



# A Novel Framework to Predict Breast Cancer Prognosis Using Immune-Associated LncRNAs

Zhijian Huang<sup>1,2†</sup>, Chen Xiao<sup>3†</sup>, Fushou Zhang<sup>4</sup>, Zhifeng Zhou<sup>5</sup>, Liang Yu<sup>6</sup>, Changsheng Ye<sup>2\*</sup>, Weiwei Huang<sup>7\*</sup> and Nani Li<sup>7\*</sup>

<sup>1</sup> Department of Breast Surgical Oncology, Fujian Medical University Cancer Hospital, Fujian Cancer Hospital, Fuzhou, China, <sup>2</sup> Breast Center, Nanfang Hospital, Southern Medical University, Guangzhou, China, <sup>3</sup> Department of Gastroenterology, Fuzhou Second Hospital Affiliated to Xiamen University, Fuzhou, China, <sup>4</sup> Department of General Surgery, The Hospital of Changle District, Fuzhou, China, <sup>5</sup> Laboratory of Immuno-Oncology, Fujian Medical University Cancer Hospital, Fujian Cancer Hospital, Fuzhou, China, <sup>6</sup> Department of Thyroid and Breast Surgery, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, <sup>7</sup> Department of Medical Oncology, Fujian Medical University Cancer Hospital, Fujian Cancer Hospital, Fuzhou, China

## OPEN ACCESS

### Edited by:

Lihong Peng,  
Hunan University of Technology,  
China

### Reviewed by:

Qiqi Xie,  
Lanzhou University, China  
Jialiang Yang,  
Genesis (Beijing) Co., Ltd., China

### \*Correspondence:

Changsheng Ye  
yechsh2006@126.com  
Weiwei Huang  
huangstudenth@163.com  
Nani Li  
linanifujian@126.com

†These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
RNA,  
a section of the journal  
Frontiers in Genetics

Received: 27 November 2020

Accepted: 31 December 2020

Published: 21 January 2021

### Citation:

Huang Z, Xiao C, Zhang F,  
Zhou Z, Yu L, Ye C, Huang W and  
Li N (2021) A Novel Framework  
to Predict Breast Cancer Prognosis  
Using Immune-Associated LncRNAs.  
Front. Genet. 11:634195.  
doi: 10.3389/fgene.2020.634195

**Background:** Breast cancer (BC) is one of the most frequently diagnosed malignancies among females. As a huge heterogeneity of malignant tumor, it is important to seek reliable molecular biomarkers to carry out the stratification for patients with BC. We surveyed immune-associated lncRNAs that may be used as potential therapeutic targets in BC.

**Methods:** LncRNA expression data and clinical information of BC patients were downloaded from the TCGA database for a comprehensive analysis of candidate genes. A model consisting of immune-related lncRNAs enriched in BC cancerous tissues was established using the univariate Cox regression analysis and the iterative Lasso Cox regression analysis. The prognostic performance of this model was validated in two independent cohorts (GSE21653 and BC-KR), and compared with known prognostic biomarkers. A nomogram that integrated the immune-related lncRNA signature and clinicopathological factors was constructed to accurately assess the prognostic value of this signature. The correlation between the signature and immune cell infiltration in BC was also analyzed.

**Results:** The Kaplan-Meier analysis showed that the OS of Patients in the low-risk group had significantly better survival than those in the high-risk group, Clinical subgroup analysis showed that the predictive ability was independent of clinicopathological factors. Univariate/multivariate Cox regression analysis showed immune lncRNA signature is an important prognostic factor and an independent prognostic marker. In addition, GSEA and GSVA analysis as well as comprehensive analysis of immune cells showed that the signature was significantly correlated with the infiltration of immune cells.

**Conclusion:** We successfully constructed an immune-associated lncRNA signature that can accurately predict BC prognosis.

**Keywords:** TCGA, GEO, immune, breast cancer, risk score, framework, prognosis

## INTRODUCTION

Globally, Cancer Statistics 2018 estimates that more than 2.1 million females are diagnosed with breast cancer (BC) every year and about 62,000 deaths occur, making it the most common female malignancy and the leading cause of cancer mortality in women. Developing countries account for nearly 60% of BC mortality (Bray et al., 2018). Although the progress of early screening and recent advances in anti-cancer strategies offer improvements in the main outcomes of BC patients (Kim et al., 2018; Siegel et al., 2020), the recurrence rate of BC remains high (Li et al., 2019). Several studies have identified a large amount of poor-prognosis related biomarkers for BC, including age, tumor size, histological grade, lymphatic vessel invasion (LVI), number of metastatic lymph nodes, hormone receptor status, c-erbB2 status, and positive margin status. However, because of the complexity of BC incidence and the heterogeneity of tumors, pre-existing prognostic markers are underpowered to increase the prediction efficiency. The emerging technique, an optimal model consisting of relevant lncRNAs, will renew hope for pronounced improvements in predicting the BC prognosis.

lncRNAs, located in the nucleus or cytoplasm, are a group of RNA molecules of greater than 200 nt in length. The majority of them have no protein-coding function, and a few can encode a limited number of polypeptides. lncRNAs are engaged in precise regulations at pre-transcription and transcription levels and a broad spectrum of biological processes, such as tumor invasion, metastasis, apoptosis, and drug resistance. Therefore, lncRNA abnormalities in the peripheral blood of BC patients may promote BC formation and progression, promoting early diagnosis and timely treatment for patients. Many studies have confirmed the regulatory effects of lncRNAs on malignancy behaviors in tumorigenesis (Mercer et al., 2009; Fang and Fullwood, 2016; Bin et al., 2018), including proliferation, adhesion, migration, and apoptosis of tumor cells (Wang et al., 2020). HOXA11-AS and MALAT1 are identified as the lncRNAs related to cell cycle progression, which can be used as prognostic biomarkers for the survival of glioma patients (Wang et al., 2016).

Immune-related lncRNAs refer to those with various functions in regulating gene expressions and critical roles (e.g., the control of the differentiation and function of immune cell types, dendritic cell activity, T cell ratio, and metabolism) in innate or adaptive immune responses (Denaro et al., 2019). Hu et al. (2013) demonstrated that a lincRNA Ccr2-5'AS was an important part of the regulatory circuit in gene expression specific to the TH2 subset of helper T cells. Yang et al. (2020) found that LNC TANCR play important roles in  $\gamma\delta$  T cell activation. Huang et al. (2018) noticed that lncRNA NKILA, regulated T cell sensitivity to Activation-induced cell death (AICD) by inhibiting NF- $\kappa$ B activity. However, details of immune-related lncRNAs related to BC are rarely reported. In this study, an immune-related lncRNA signature was constructed and verified in an independent set and its association with immune cell infiltration was analyzed to evaluate the clinical predictive value of the signature.

## MATERIALS AND METHODS

### Data Download, Pre-processing and Screening of Prognostic IRGs

To identify BC-related lncRNAs, lncRNA profile, and clinical data from the TCGA training set were downloaded from the UCSC Xena database<sup>1</sup>; To verify the multi-lncRNA signature, gene profile and clinical information of the Breast Cancer-KR (BC-KR) data set were downloaded from ICGC<sup>2</sup>; GEOquery package with R language (Davis and Meltzer, 2007) was used to download and analyze the reliable BC expression data set GSE21653 from the GEO database, the source of species was human, and the platform was based on the GPL570 ([HG-U133\_Plus\_2] Affymetrix Human Genome U133 Plus 2.0 Array), including 266 BC samples included in this study. The BC-KR data set and the GSE21653 data set were used as verification sets. Firstly, the gene expression data and the corresponding clinical data of BC patients were pre-processed; the immune response genes (IRGs) were downloaded and sorted out from ImmPort database<sup>3</sup>. The prognostic immune-related lncRNAs were screened out from the TCGA data set using the univariate Cox regression method, and the statistical standard of  $p < 0.05$  was used. We performed GO and KEGG enrichment analysis to extract prognostic IRGs (Supplementary Figure 1).

### Screening, Co-expression Analysis and GO Enrichment Analysis of Immune-Related lncRNAs

The gene list of expression matrix was annotated and classified by mRNA and lncRNA, the mRNA expression matrix and lncRNA expression matrix were distinguished, and Pearson coefficient between lncRNA gene expression and prognostic IRGs was calculated. Pearson coefficient  $< -0.4$  and  $P < 0.01$  were set as the identification criteria of lncRNAs related to immune. In order to explore the possible function of lncRNAs related to immune in the occurrence and development of BC, co-expression analysis was performed on mRNA expression matrix and immune-related lncRNAs expression matrix. GO enrichment analysis was performed to annotate the immune functions of target immune-related lncRNAs using the ClusterProfiler package (Yu et al., 2012). Adj  $p$  value  $< 0.05$  was considered statistically significant.

