S, Hong X. Long non-coding RNA LINC00511/miR-150/MMP13 axis promotes breast cancer proliferation, migration and invasion. Biochim Biophys Acta Mol Basis Dis 2021;1867:165957.

23. Chen Z, Wu H, Zhang Z, Li G, Liu B. LINC00511 accelerated the process of gastric cancer by targeting miR-625-5p/NFIX axis. Cancer Cell Int 2019;19:351.

24. Sun CB, Wang HY, Han XQ, Liu YN, Wang MC, Zhang HX, et al. LINC00511 promotes gastric cancer cell growth by acting as a ceRNA. World J Gastrointest Oncol 2020;12:394–404.

25. Peng X, Li X, Yang S, Huang M, Wei S, Ma Y, et al. LINC00511 drives invasive behavior in hepatocellular carcinoma by regulating exosome secretion and invadopodia formation. J Exp Clin Cancer Res 2021;40:183.

26. Hu WY, Wei HY, Li KM, Wang RB, Xu XQ, Feng R. LINC00511 as a ceRNA promotes cell malignant behaviors and correlates with prognosis of hepatocellular carcinoma patients by modulating miR-195/EYA1 axis. Biomed Pharmacother 2020;121:109642.

27. Hu Y, Zhang Y, Ding M, Xu R. LncRNA LINC00511 acts as an oncogene in colorectal cancer via sponging miR-29c-3p to upregulate NFIA. OncoTargets Ther 2021;13:13413–24.

28. Sun S, Xia C, Xu Y. HIF-1α induced lncRNA LINC00511 accelerates the colorectal cancer proliferation through positive feedback loop. Biomed Pharmacother 2020;125:110014.

29. Dong LM, Zhang XL, Mao MH, Li YP, Zhang XY, Xue DW, et al. LINC00511/miRNA-143-3p modulates apoptosis and malignant phenotype of bladder carcinoma cells via PCMT1. Front Cell Dev Biol 2021;9:650999.

30. Li J, Li Y, Meng F, Fu L, Kong C. Knockdown of long non-coding RNA linc00511 suppresses proliferation and promotes apoptosis of bladder cancer cells via suppressing Wnt/beta-catenin signaling pathway. Biosci Rep 2018;38:BSR20171701.

31. De Palma FDE, Del Monaco V, Pol JG, Kremer M, D'Argenio V, Stoll G, et al. The abundance of the long intergenic non-coding RNA 01087 differentiates between luminal and triple-negative breast cancers and predicts patient outcome. Pharmacol Res 2020;161:105249.

32. She JK, Fu DN, Zhen D, Gong GH, Zhang B. LINC01087 is highly expressed in breast cancer and regulates the malignant behavior of cancer cells through miR-335-5p/Rock1. Onco Targets Ther 2020;13:9771–83.

33. Guo Z, Yao L, Guo A. Clinical and biological impact of LINC02544 expression in breast cancer after neoadjuvant chemotherapy. Eur Rev Med Pharmacol Sci 2020;24:10573–85.

34. Wu J, Wang N, Yang Y, Jiang G, Mu Q, Zhan H, et al. LINC01152 upregulates MAML2 expression to modulate the progression of glioblastoma multiforme via Notch signaling pathway. Cell Death Dis 2021;12:115.

35. Chen T, Pei J, Wang J, Luo R, Liu L, Wang L, et al. HBx-related long non-coding RNA 01152 promotes cell proliferation and survival by IL-23 in hepatocellular carcinoma. Biomed Pharmacother 2019;115:108877.

36. Acha-Sagredo A, Uko B, Pantazi P, Bediaga NG, Moschandrea C, Rainbow L, et al. Long non-coding RNA dysregulation is a frequent event in non-small cell lung carcinoma pathogenesis. Br J Cancer 2020;122:1050–8.

37. Che H, Che Y, Zhang Z, Lu Q. Long non-coding RNA LINC01929 accelerates progression of oral squamous cell carcinoma by targeting the miR-137-3p/FOXC1 Axis. Front Oncol 2021;11:657876.

38. Pardoll DM. Immunology beats cancer: A blueprint for successful translation. Nat Immunol 2012;13:1129–32.