European Journal of Gynaecological Oncology

IMR Press

Original Research

Breast cancer overall-survival can be predicted with a 19 lncRNA tissue signature

Xiao-Peng Yu[1], Cai-Gang Liu[2], Fang Qiu[1], Fei Xing[1], Si-Jia Han[1], Ye Han[1], Gui-Jin He[1,*]

[1] *Secondary Department of Breast Surgery, Shengjing Hospital of China Medical University, 110004 Shenyang, Liaoning, China*

[2] *First Department of Breast Surgery, Shengjing Hospital of China Medical University, 110004 Shenyang, Liaoning, China*

*Correspondence: hegi@sj-hospital.org (Gui-Jin He)

DOI:10.31083/j.ejgo4205128

Submitted: 5 November 2019 Revised: 6 May 2020 Accepted: 13 May 2020 Published: 15 October 2021

Objective: Studying the prognosis of breast cancer (BRCA) is of great significance for clinical treatment. LncRNA has been shown to be significantly important in breast cancer, but only few studies exist that relate to the prognosis of lncRNA. This study aimed to build a lncRNA-based breast cancer prognosis risk model using the data from TCGA datasets. *Methods*: we used the TCGA public database to explore the differential expression of lncRNA and cancer prognosis in breast cancer patients. The RNA-Seq data and clinical data pertaining to 1090 BRCA patients in the TCGA database were downloaded and analyzed. The prognosis-related lncRNAs in BRCA patients were identified in the training set and validated in the test set and the complete data set. ROC was performed to determine the optimal cut-off point for patient risk classification, and survival analysis was performed to determine its significance in prognosis prediction. *Results*: A total of 19 prognosis-associated differentially expressed lncRNAs (LSINCT5, TRG-AS1, CH17-189H20.1, RP11-1399P15.1, RP11-344P13.6, RP5-1028K7.2, AL022344.7, USP30-AS1, RP11-522I20.3, AL122127.25, BHLHE40-AS1, CHRM3-AS2, LINC00704, RP5-1073O3.2, RP11-316M21.6, CTA-384D8.31, RP11-10J5.1, RP11-426L16.3, RP11-344B5.2) were screened out. The BRCA prognosis risk assessment model based on 19-lncRNA can predict the survival rate of breast cancer patients. *Conclusion*: This model can predict the prognosis of breast cancer patients and these 19 lncRNAs can be used as potential molecular markers for breast cancer prognosis prediction.

Keywords

Breast cancer; Long non-coding RNA; Prognosis; TCGA; Predicting model

1. Introduction

Breast cancer is one of the most common malignant tumors that threaten women's health worldwide. With the continuous development of surgery, radiotherapy and chemotherapy, as well as the wide application of molecular diagnostics and genomics, the survival rate of breast cancer patients has, to some extent, increased. However, there are still about 20%–30% of early stage breast cancer patients that experience recurrence of the cancer or development of distant metastasis [1, 2]. Therefore, even though it is challenging to identify molecular markers that affect or predict the prognosis of breast cancer, it is of great significance to do so; moreover, this is necessary to improve its accuracy of early diagnosis and prognosis [3]. The development of breast cancer is the result of multi-factor, multi-stage, multi-gene interactions and long non-coding RNA (lncRNA) is among the factors involved in the development of cancer [4].

The lncRNAs, a class of RNA with a length greater than 200 nt, are thought to be an important regulatory molecule of gene expression. While their mechanisms of action are not fully understood, the past 10 years of research led to a more comprehensive picture [5–7]. It can affect apoptosis, signaling pathways, tumor invasion and metastasis of tumor cells, and plays a key role in tumorigenesis [8]. At present, numerous studies have found that the expression of lncRNAs in breast cancer patients significantly change during the development of breast cancer [9, 10]; however, there are few studies related to lncRNA in terms of prognosis estimation. K. P. Sørensen found that HOTAIR could be used as an independent risk factor for breast cancer metastasis and poor prognosis, since the tumor-free survival and overall survival of patients with high expression of HOTAIR were significantly lower than those with low expression of HOTAIR [11]. In breast cancer cells, the reduced expression of NBAT1 is associated with metastasis and poor prognosis of cancer cells, suggesting that NBAT1 is expected to be a prognostic biomarker and metastatic therapeutic target for breast cancer [12]. High expression of EPB41L4A-AS2 can inhibit tumor cell proliferation; moreover, breast cancer patients with overexpression of EPB41L4A-AS2 have a good prognosis [13]. However, few models exist that systematically predict the prognosis of breast cancer using lncRNA.