### Construction and Evaluation of a Prognosis Model

A least absolute shrinkage and selection operator (Lasso) Cox regression model was used to identify optimal lncRNA candidates and construct the immune-related lncRNA signature. The process of cross-verification is to select samples randomly. The optimal value of  $\lambda$  was identified to train the model. Optimal lncRNA candidates were included during the training and stopped when the area under the curve (AUC) of

<sup>1</sup><https://xenabrowser.net/datapages/>

<sup>2</sup><https://dcc.icgc.org/>

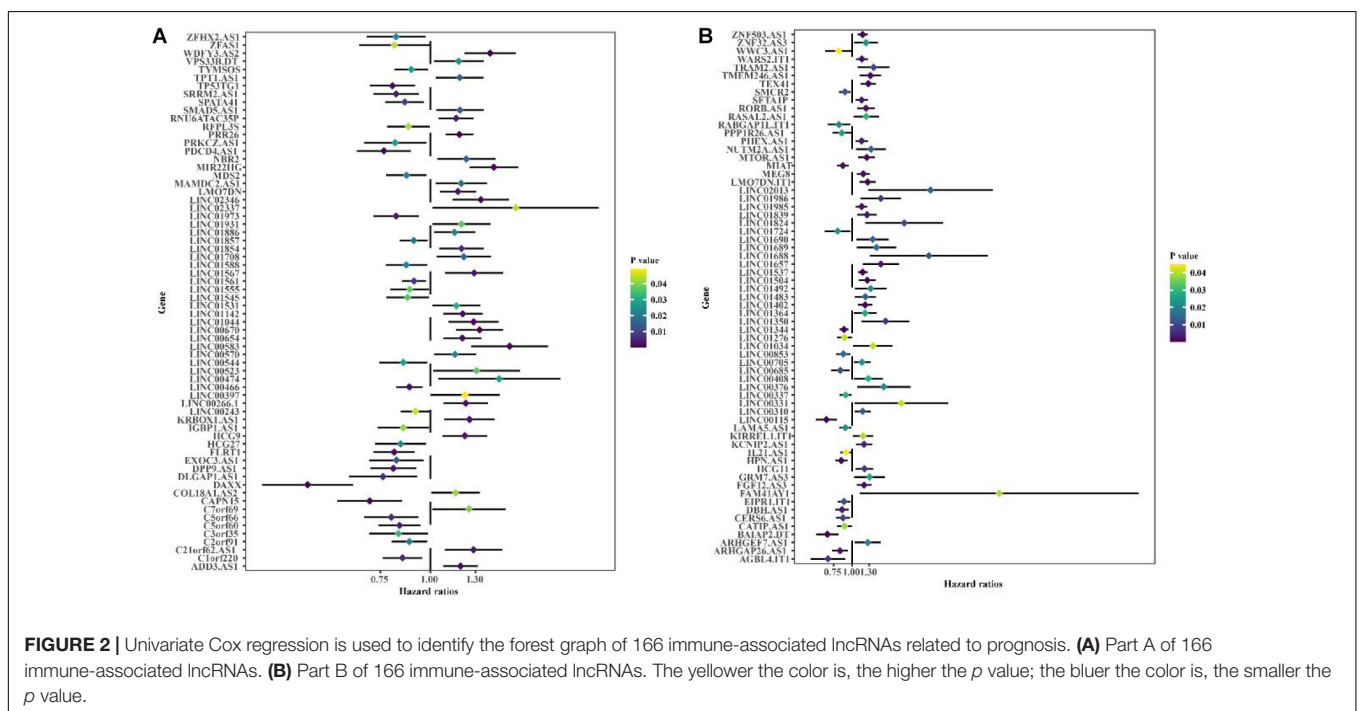
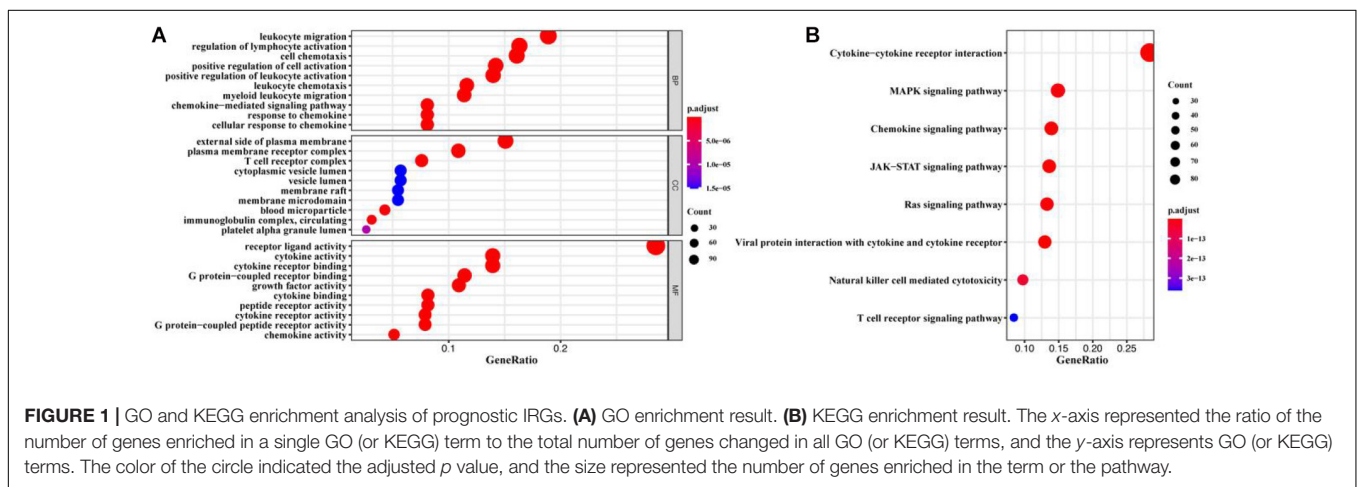
<sup>3</sup><https://ImmPort.niaid.nih.gov>

ROC reached the peak. Therefore, the optimal multi-lncRNA signature containing the least number of prognostic immune-related lncRNAs was refined (Tibshirani, 1997). Subsequently, a univariate Cox regression analysis was implemented to identify and confirm prognostic immune-related lncRNAs in the signature, with  $p < 0.05$  as the standard. An iterative Lasso Cox regression analysis was performed to construct the prognostic immune-related lncRNA signature using the glmnet package (Friedman et al., 2010). The specific method was to count the consensus genes whose gene frequency is more than 100 after 1000 Lasso Cox regression. Then, a multivariate Cox regression analysis was conducted to assess the prognostic capacity of the refined prognostic gene signature, using the expression values of genes as covariates. Subsequently, a prognostic risk signature was conducted for each patient from the TCGA training cohort

according to the refined prognostic gene-expression signature. Using the median risk score as the threshold, these patients were divided into the high- and low-risk groups and the survival differences between the high- and low-risk groups were analyzed using the R software.

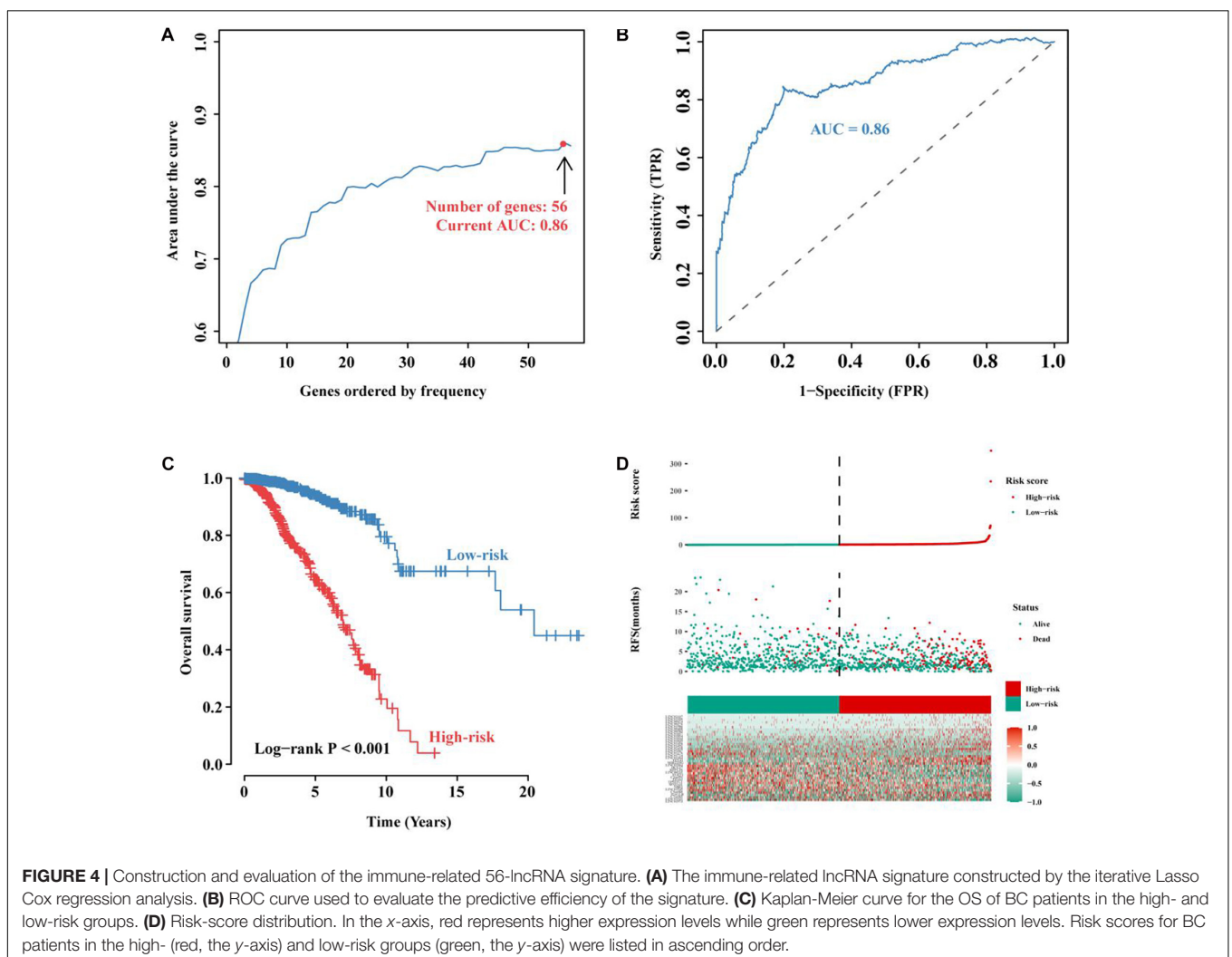
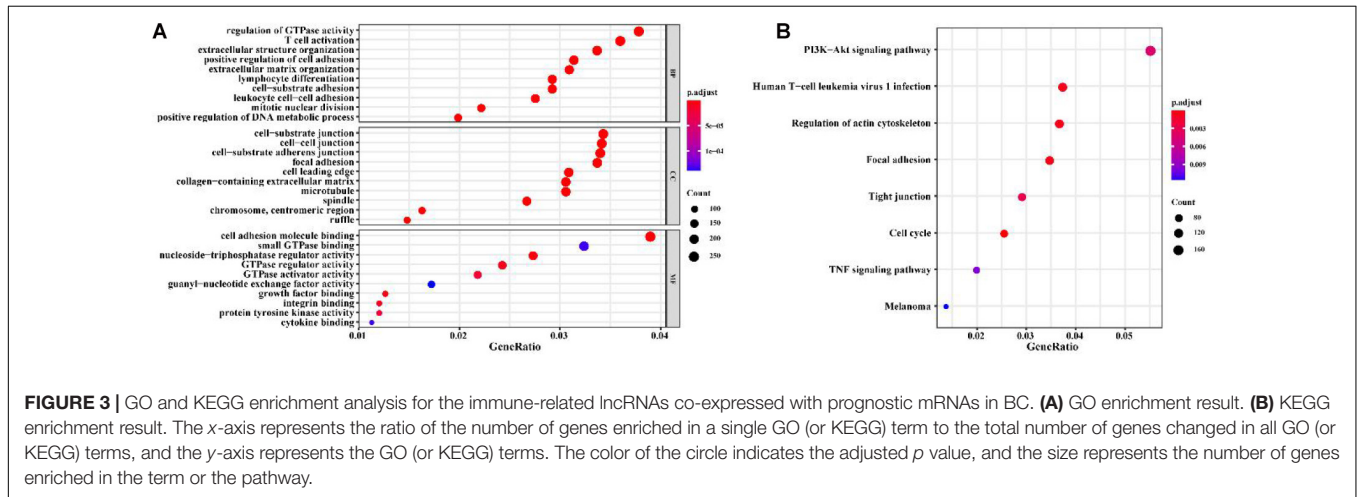
### Comparison of the Prognostic Multi-LncRNA Signature With Known Biomarkers and Validation of the Signature

