

Exploring Long Non-coding RNA Potential Markers for Prognosis Risk based on Double Hormone Receptor Positive Breast Cancer CeRNA Network

Qi Huang (✉ huangqi214@163.com)

Second Affiliated Hospital of Anhui Medical University <https://orcid.org/0000-0002-0261-4562>

Yao Peng

Second Affiliated Hospital of Anhui Medical University

Jingjing Wu

Second Affiliated Hospital of Anhui Medical University

Na Li

Second Affiliated Hospital of Anhui Medical University

Research Article

Keywords: Long non-coding RNA, breast cancer, ceRNA network, biomarker, prognosis.

Posted Date: July 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-663806/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: The aim of this study is to search for key long non-coding RNA (lncRNAs) able to influence the prognosis of double hormone receptor positive (ER+PR+) breast cancer by constructing a competing endogenous RNA (ceRNA) topology network of RNA-seq data and miRNA-seq data.

Materials and methods: Based on The Cancer Genome Atlas (TCGA), we downloaded RNA-seq data and miRNA-seq data of ER+PR+ breast cancer and obtained lncRNA-miRNA relationship pairs by comparing with the data in moRcode database. MiRDB, miRTarBase and TargetScan databases were used to predict target genes, and thereby gaining miRNA-mRNA relationship pairs. The ceRNA network was constructed using Cytoscape, then functional enrichment analysis was performed. Cox survival analysis was carried out to explore the potential prognostic lncRNAs.

Results: A ceRNA topology network was constructed and six lncRNAs, PSORS1C3, AC010536.2, IQCF5-AS1, AL021707.1, FGD5-AS1, and PART1, were found to be mostly associated with the survival of ER+PR+ breast cancer.

Conclusion: Six lncRNAs with prognostic significance were identified in ER+PR+ breast cancer, providing new ideas for the diagnose and treatment of this subtype.

Introduction

Breast cancer (BC) is one of the most common malignant tumors in women and consists of highly heterogeneous subtypes(1–3). According to statistics, the incidence rate is 8–12% among all body malignant tumors. Its incidence is considered much related to heredity and is higher among women aged 40–60 and perimenopause(4, 5). Breast cancer is defined by distinct steroid hormone receptors (HR): estrogen receptors (ER) and progesterone receptors (PR) status. HR status has been shown to be a significant prognostic factor and predictive marker in response to endocrine therapy during breast cancer treatment(6, 7). The double hormone receptor positive (ER + PR+) breast cancer accounts for the vast majority of breast cancer and is known to have a favorable prognosis and good response to endocrine therapy compared with the double negative phenotype(8, 9). With the continuous development of science and technology, great progress has been made in the diagnosis and treatment of breast cancer, the survival rate of patients has also been greatly improved(10, 11). However, there are still a small number of ER + PR + breast cancer patients who are less sensitive to endocrine therapy leading to a poor prognosis, the reason still remains unknown. Recently, more and more researches have focused on the study of tumor biomarkers. Biomarkers with predictive and diagnostic significance have provided new clinical ideas for diagnosing and treating this breast cancer phenotype(12–14).

For the past few years, the abnormal expression of protein products during transcription are attracting widespread attention, long non-coding RNA (lncRNA) has gradually become one of the hot issues in cancer research(15). LncRNA is one of the two categories of noncoding RNAs and is longer than 200 nucleotides in length(16). Another category is microRNA (miRNA) with small or short RNA length that can

induce mRNA degradation or inhibit translation to silence target gene expression(17). With the rapid development of whole genome and transcriptome sequencing technologies, increasing evidence demonstrates that lncRNAs play significant regulatory roles in various biological processes, including cancer progression(18–20). According to recent studies, lncRNAs form a competitive endogenous RNA (ceRNA) through sponge adsorption of miRNA to regulate mRNA expression(21). The ceRNA hypothesis believes that lncRNAs and other transcripts can competitively bind miRNAs to prevent them from binding to mRNA thereby regulate mRNA expression(22).

At present, the mechanism of ceRNA interactions in double hormone receptor positive breast cancer is not clear yet. Exploring the interaction relationship is of great significance for elucidating the mechanism of breast cancer development.