In this study, we intend to explore the TCGA public database to identify the ability of differentially expressed lncRNA in breast cancer patients to predict the prognosis of breast cancer, and to establish a lncRNAs-based model that predicts the prognosis of breast cancer.

2. Materials and method

2.1 Data collection and preprocessing

14371 lncRNA sequencing data (**Supplementary Table 1**) and corresponding clinical information of 1090 breast

cancer patients were downloaded from the TCGA public database (https://portal.gdc.cancer.gov/projects/TCGA-BRCA) on September 20, 2018. Clinical information includes survival status and survival time (see **Supplementary Table 2**). The data used in this study meets the requirements of the official TCGA published data and is publicly available. Standardization of the 14371 lncRNAs was performed by log2 (x + 1). Then, lncRNAs with expression levels >0 and expressed in more than 50% of the samples were selected as enriched expression lncRNAs. A total of 7272 enriched expression of lncRNAs were obtained.

2.2 Selection of prognosis-related lncRNAs

Using random number generation, 1090 breast cancer patients were randomized into a training set (n = 545) and a test set (n = 545). The training set is used to learn sample features, screen lncRNA and build the model. And the test set is used as an internal validation to verify the prognostic effect of the model. Univariable Cox regression analysis of lncRNAs and survival time in the training population was performed using the Survival Package of R software [14]. LncRNAs with significant differences in expression ($p \leq 0.05$) were selected as target lncRNAs. Then Rbsurv package of R were used to screen the prognostic-related lncRNAs by Robust likelihood-based survival analysis [15].

The specific protocols are as follows: (1) Randomly divide 1090 breast cancer patients into training set (N*(1-*P*)) and test set (N**P*) (p = 1/3); (2) each target lncRNA is fitted to the univariable cox regression model of the training set, and obtained the corresponding parameter estimation, and then the log-likelihood of the parameter estimation of each target lncRNA is calculated in the test set; (3) repeating step (2) 10 times, ten log-likelihoods of one gene were obtained, and the mean values were calculated; (4) steps (2) and (3) were used to search for each target lncRNA, and the gene with the highest log-likelihood was selected as the optimal gene; (5) Subsequently, we searched the next best gene by evaluating every two-gene model and selected an optimal one with the largest mean log likelihood; (6) continued this forward gene selection procedure, resulting in a series of models. Akaike's information criterions (AICs) for all the candidate models were computed and an optimal model with the smallest AIC was selected finally. Using this model, prognostic-related lncRNAs were screened out.

2.3 Establishment and verification of risk assessment formula

All prognostic-related lncRNAs were included in the risk assessment formula and then weighted according to the estimated regression coefficients of the multivariate cox regression analysis of the training set. Use this formula to calculate the risk scores of patients in the training set, and the patients are divided into high risk group and low risk group according to the median risk score. The ROC (receiver operating characteristic) curve was plotted using survival ROC package of R and the optimal cut-off point was selected based on maximum sensitivity and specificity. The training set patients were re-

Table 1. Top 20 lncRNAs associated with the survival time of patients in the training set (N = 545).

lncRNA	HR	Cox.*p*.value
LINC00704	4.852630796	1.05E-05
AL022344.7	6.454117481	1.19E-05
RP11-344B5.2	1.856484282	4.79E-05
RP11-247C2.2	4.555215339	0.000145897
RP11-426L16.3	0.093780312	0.000192581
LSINCT5	7.786037821	0.000299324
LINC01451	13.275238	0.000393814
WARS2-IT1	2.054981475	0.000464166
RP11-129I19.2	1.900054628	0.000478849
STXBP5-AS1	10.79819638	0.000620234
AC002454.1	2.717891361	0.000736211
RP11-553A10.1	2.112912895	0.000788281
RP5-1073O3.2	8.61E-06	0.000935167
LINC00163	11.37727408	0.001096781
RP11-522I20.3	2.100554809	0.00114148
RP11-879F14.2	2.281790457	0.001449614
RP11-732A21.2	1788.258101	0.001569148
RP11-1260E13.3	6.108488937	0.001737004
RP11-316M21.6	0.053647998	0.001991235
RP11-598F7.3	0.174805023	0.002052405

grouped (high risk group and low risk group) according to the optimal cut-off point and evaluated by Kaplan Meier method and compared by log-rank test. The risk assessment formula is then validated in the test set and the entire data set.