To explore the prognostic capability and prediction efficiency of our signature for BC, a comparative survival analysis was used to compare the prognostic capability between the signature and known prognostic biomarkers, and



the ROC curve was used to analyze the sensitivity and specificity of the prognostic risk signature. Based on the prognostic multi-lncRNA signature in the training

set, its prognostic performances were verified in two independent validation sets (BC-KR and GSE21653 validation sets).





## Subgroup Survival Analysis, the Prognostic Value of the Signature, and Construction of Nomograms

The prognostic value of any biomarker can be evaluated by whether they are independent of the existence of clinicopathological prognostic factors. In this study, to evaluate the independence and applicability of the validated signature, the BC patients in the TCGA cohort were reassigned according to different clinicopathological characteristics. Kaplan-Meier survival analysis was performed to compare the prognostic capability between subgroups. Decision curve analysis was used to evaluate the net benefit and clinical value of the signature. Univariate Cox regression and multivariate Cox regression were used to evaluate the associations between clinical features and the risk score values, which were illustrated as nomograms.

## Phenotype Differences Between the High- and Low-Risk Groups Using GSEA and GSVA Analysis

Differences in functional phenotypes between the high- and low-risk groups were analyzed using the gene set enrichment analysis (GSEA) and GSVA analysis. We selected “c2.all.v7.1.symbols.gmt,” “c5.all.v7.1.symbols.gmt,” and “h.all.v7.1.symbols.gmt” as the reference gene sets. The differences in activated pathways between the high- and low-risk groups were identified using GSEA v4.0.3 software, and permutations of the gene sets were performed 1000 times for each analysis. A nominal  $p < 0.05$  and a false discovery rate (FDR)  $< 0.05$  were considered as statistical significance. We

selected “h.all.v7.1.symbols.gmt” as the reference gene set. The differences in activated pathways were also analyzed using the ClusterProfiler package and GSVA package (Hänzelmann et al., 2013). A  $p$  value  $< 0.05$  was considered statistically different.

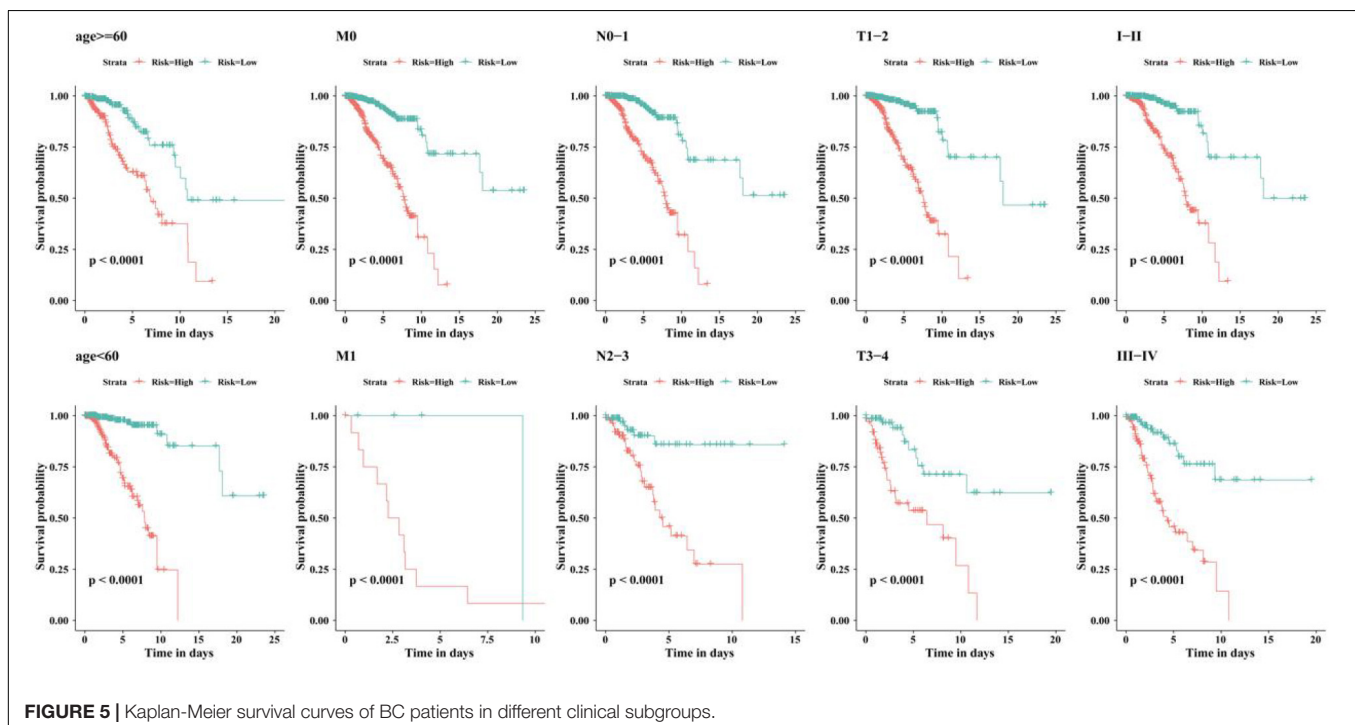
## Correlation Between the Prognostic Signature and Immune Cell Infiltration

To explore the association between the lncRNA signature and immune cell infiltration in BC, gene expression matrix data was uploaded to CIBERSORT (Newman et al., 2015) for cell type-specific gene expression purification. Infiltrating immune cells were filtered for analysis, and the abundances of infiltrating immune cells between the high- and low-risk groups were compared with CIBERSORT  $p < 0.05$  for the eligible samples. Survival-related infiltrating immune cells were identified using Kaplan-Meier survival analysis. Finally, the correlation analysis was performed for the obtained prognostic biomarkers and immune cell infiltration. The results were visualized by using ggplot2 package (Ginestet, 2011).

## RESULTS

### Screening of Immune-Related LncRNAs for the Prognosis of BC Patients

A total of 1900 IRGs were retrieved from the ImmPort database. Univariate Cox regression analysis showed that 516 IRGs (Supplementary Table 1) were associated with the prognosis of BC patients. We further identified immune functions of the prognostic IRGs via the gene GO and KEGG enrichment. We found that the 516 prognostic IRGs



were significantly enriched in GO terms related to leukocyte migration, regulation of lymphocyte activation, cell chemotaxis, external side of plasma membrane, plasma membrane receptor complex, T cell receptor complex, and receptor-ligand activity, and in KEGG pathways related to cytokine-cytokine receptor interaction, MAPK signaling pathway, chemokine signaling pathway, and JAK-STAT signaling pathway (Figure 1 and Supplementary Table 2).

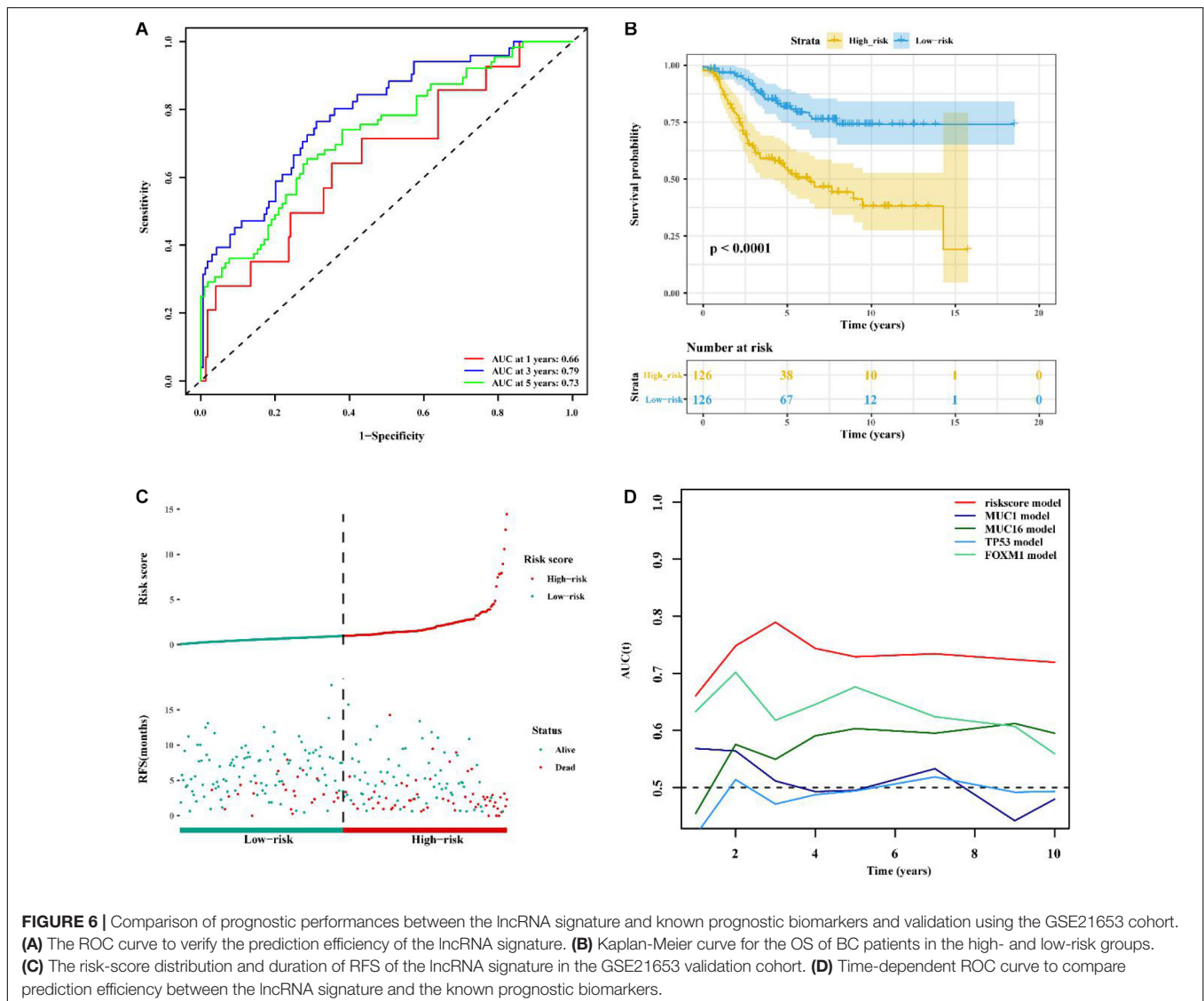
### Co-expression Analysis and GO Enrichment Analysis of Immune-Related lncRNAs

Co-expression analysis showed 948 immune-related lncRNAs linked to immune responses in BC (Supplementary Table 3). To explore the prognostic capacity of the immune-related lncRNAs, univariate Cox regression analysis was implemented and identified that 166 immune-related lncRNAs (Supplementary

Table 4) were related to the maximum prognostic value ( $P < 0.05$ , Figure 2;  $P < 0.05$ ). To understand the immune function of the 166 genes, we performed functional enrichment analysis for these genes. GO enrichment analysis revealed that these genes were mainly enriched in GO terms related to the regulation of GTPase activity, T cell activation, extractable structure organization, cell-substitute junction, cell adhesion, and molecular binding (Figure 3 and Supplementary Table 5). These results suggest that immune-related lncRNAs may participate in the occurrence and development of BC by regulating immune responses.

### Construction and Evaluation of the Prognostic Model

To further improve the prediction efficiency of the signature, we performed iterative Lasso Cox regression analysis and built an optimal signature containing 56 immune-related lncRNAs (Figure 4A). To verify its predictive efficiency, we



performed ROC analysis and showed that the immune-related 56-lncRNA signature had a good prognostic performance in BC (AUC = 0.86, **Figure 4B**). Kaplan-Meier survival analysis showed that the OS of patients in the high-risk group was significantly worse than that in the low-risk group ( $P < 0.001$ , **Figure 4C**). The distribution risk scores, duration of recurrence-free survival (RFS), and lncRNA expression levels were illustrated in **Figure 4D**. The number of deaths in the high-risk group was significantly higher than that in the low-risk group. These results suggest that the immune-related 56-lncRNA signature can distinguish high-risk patients from low-risk individuals, and it is of great significance in predicting the prognosis of BC.

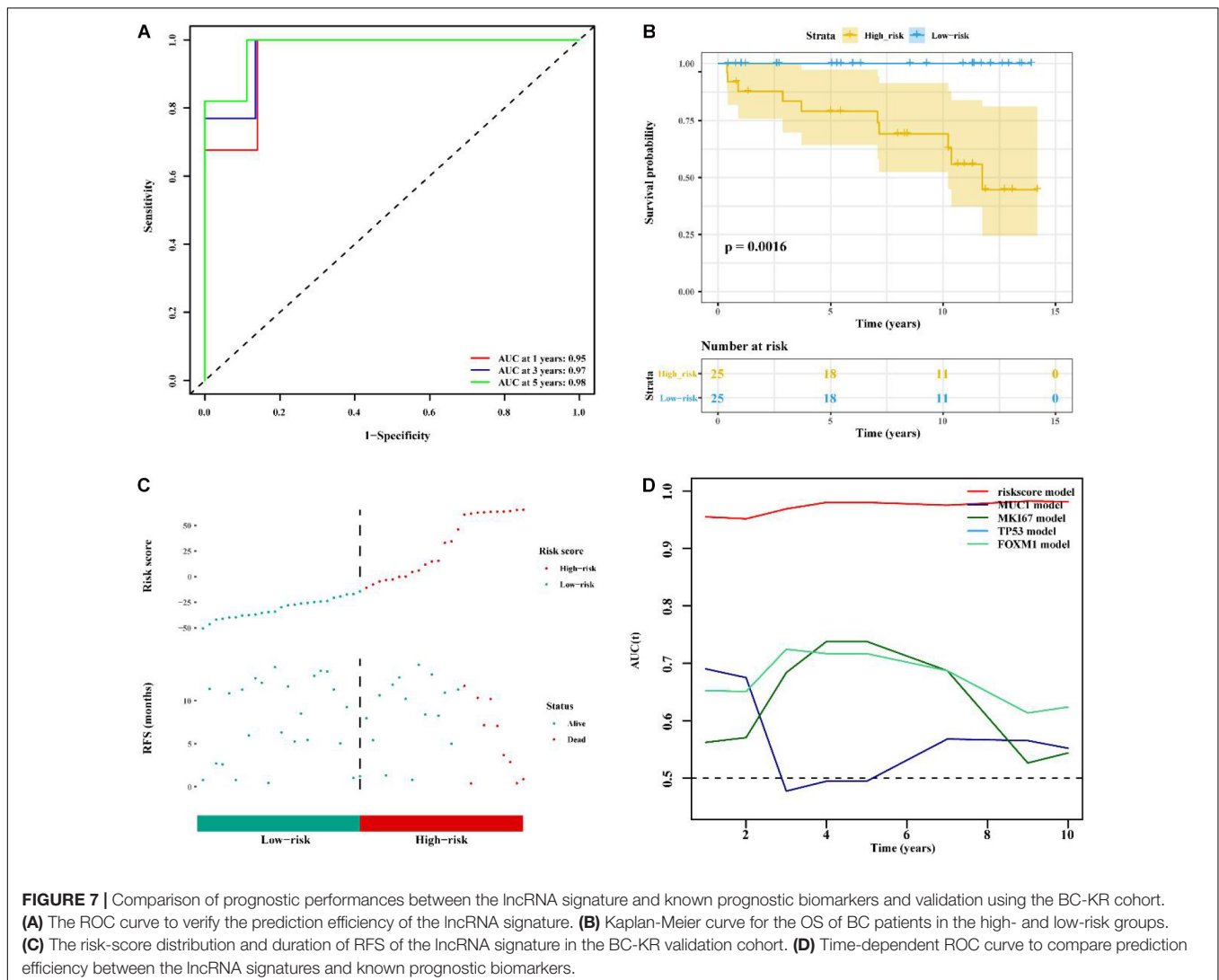
### The lncRNA Signature as an Independent Prognostic Predictor

We examined the independent prognostic ability of the immune-related lncRNA signature. The BC patients from the TCGA cohort were reassigned according to different

prognostic and clinicopathological factors consisting of age, metastasis, and tumor staging (TNM and clinical staging). Kaplan-Meier survival analysis was performed for each subgroup and showed that the survival time of BC patients in the low-risk group was significantly prolonged regardless of age, TNM staging, and clinical staging ( $p < 0.0001$ , **Figure 5**). This indicates that the immune-related 56-lncRNA signature is independent of gender, age, and metastasis in survival prediction, which can be considered to be an independent predictor of the prognosis of BC patients.