3. Result

3.1 Screening for prognostic-related lncRNA

Data pertaining to the expression of 14371 lncRNAs from 1090 patients were obtained from the TCGA database. We further screened the expression of 7272 lncRNAs in breast cancer patients. See **Supplementary Table 3** for screening criteria. Subsequently, these 1090 patients were randomly divided into a training set (**Supplementary Table 4**) and a test set (**Supplementary Table 5**). Multivariate Cox regression analysis of the training set identified 326 lncRNAs with significant differences in expression ($p < 0.05$) as target lncRNAs. The top 20 lncRNAs where the p value was the smallest were selected and listed in Table 1. Robust likelihood-based survival analysis of target lncRNAs screened 19 prognostic-related lncRNAs (Table 2).

3.2 Risk prediction model for breast cancer prognosis based on 19 lncRNAs

To comprehensively study the relationship between the above 19 lncRNAs and breast cancer prognosis, we developed a risk prediction formula based on these 19 lncRNAs according to the cox coefficient. Risk score = (-0.401903149297177* BHLHE40-AS1) + (-0.552395828746267* TRG-AS1) + (0.353788296439561* RP11-10J5.1) + (3.34553443422715* LSINCT5) + (-0.184362518238646* CTA-384D8.31) + (-4.07228858077052* RP5-1073O3.2) + (0.354513130938573*

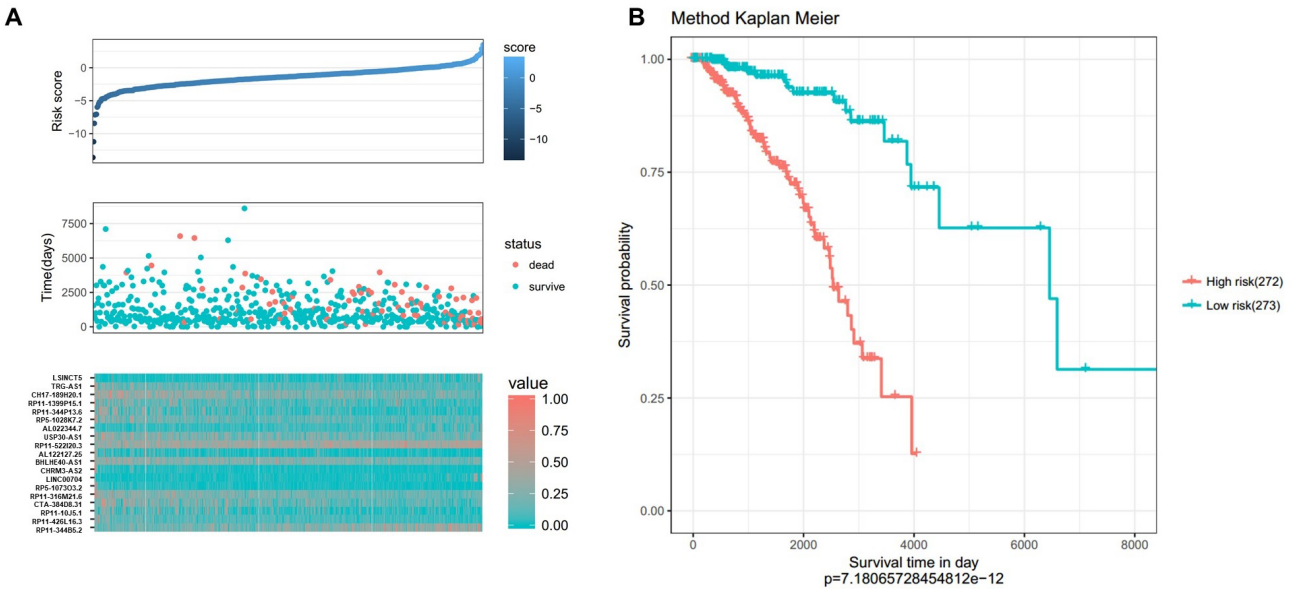

Fig. 1. LncRNA risk score analysis of the training set (N = 545). (A) The distribution of 19-lncRNA based risk core, patients' survival and heat-map of the lncRNA expression signature. Rows represent lncRNAs, and columns represent patients. (B) Kaplan-Meier estimates of patients' survival status and time using the median lncRNA risk score cut-off which divided patients into low-risk and high-risk groups.