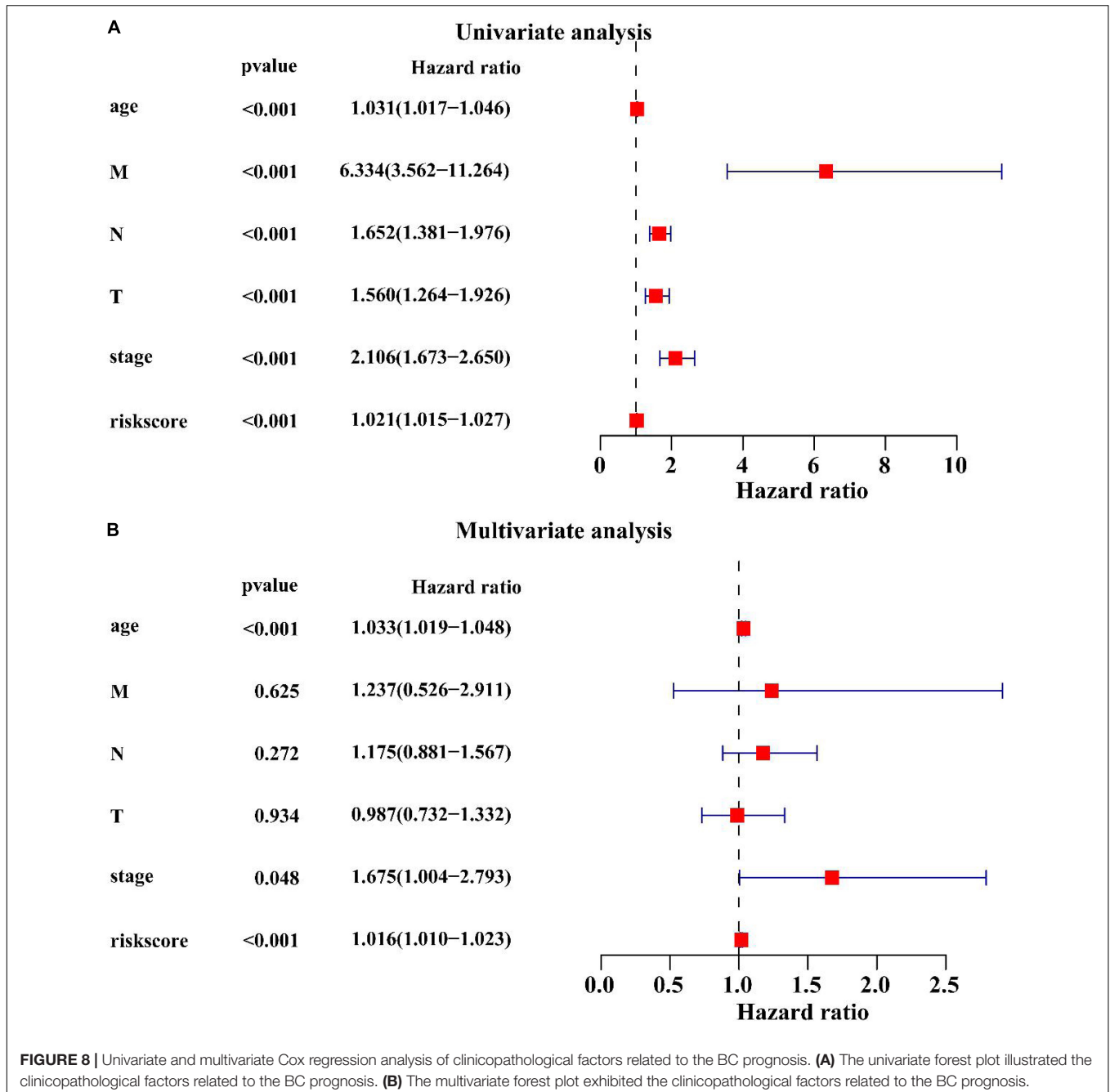
### Validation of the lncRNA Signature and the Comparison Between the Signature and Known Prognostic Biomarkers

To determine whether the immune-related 56-lncRNA signature was superior to known prognostic biomarkers in terms of prognostic capability and prediction efficiency, the validation was performed in the other two independent verification set



GSE21653 data set and the BC-KR data set cohort. The ROC analysis showed a similar prediction efficiency between the training cohort and the GSE21653 validation cohort (Figure 6A). Kaplan-Meier survival analysis showed that the OS of BC patients in the high-risk group was significantly worse than that in the low-risk group ( $P < 0.001$ , Figure 6B). The distribution of risk scores and duration of RFS were illustrated in Figure 6C. The number of deaths in the high-risk group was significantly higher than that in the low-risk group. ROC analysis showed that the AUC of the lncRNA signature was higher than that

of known prognostic biomarkers (Figure 6D). The lncRNA signature had better stability and reliability in predicting the survival of BC patients, which should be considered as a good prognostic biomarker. The validation using the BC-KR cohort showed similar results in prognostic capability and prediction efficiency of the signature, more deaths in the high-risk group, and the superiority over known biomarkers (Figure 7). These results suggest that the immune-related 56-lncRNA signature can predict the survival of BC patients in other independent cohorts.





## Univariate and Multivariate Cox Regression Analysis of Clinicopathological Factors Related to the BC Prognosis

Univariate Cox regression analysis showed that age (HR = 1.031,  $P < 0.001$ ), M stage (HR = 6.334,  $P < 0.001$ ), N stage (HR = 1.652,  $P < 0.001$ ), T stage (HR = 1.560,  $P < 0.001$ ), clinical stage (HR = 2.106,  $P < 0.001$ ), the lncRNA signature (HR = 1.021,  $P < 0.001$ ) were associated with the BC prognosis (Figure 8A). Multivariate Cox regression analysis revealed that age (HR = 1.033,  $P < 0.001$ ), clinical stage (HR = 1.675,  $P < 0.001$ ) and the lncRNA signature (HR = 1.016,  $P < 0.001$ ) were significantly associated with the BC prognosis (Figure 8B). These results suggest that the lncRNA signature is an important prognostic factor for BC.

## Construction of a Nomogram to Predict the OS of Patients With BC

Combined with clinicopathological characteristics, we constructed a nomogram containing the immune-related 56-lncRNA signature (Figure 9) to predict the 3-, 5-, and 8-year survival rate of BC patients. The results suggest that the nomogram can independently evaluate the survival of BC patients, which may help physicians make better medical decisions and follow-up plans.

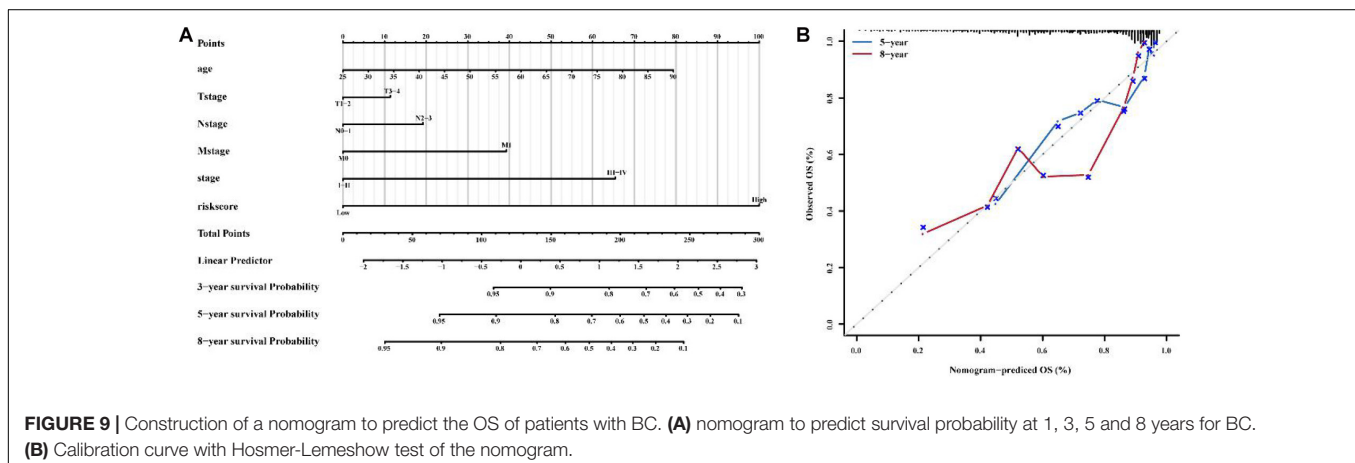
## GSEA and GSVA Analysis for Phenotype Differences Between the High- and Low-Risk Groups

We conducted GSEA and GSVA analysis to identify differences in main functional phenotypes between the high- and low-risk groups. As shown in Figures 10A,B and Supplementary Figure 2A, GSEA analysis revealed that the pathways involved in BC patients from the high-risk group consisted of CELL\_CELL\_JUNCTION, DUTERTRE ESTRADIOL\_RESPONSE\_24HR\_DN, EPITHELIAL\_MESENCHYMAL\_TRANSITION, KRAS\_SIGNALING\_Up, Zhang\_INTERFERON\_RESPONSE, and POSITIVE

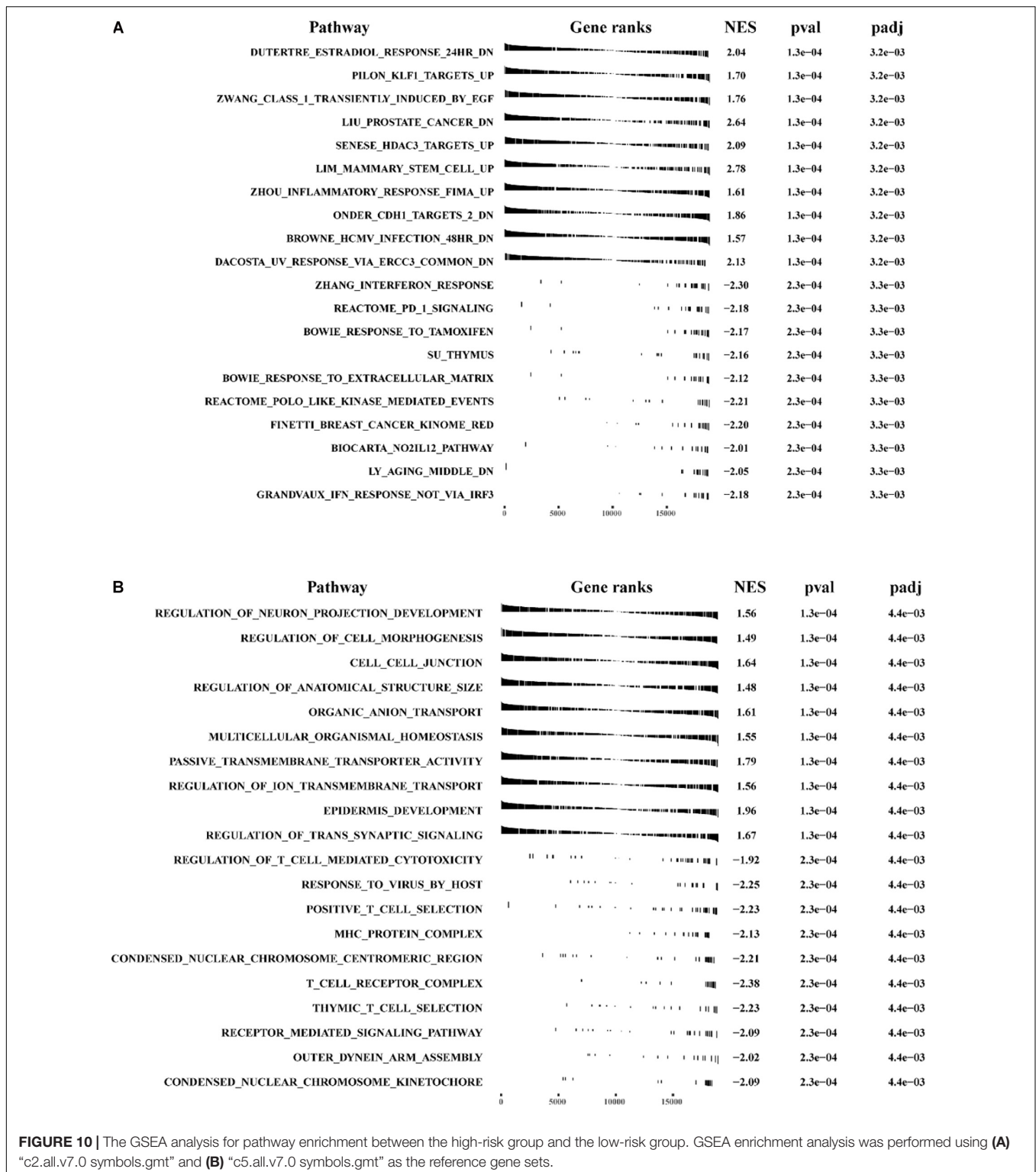
\_T\_CELL\_SELECTION. GSVA analysis showed that HEME\_METABOLISM, TGF\_BETA\_SIGNALING, KRAS\_SIGNALING\_UP, IL6\_JAK\_STAT3\_SIGNALING, INFLAMMATORY\_RESPONSE were activated in BC patients from the high-risk group, and ESTROGEN\_RESPONSE\_LATE, MYC\_TARGETS\_V1, DNA\_REPAIR, and INTERFERON\_GAMMA\_RESPONSE were inhibited (Supplementary Figure 2B and Supplementary Table 6).

## Correlation Between Immune Cell Infiltration and the lncRNA Signature

Based on the mRNA-lncRNA co-expression network, GO enrichment analysis had confirmed that the immune-related lncRNA signature was involved in the occurrence and development of BC by regulating immune responses. To explore the association between the lncRNA signature and immune cell infiltration, a comprehensive analysis was carried out to figure out the composition of infiltrating immune cells in the tumor microenvironment. Correlation analysis showed that cytotoxic cells were positively correlated with T cells and B cells, while Th2 cells were negatively correlated with other immune cells (Figure 11A). We predicted the interaction network between immune cells, and the results showed that CD8<sup>+</sup> T cells, Th1 cells, and B cells had a strong interaction with other immune cells, while Tgd, T helper cells, and Th17 cells had a weak interaction with other immune cells (Figure 11B). The analysis of immune cell composition showed that BC patients had higher numbers of T helper cells, CD8<sup>+</sup> T cells, and T helper cells but lower numbers of B cells and regulatory T (TReg) cells in the tumor microenvironment (Figure 11C). We compared the difference in the expressions of infiltrating immune cells in BC patients between the high-risk group and the low-risk group. As shown in Figure 11D, the levels of aDC, Macrophages, Tgd, Macrophages, and TReg were significantly different between the two groups ( $P < 0.01$ ). The prognosis analysis for infiltrating immune cells showed that aDC, Neutrophils, Tem, DC, NK CD56dim cells, Tgd, Mast cells, TCM, and TReg were correlated with the prognosis of BC patients (Figure 12). The correlation analysis confirmed a significant correlation between the infiltrating



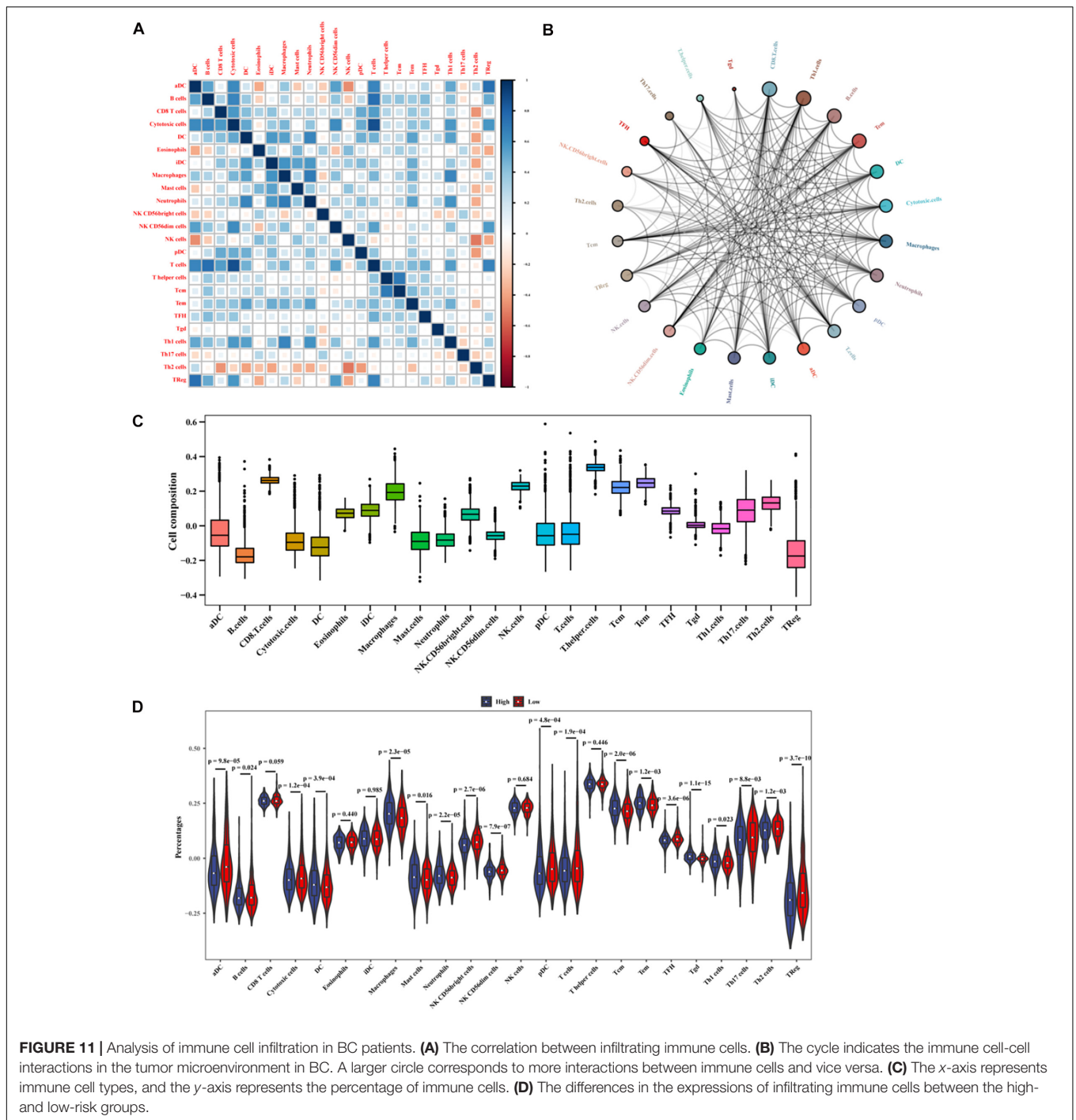
**FIGURE 9 |** Construction of a nomogram to predict the OS of patients with BC. (A) nomogram to predict survival probability at 1, 3, 5 and 8 years for BC. (B) Calibration curve with Hosmer-Lemeshow test of the nomogram.



immune cells and the lncRNAs in the signature (Supplementary Figure 3). These results suggest that the lncRNAs in the signature have a complex regulatory relationship with immune cell infiltration, which may be an important way to participate in the development of BC and have high research value.