Table 2. Prognosis related lncRNA signature screened using forward selection in the training set (N = 545).

lncRNA	nloglik	AIC
BHLHE40-AS1	380.47	762.94*
TRG-AS1	378.21	760.42*
RP11-10J5.1	377.48	760.95*
LSINCT5	372.75	753.51*
CTA-384D8.31	370.31	750.63*
RP5-1073O3.2	366.61	745.21*
RP11-522I20.3	364.22	742.43*
RP5-1028K7.2	364.2	744.41*
RP11-344B5.2	359.65	737.29*
AL122127.25	359.43	738.85*
RP11-1399P15.1	358.3	738.61*
CHRM3-AS2	357.7	739.39*
USP30-AS1	357.69	741.37*
RP11-316M21.6	354.99	737.97*
LINC00704	348.85	727.7*
RP11-426L16.3	347.33	726.66*
AL022344.7	344.1	722.2*
RP11-344P13.6	340.54	717.09*
CH17-189H20.1	332.51	703.02*

RP11-522I20.3) + (0.478673660175568* RP5-1028K7.2) + (0.548011782590826* RP11-344B5.2) + (-0.226325070280744* AL122127.25) + (-0.199054150971782* RP11-1399P15.1) + (-2.25757571437516* CHRM3-AS2) + (0.269133951150645* USP30-AS1) + (-3.36879822214685* RP11-316M21.6) + (1.61141284580778* LINC00704) + (-1.15282439300655* RP11-426L16.3) + (1.62700973053501* AL022344.7) + (-0.831276986176053* RP11-344P13.6) + (-0.773367909620609* CH17-189H20.1)

We then calculated the risk score for each patient in the training set, sorted them by risk score, and divided them into high-risk groups (n = 272) and low-risk groups (n = 273). Fig. 1A shows that the survival time of breast cancer patients was negatively correlated with the risk value (PCC = -1.4783). Most patients who died had higher risk scores, while patients with better prognoses had lower risk scores. LINC00704 and RP11-344B5.2 were highly expressed in the high risk group. Kaplan-Meier curves and log-rank tests revealed longer survival times of patients in the low-risk group compared with those in the high-risk group (Fig. 1B, p = 7.18e-12). The p value of the log-rank test further indicates that the prognostic ability of the risk formula is stronger than the predictive ability of any lncRNAs that were identified alone.

3.3 Effective indicators of breast cancer prognosis

We performed ROC analysis to evaluate the sensitivity and specificity of the risk formula. Fig. 2A shows that most of the cut-off points were well classified, and the ROC AUC was 0.83. The optimal cut-off point was -0.075, at which point the sensitivity and specificity were optimal. According to this cut-off point, breast cancer patients can be reclassified: high-risk group (n = 92) and low-risk group (n = 453). Kaplan-Meier curves and log-rank test analyses of these two groups showed significant differences in survival time between the high-risk group and the low-risk group (Fig. 2B, p = 8.70e-18, $p < 0.0001$).

3.4 Effectiveness of breast cancer prognosis risk prediction model

Subsequently, we verified the sensitivity and specificity of the risk formula for 19 lncRNAs in the entire data set and test set. Using the optimal cut-off point of the risk formula,