## DISCUSSION

Breast cancer remains one of the most lethal malignancies in the world. Although currently available cancer treatments (such as surgery, chemotherapy, radiotherapy, endocrine therapy,



targeted therapy, and immunotherapy) have a positive impact on BC survival, increasing life expectancy in a short period seems unrealistic. The risk of BC recurrence is particularly high, suggesting that BC recurrence is one of the most important causes of mortality. The difficulty of predicting the BC prognosis among a complex tumor microenvironment comprised of molecular and cellular heterogeneity between some commonly used markers in several cancers, such as CA 15-3, CA 27.29, CEA, MUC1, MUC16, TP53, FOXM1, ER, PR, HER2, KI 67, and others (Børresen-Dale, 2003;

Lakshmanan et al., 2012; Graham et al., 2014; Kabel, 2017; Song et al., 2017; Stergiou et al., 2019), makes a prediction inaccurate. Therefore, the proposal of an effective and reliable biomarker using a risk prediction model and cohort-specific imputation of gene expression to encourage the implementation of relevant clinical trials to validate this biomarker is currently possible. Advantages of a reliable biomarker are that high-risk populations can be closely monitored, management for the prevention and early detection of BC recurrence can be implemented, and once the recurrence is timely detected,

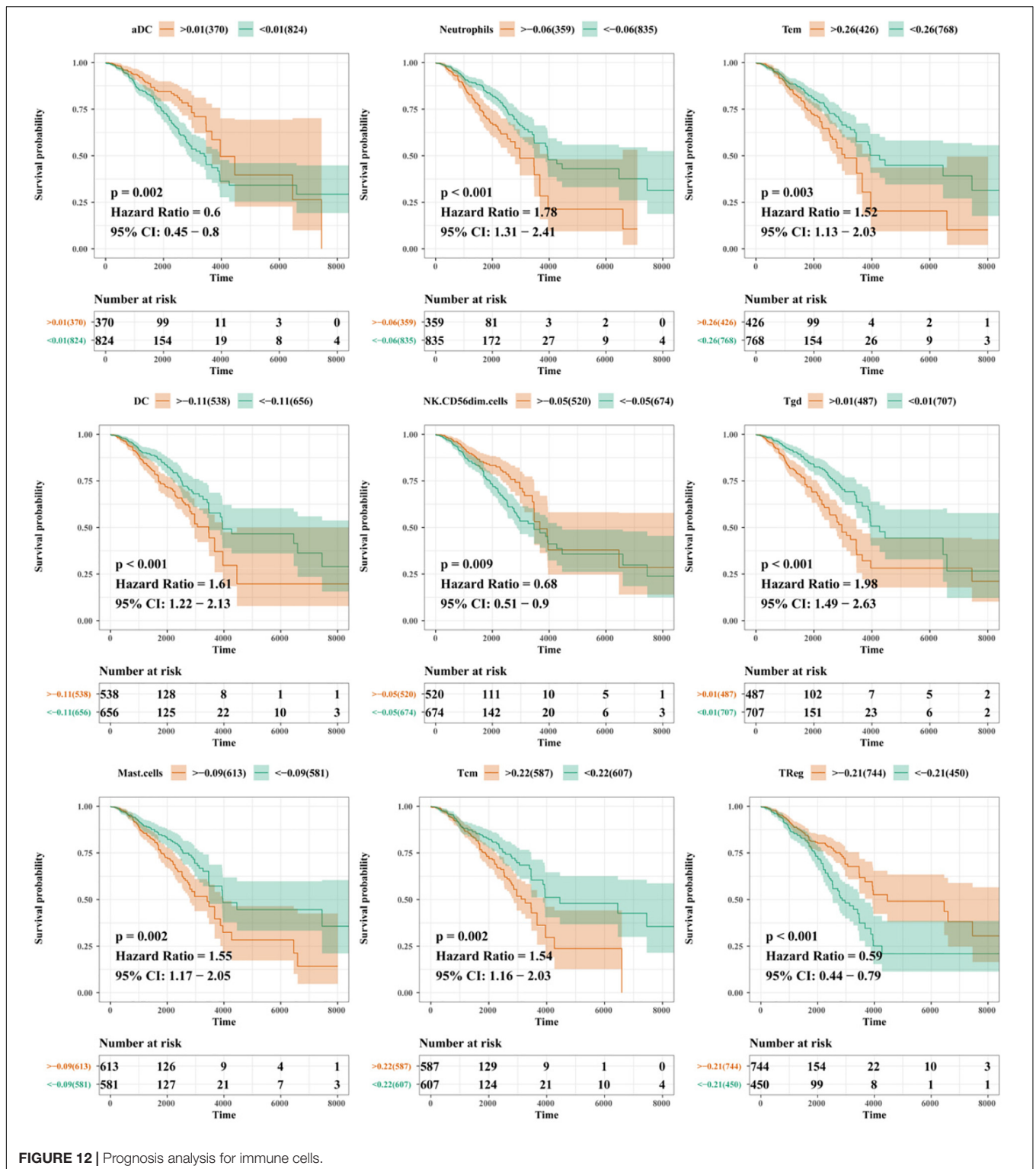


FIGURE 12 | Prognosis analysis for immune cells.

individualized treatment plans for prolonged survival can be prescribed.

Previous studies of prognostic biomarkers of BC have focused on coding genes and microRNAs. lncRNAs, especially immune-related lncRNAs, recently emerging as novel biomarkers

have gained primary results. In-depth sequencing studies have found that lncRNAs directly interact with a variety of signaling pathways and regulators of oncogenes or tumor suppressor genes, thereby affecting tumor occurrence and development (Lin and Yang, 2018).



In this study, we initially identified 948 immune-related lncRNA candidates. Using a signature risk prediction model and the iterative Lasso Cox regression analysis, 56 lncRNAs with the maximum prognostic values were selected from the candidates to form a signature. ROC analysis showed good prognostic efficiency, with an AUC of 0.86. Finally, a 56-lncRNA signature was constructed. In the test set, BC patients were divided into the high- and low-risk groups according to their risk scores. Kaplan-Meier curve showed that patients in the low-risk group had a significantly longer OS than those in the high-risk group. Moreover, multivariate Cox regression analysis showed that the lncRNA signature was an independent predictor for BC prognosis in the two other validation cohorts from GEO and ICGC. Besides, the lncRNA signature offers a more effective prediction than known biomarkers. Clinical subgroup analyses showed the signature was independent of the presence of clinicopathological factors (gender, age, metastatic lymph nodes, and tumor metastasis) related to BC prognosis, indicating that the signature is an independent predictor of BC survival and the perspective of its clinical applicability is expected. The construction of a nomogram confirmed that the lncRNA signature can independently evaluate BC survival.

Cancer immunotherapy is a rapidly developing and exciting field of oncology. Eliciting adaptive immunity, particularly the adaptive infiltrating immune system which is an important component in the tumor microenvironment, is the goal of immunotherapy for effective and sustained anti-tumor responses to limit tumor growth and metastasis. Studies have shown that a combination of immunotherapy and other therapies can bring better benefits to patients with BC, especially advanced-stage, triple-negative BC (Schmid et al., 2018). In this study, GEVA analysis of mRNAs that are co-expressed with immune-related lncRNAs showed that TGF\_BETA\_SIGNALING and IL-6/JAK/STAT3TGF were the main enriched pathways. Hao et al. (2017) reported that TGF- $\beta$  signaling pathway plays an important role in regulating the migration and invasion of lung cancer and BC cells. The IL-6/JAK/STAT3 signaling pathway is critical in regulating cell growth, survival, and differentiation in tumor cells, and also in mediating the differentiation and activation of T lymphocytes (Eskiler et al., 2019). Immune-related lncRNAs may participate in the occurrence and development of BC via directly interacting with a variety of signaling pathways and regulators of oncogenes or tumor suppressor genes to regulate adaptive immune responses. We further studied the relationship between lncRNAs in the signature and immune cell infiltration, and found that the levels of aDC, B cells, NK, CD56dim cells, T cells, TFH, and TReg in the low-risk group were higher than those in the high-risk group, aDC, NK, CD56DIM cells and TReg were correlated with BC prognosis, and aDC, NK, CD56DIM cells were positively correlated with T cells. Tumor-infiltrating lymphocytes are an important component of the tumor microenvironment in BC. Pantelidou et al. (2019) reported that Olaparib, a PARP inhibitor, inhibited tumor via activating the cGAs/STING pathway in tumor cells to induce T cell recruitment.