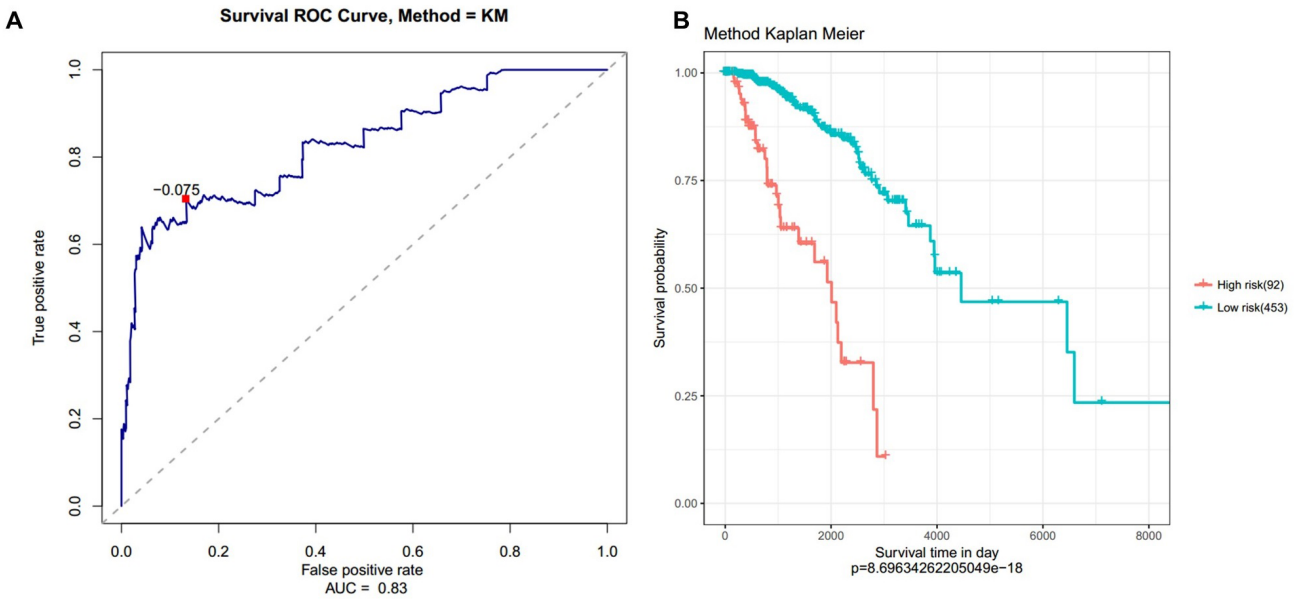

Fig. 2. (A) Receiver operating characteristic (ROC) analysis of the sensitivity and specificity of the survival time by the 19-lncRNA signature based risk score. (B) Kaplan-Meier estimates of the survival time of patients from the training set using the 19-lncRNA signature based risk score.

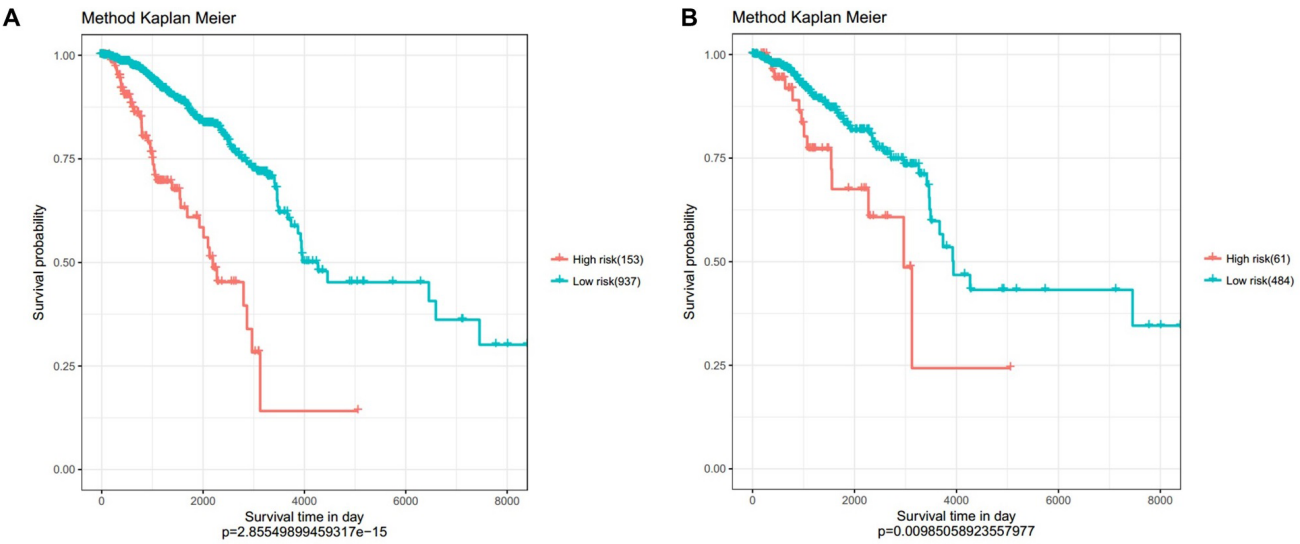

Fig. 3. (A) Kaplan-Meier estimates of the survival time of patients from the complete set using the 19-lncRNA signature based risk score. (B) Kaplan-Meier estimates of the survival time of patients from the testing set using the 19-lncRNA signature based risk score.