Pei et al. (2018) reported that LNC SNHG1 functioning as a competing endogenous RNA (CeRNA) inhibited the differentiation of TReg cells, thereby preventing BC immune escape. Wang et al. (2019) reported that TReg cells play

a major role in the development of immunosuppressive tumor microenvironment mainly through cytokine signal transductions—a key determinant of immunosuppressive potential and unfavorable clinical outcomes. Inconsistent with these findings, we found that BC patients with a high TReg level had a better prognosis than those with a low TReg level (Takasato et al., 2020). The origin of tumor-infiltrating TReg cells and their association with circulating TReg cells are rarely reported, which will be explained in the follow-up study in our following publications.

Besides, there are some limitations of our study. Further cell and animal experiments are needed to verify the correlation between lncRNA expressions and immunophenotype in BC. To explore the underlying immune mechanisms, a larger data set is needed to verify the accuracy of the prediction model.

## CONCLUSION

In conclusion, we have constructed an immune-related 56-lncRNA signature which can be used as an independent prognostic marker for the survival risk subgroup of BC patients. This lncRNA signature is superior to known predictors in terms of prognostic performance in risk stratification, which will renew hope for experimental and clinical validations as well as clinical application in the future.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

CY, WH, and NL conceived and designed the study. ZH, CX, and LY performed the experiments. FZ, WH, and ZZ analyzed the data. ZH wrote the manuscript. All authors have read and approved this manuscript.

## FUNDING

This work was supported by Science and Technology Program of Fujian Province, China (No. 2018Y2003); the Natural Science Foundation of Fujian Province (No. 2020J011112); and the Natural Science Foundation of Fujian Province (No. 2018J01265).

## ACKNOWLEDGMENTS

The authors thank reviewers for helpful comments on the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.634195/full#supplementary-material>

## REFERENCES

- Bin, X., Hongjian, Y., Xiping, Z., Bo, C., Shifeng, Y., and Binbin, T. (2018). Research progresses in roles of lncRNA and its relationships with breast cancer. *Cancer Cell Int.* 18:179. doi: 10.1186/s12935-018-0674-0
- Børresen-Dale, A.-L. (2003). TP53 and breast cancer. *Hum. Mutat.* 21, 292–300. doi: 10.1002/humu.10174
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Davis, S. R., and Meltzer, P. S. (2007). GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Cancer Res.* 23, 1846–1847. doi: 10.1093/bioinformatics/btm254
- Denaro, N., Merlano, M. C., and Lo Nigro, C. (2019). Long noncoding RNA s as regulators of cancer immunity. *Mol. Oncol.* 13, 61–73.
- Eskiler, G. G., Bezdegumeli, E., Ozman, Z., Ozkan, A., Deveci, Bilir, C., Kucukakca, B. N., et al. (2019). IL-6 mediated JAK/STAT3 signaling pathway in cancer patients with cachexia[J]. *Bratislava Med. J.* 120, 819–826.
- Fang, Y., and Fullwood, M. J. (2016). Roles, functions, and mechanisms of long non-coding RNAs in cancer. *Genomics Proteom. Bioinform.* 14, 42–54. doi: 10.1016/j.gpb.2015.09.006
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22.
- Ginestet, C. (2011). ggplot2: elegant graphics for data analysis. *R. Stat. Soc.* 174, 245–246.
- Graham, L. J., Shupe, M. P., Schneble, E. J., Flynt, F. L., Clemenshaw, M. N., Kirkpatrick, A. D., et al. (2014). Current approaches and challenges in monitoring treatment responses in breast cancer. *J. Cancer* 5, 58–68. doi: 10.7150/jca.7047
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 14:7. doi: 10.1186/1471-2105-14-7
- Hao, Y., Yang, X., Zhang, D., Luo, J., and Chen, R. (2017). Long noncoding RNA LINC01186, regulated by TGF- $\beta$ /SMAD3, inhibits migration and invasion through Epithelial-Mesenchymal-Transition in lung cancer[J]. *Gene* 608, 1–12.
- Hu, G., Tang, Q., Sharma, S., Yu, F., Escobar, T. M., Muljo, S. A., et al. (2013). Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nat. Immunol.* 14, 1190–1198. doi: 10.1038/ni.2712
- Huang, D., Chen, J., Yang, L., Ouyang, Q., Li, J., Lao, L., et al. (2018). NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death[J]. *Nat. Immunol.* 19, 1112–1125.
- Kabel, A. M. (2017). Tumor markers of breast cancer: new perspectives. *J. Oncol. Sci.* 3, 5–11. doi: 10.1016/j.jons.2017.01.001
- Kim, Y. A., Cho, H., Lee, N., Jung, S. Y., Sim, S. H., Park, I. H., et al. (2018). Doxorubicin-induced heart failure in cancer patients: a cohort study based on the Korean National Health Insurance Database. *Cancer Med.* 7, 6084–6092. doi: 10.1002/cam4.1886
- Lakshmanan, I., Ponnusamy, M. P., Das, S., Chakraborty, S., Haridas, D., Mukhopadhyay, P., et al. (2012). MUC16 induced rapid G2/M transition via interactions with JAK2 for increased proliferation and anti-apoptosis in breast cancer cells. *Oncogene* 31, 805–817. doi: 10.1038/onc.2011.297
- Li, N., Deng, Y., Zhou, L., Tian, T., Yang, S., Wu, Y., et al. (2019). Global burden of breast cancer and attributable risk factors in 195 countries and territories, from 1990 to 2017: results from the Global Burden of Disease Study 2017. *J. Hematol. Oncol.* 12:140. doi: 10.1186/s13045-019-0828-0
- Lin, C., and Yang, L. (2018). Long noncoding RNA in cancer: wiring signaling circuitry. *Trends Cell. Biol.* 28, 287–301. doi: 10.1016/j.tcb.2017.11.008
- Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155–159. doi: 10.1038/nrg2521
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337
- Pantelidou, C., Sonzogni, O., De Oliveria Taveira, M., Mehta, A. K., Kothari, A., Wang, D., et al. (2019). PARP inhibitor efficacy depends on CD8 T-cell recruitment via intratumoral STING pathway activation in BRCA-deficient models of triple-negative breast cancer. *Cancer Discov.* 9, 722–737. doi: 10.1158/2159-8290.Cd-18-1218
- Pei, X., Wang, X., and Li, H. (2018). 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.* 118, 24–30. doi: 10.1016/j.ijbiomac.2018.06.033
- Schmid, P., Adams, S., and Rugo, H. S. (2018). Atezolizumab and NabPaclitaxel in advanced triple-negative breast cancer. *N. Engl. J. Med.* 379, 2108–2121.
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590
- Song, X., Fiati Kenston, S. S., Zhao, J., Yang, D., and Gu, Y. (2017). Roles of FoxM1 in cell regulation and breast cancer targeting therapy. *Med. Oncol.* 34:41. doi: 10.1007/s12032-017-0888-3
- Stergiou, N., Gaidzik, N., Heimes, A.-S., Dietzen, S., Besenius, P., Jäkel, J., et al. (2019). Reduced breast tumor growth after immunization with a tumor-restricted MUC1 glycopeptide conjugated to tetanus toxoid. *Cancer Immunol. Res.* 7, 113–122. doi: 10.1158/2326-6066.Cir-18-0256
- Takasato, Y., Kurashima, Y., Kiuchi, M., Hirahara, K., Murasaki, S., Arai, F., et al. (2020). Orally desensitized mast cells form a regulatory network with Treg cells for the control of food allergy. *Mucosal Immunol.* [Epub ahead of print]. doi: 10.1038/s41385-020-00358-3
- Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Stat. Med.* 16, 385–395. doi: 10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380<3.0.co;2-3
- Wang, J., Yang, S., Ji, Q., Li, Q., Zhou, F., Li, Y., et al. (2020). Long non-coding RNA EPIC1 promotes cell proliferation and motility and drug resistance in glioma. *Mol. Ther. Oncolytics* 17, 130–137. doi: 10.1016/j.omto.2020.03.011
- Wang, L., Simons, D. L., Lu, X., Tu, T. Y., Solomon, S., Wang, R., et al. (2019). Connecting blood and intratumoral T cell activity in predicting future relapse in breast cancer. *Nat. Immunol.* 20, 1220–1230. doi: 10.1038/s41590-019-0429-7
- Wang, Q., Zhang, J., Liu, Y., Zhang, W., Zhou, J., Duan, R., et al. (2016). A novel cell cycle-associated lncRNA, HOXA11-AS, is transcribed from the 5-prime end of the HOXA transcript and is a biomarker of progression in glioma. *Cancer Lett.* 373, 251–259. doi: 10.1016/j.canlet.2016.01.039
- Yang, C., Feng, T., Lin, F., Gong, T., Yang, S., Tao, Y., et al. (2020). Long noncoding RNA TANC1 promotes  $\gamma\delta$  T cells activation by regulating TRAIL expression in cis. *Cell Biosci.* 10:15. doi: 10.1186/s13578-020-00383-6
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Huang, Xiao, Zhang, Zhou, Yu, Ye, Huang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.