breast cancer patients in the entire data set were divided into high-risk groups (n = 153) and low-risk groups (n = 937), and survival time differences amongst the high risk group and low-risk group were determined (Fig. 3A, $p < 0.0001$). Similarly, patients in the test set were classified according to the optimal cut-off point of the risk formula: 61 breast cancer patients were divided into high-risk groups, and 464 breast cancer patients were divided into low-risk groups. Survival analysis showed significant differences in survival between the high-risk group and the low-risk group (Fig. 3B, $p < 0.01$).

4. Discussion

In this study, 19 breast cancer prognostic-related lncRNAs were identified from the training set, namely: LSINCT5, TRG-AS1, CH17-189H20.1, RP11-1399P15.1, RP11-344P13.6, RP5-1028K7.2, AL022344.7, USP30-AS1, RP11-522I20.3, AL122127.25, BHLHE40-AS1, CHRM3-AS2, LINC00704, RP5-1073O3.2, RP11-316M21.6, CTA-384D8.31, RP11-10J5.1, RP11-426L16.3, and RP11-344B5.2. Based on the cox coefficient of multivariate regression analysis, we developed a prognostic risk formula for breast cancer. In addition, we obtained the optimal cut-off point by ROC analysis and re-grouped the training set data. Subsequently, we verified the risk formula and the

optimal cut-off point in the test set as well as the entire data set. It was found that patients in the high-risk group had shorter survival days and a higher mortality rate. RP11-10J5.1, RP11-522I20.3, LSINCT5, RP5-1028K7.2, RP11-344B5.2, USP30-AS1, LINC00704 and AL022344.7 are highly expressed in high-risk breast cancer patients with a poor prognosis and BHLHE40-AS1, TRG-AS1, CTA-384D8.31, RP5-1073O3.2, AL122127.25, CHRM3-AS2, RP11-316M21.6, RP11-426L16.3, RP11-344P13.6, RP11-1399P15.1 and CH17-189H20.1 are highly expressed in low-risk breast cancer patients with a good prognosis. These results demonstrated that lncRNAs could be prognostic markers of breast cancer.

In recent years, studies have found that lncRNA has a variety of biological activities, involved in X chromosome silencing, chromosome modification, transcriptional activation and inhibition, and RNA cleavage, and plays important roles in cell differentiation, embryonic and tissue development, and tumor progression [16–18]. But the numbers of lncRNA are many, and their biological functions are not yet fully understood. This study identified 19 lncRNAs and only 3 lncRNAs - LINC00704 (also known as MANCR), AL022344.7, and RP11-344B5.2 (also known as lnc-PTPA-3) have been previously reported, the other 16 lncRNAs have not been reported before. This is the first time these lncRNAs have been found to be expressed in breast cancer patients, suggesting that these lncRNAs may be associated with a poor prognosis in breast cancer patients. But this discovery requires further experimental and mechanistic validation. LINC00704 has been reported to be associated with recurrence and metastasis of breast cancer patients and can be used as a biomarker for predicting recurrence in breast cancer patients [19]. LINC00704 was found to be upregulated in breast cancer specimens and cells. Depletion of MANCR significantly inhibits triple-negative breast cancer cell proliferation with a concomitant increases in DNA damage and an increase in the incidences of cytokinesis and cell apoptosis by affecting the expression of $>$2000 genes that are enriched in cell-cycle regulation pathways [20]. In addition, Lu W *et al.* (2018) found that LINC00704 was upregulated in thyroid cancer, and that the shorter Overall Survival (OS) time in thyroid cancer patients was associated with higher expression levels. Meanwhile knockdown of LINC00704 significantly impaired proliferation and colony formation capacity of thyroid cancer cells and induced cell G1/G0 phase arrest and cell apoptosis, and inhibited a cells invasive ability in thyroid cancer [21]. These findings indicate that LINC00704 might play important roles in cancer tumorigenesis and progression. AL022344.7 has been reported to be upregulated in NSCLC cell lines due to promoter hypomethylation [22]. RP11-344B5.2 has not been reported before, and this is the first time it has been found to be highly expressed in breast cancer patients.

Based on the identified lncRNAs, we developed a risk model for breast cancer prognosis prediction. When applying it to the training set, we found significant differences in survival curves between high-risk breast cancer patients and low-risk breast cancer patients. Patients in the high-risk breast cancer group had a poor prognosis and a shorter survival time, while patients in the low-risk breast cancer group had a longer survival time. This study aims to develop a risk model for predicting survival time in breast cancer patients. The results of this study further emphasizes the importance of lncRNA-based risk scores in cancer prognosis [23–25]. We optimized the risk classification of breast cancer patients by identifying the optimal cut-off, and we obtained a more significant difference in survival analysis. We validated the model in the entire data set and test set, which indicates that the model has great repeatability and stability. The limitation of this paper is that it did not study the causal relationship between lncRNAs and breast cancer patients; however, this breast cancer prognosis prediction model can be used for scientific research and subsequent clinical application.

HOTAIR and FSIP1 are two of the early identified lncRNAs and play significant roles in gene. However, in the unicox of this study, the *p*-values of HOTAIR and FSIP1 are 0.80 and 0.92, respectively, which is not suitable for multi-factor Cox regression analysis of this dataset. Therefore, this model did not include in HOTAIR and FSIP1. Alternatively, the signatures identified in this study may be more potentially as diagnostic markers than HOTAIR. The test data of this model comes from the breast cancer tissue of TCGA. Because the expression of lncRNAs has tissue specificity, it is impossible to evaluate the expression of lncRNA in blood. If it is applied to clinic, the model based on blood samples is more meaningful. The significance of this study is to provide method guidance for model construction in blood.

5. Conclusions

We identified 19 prognostic related lncRNAs of breast cancer patients. The target genes and related biological functions of these lncRNAs provide a basis for studying the occurrence and development of breast cancer. In addition, we also determined that the risk formula based on 19-lncRNAs expression can predict the survival time of breast cancer patients, which is of great significance for prognosis prediction and patient management of breast cancer patients.

Author contributions

GJH and CGL conceived and designed the study; XPY, SJH and YH contributed materials and analyzed the data; XPY, FQ and FX wrote the paper. All authors have read and approved the manuscript.

Ethics approval and consent to participate

Not applicable.

Acknowledgment

We would like to express our gratitude to all colleague who helped us during the writing of this manuscript. The authors also thank the peer reviewers for careful review of the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (Grants 81102029 and 81872159 to Caigang Liu).

Conflict of interest

The authors declare no conflict of interest.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at https://ejgo.imrpress.com/EN/10.31083/j.ejgo4205128.

References

[1] Tong CWS, Wu M, Cho WCS, To KKW. Recent advances in the treatment of breast cancer. Frontiers in Oncology. 2018; 8: 227.

[2] Brewster AM, Hortobagyi GN, Broglio KR, Kau SW, Santa-Maria CA, Arun B, *et al.* Residual risk of breast cancer recurrence 5 years after adjuvant therapy. Journal of the National Cancer Institute. 2008; 100: 1179–1183.

[3] Liu C, Jiang Y, Gu X, Xu Z, Ai L, Zhang H, *et al.* Predicting level 2 axillary lymph node metastasis in a Chinese breast cancer population post-neoadjuvant chemotherapy: development and assessment of a new predictive nomogram. Oncotarget. 2017; 8: 79147–79156.

[4] Arun G, Diermeier SD, Spector DL. Therapeutic targeting of long non-coding RNAs in cancer. Trends in Molecular Medicine. 2018; 24: 257–277.

[5] Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. Nature Reviews Genetics. 2016; 17: 47–62.

[6] Latgé G, Poulet C, Bours V, Josse C, Jerusalem G. Natural antisense transcripts: Molecular mechanisms and implications in breast cancers. International Journal of Molecular Sciences. 2018; 19: 123.

[7] Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigó R, Johnson R. Towards a complete map of the human long non-coding RNA transcriptome. Nature Reviews Genetics. 2018; 19: 535–548.

[8] Gutschner T, Richtig G, Haemmerle M, Pichler M. From biomarkers to therapeutic targets-the promises and perils of long non-coding RNAs in cancer. Cancer and Metastasis Reviews. 2018; 37: 83–105.

[9] Reiche K, Kasack K, Schreiber S, Lüders T, Due EU, Naume B, *et al.* Long non-coding RNAs differentially expressed between normal versus primary breast tumor tissues disclose converse changes to breast cancer-related protein-coding genes. PLoS ONE. 2014; 9: e106076.

[10] Ding X, Zhu L, Ji T, Zhang X, Wang F, Gan S, *et al.* Long intergenic non-coding RNAs (LincRNAs) identified by RNA-seq in breast cancer. PLoS ONE. 2014; 9: e103270.

[11] Sørensen KP, Thomassen M, Tan Q, Bak M, Cold S, Burton M, *et al.* Long non-coding RNA HOTAIR is an independent prognostic marker of metastasis in estrogen receptor-positive primary breast cancer. Breast Cancer Research and Treatment. 2013; 142: 529–536.

[12] Hu P, Chu J, Wu Y, Sun L, Lv X, Zhu Y, *et al.* NBAT1 suppresses breast cancer metastasis by regulating DKK1 via PRC2. Oncotarget. 2015; 6: 32410–32425.

[13] Xu S, Wang P, You Z, Meng H, Mu G, Bai X, *et al.* The long noncoding RNA EPB41L4a-as2 inhibits tumor proliferation and is associated with favorable prognoses in breast cancer and other solid tumors. Oncotarget. 2016; 7: 20704–20717.

[14] O'Quigley J, Moreau T. Cox's regression model: computing a goodness of fit statistic. Computer Methods and Programs in Biomedicine. 1986; 22: 253–256.

[15] Renaud G, Stenzel U, Maricic T, Wiebe V, Kelso J. DeML: robust demultiplexing of Illumina sequences using a likelihood-based approach. Bioinformatics. 2015; 31: 770–772.

[16] Ge Y, Yan X, Jin Y, Yang X, Yu X, Zhou L, *et al.* fMiRNA-192 and miRNA-204 directly suppress lncRNA HOTTIP and interrupt GLS1-mediated glutaminolysis in hepatocellular carcinoma. PLOS Genetics. 2015; 11: e1005726.

[17] Mehta SL, Kim T, Vemuganti R. Long noncoding RNA FosDT promotes ischemic brain injury by interacting with REST-associated chromatin-modifying proteins. Journal of Neuroscience. 2015; 35: 16443–16449.

[18] Sahu A, Singhal U, Chinnaiyan AM. Long noncoding RNAs in cancer: from function to translation. Trends in Cancer. 2015; 1: 93–109.

[19] Liu H, Li J, Koirala P, Ding X, Chen B, Wang Y, *et al.* Long noncoding RNAs as prognostic markers in human breast cancer. Oncotarget. 2016; 7: 20584–20596.

[20] Tracy KM, Tye CE, Ghule PN, Malaby HLH, Stumpff J, Stein JL, *et al.* Mitotically-associated lncRNA (MANCR) affects genomic stability and cell division in aggressive breast cancer. Molecular Cancer Research. 2018; 16: 587–598.

[21] Lu W, Xu Y, Xu J, Wang Z, Ye G. Identification of differential expressed lncRNAs in human thyroid cancer by a genome-wide analyses. Cancer Medicine. 2018; 7: 3935–3944.

[22] Horie M, Kaczkowski B, Ohshima M, Matsuzaki H, Noguchi S, Mikami Y, *et al.* Integrative CAGE and DNA methylation profiling identify epigenetically regulated genes in NSCLC. Molecular Cancer Research. 2017; 15: 1354–1365.

[23] Li J, Chen Z, Tian L, Zhou C, He MY, Gao Y, *et al.* LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut. 2014; 63: 1700–1710.

[24] Iguchi T, Uchi R, Nambara S, Saito T, Komatsu H, Hirata H, *et al.* A long noncoding RNA, lncRNA-ATB, is involved in the progression and prognosis of colorectal cancer. Anticancer Research. 2015; 35: 1385.

[25] Zhou M, Zhao H, Wang Z, Cheng L, Yang L, Shi H, *et al.* Identification and validation of potential prognostic lncRNA biomarkers for predicting survival in patients with multiple myeloma. Journal of Experimental & Clinical Cancer Research. 2015; 34: 102.