Materials And Methods

Data download and preprocessing

We downloaded breast cancer RNA-seq data and miRNA-seq data from the TCGA database. The RNA-seq data contained 113 normal samples and 1102 cancer samples. Likewise, the miRNA-seq data contained 104 normal samples and 1096 cancer samples. Among these samples, we filtered out 12 male samples and then selected out data with both RNA-seq and miRNA-seq. Both estrogen and progesterone receptors should be positive in cancer samples, thus we got the final 753 samples, including 102 normal samples and 651 cancer samples. Then we extracted the mRNA expression profile and the lncRNA expression profile from the matrix files obtained from the RNA-seq data. Since many miRNAs express 0 in most samples and these miRNAs may affect our follow-up relationship construction, we removed miRNAs with a ratio of 0 to 75% in all samples. Therefore, we finally obtained three matrix files of mRNA, lncRNA and miRNA expression profiles and ready for downstream analysis.

Forecast follow-up relationship

MiRcode is a database website to predict miRNA targets based on a complete GENCODE annotation transcriptome of human. Through this database prediction, we got the lncRNA-miRNA relationship files. MiRDB is a miRNA target gene prediction database based on high-throughput sequencing experiments. MiRTarBase is a miRNA targets database validated by experiments. TargetScanS predicts miRNA target genes by searching for conserved 8mer and 7mer sites that match the miRNA seed sequence. These three databases are currently more used to predict miRNA target genes. We used miRDB, MiRTarBase and TargetScan these three databases to predict the target genes of our miRNAs, and defined the genes predicted concurrently by these three databases as our final target genes. In this way, we got the miRNA-mRNA relationship files.

Screening relationship

Based on the ceRNA hypothesis, lncRNAs can competitively bind miRNAs to prevent them from binding to mRNA in the model of ceRNA. Therefore, the expression level of lncRNA that can theoretically exert

ceRNA function is inversely related to the expression level of miRNA. The mRNA expression level in the ceRNA relationship is also negatively correlated with the miRNA expression. We calculated the Pearson correlation coefficient of lncRNA and miRNA in lncRNA-miRNA relationship pairs, and screened out the pairs with significantly negative correlation ($P < 0.05$). Similarly, we calculated the Pearson correlation coefficient and its P value of miRNA and mRNA in miRNA-mRNA relationship pairs, and then screened out significant negative correlation ($P < 0.05$) relationship pairs. These pairs acted as preselected ceRNA regulatory relationship pairs.

Building a ceRNA network

The ceRNA network is a lncRNAs and mRNAs network using miRNA as a medium. In this network, each miRNA must be linked to lncRNA and mRNA simultaneously. That is to say, in our relationship files, only the lncRNA-miRNA relationship files or the miRNA-mRNA relationship files exist. Thus, we further filtered out the incompatible relationships between the two relationship pairs, combined with the screened lncRNA-miRNA and miRNA-mRNA pairs, visualized the topological network with Cytoscape software, and finally obtained a topological network diagram of ceRNAs.

Analysis of mRNA enrichment in ceRNA networks

To observe the function of the constructed ceRNA network, we selected all the mRNAs in the network and used the R package clusterProfiler⁸ to perform functional enrichment analysis.

CeRNA network topology analysis and stability analysis

The network topology quality is analyzed by using the plug-in NetworkAnalyzer tool kit in Cytoscape. NetworkAnalyzer network analysis mainly includes graphs of network diameter, number of connections, average aggregation coefficient, and so on. In this study, we calculated the number of connections, the length of the paths, and the proximity of the nodes.

Prognostic analysis of ceRNA modules

We downloaded the clinical information of the above samples from the TCGA website and extracted the survival data of each sample. Then, combined with the expression data obtained before, Cox survival analysis of lncRNA nodes in the network was carried out to explore the potential lncRNA prognostic markers.

lncRNA prognostic model construction

We used the lncRNA with prognostic value to further construct a prognostic model, and observed the impact of its classification on the prognosis.

Results

Relationship construction and screening

We have a total of 14370 expression profiles of lncRNAs. Using the miRcode database, we obtained 7098 lncRNA-miRNA pairs containing 1449 lncRNAs and 35 miRNAs. Also, we had 2253 expression profiles of miRNAs and removed samples with a ratio of 0 to 75%, 875 miRNAs were left. Using miRDB, miRTarBase and TargetScan databases, we performed target gene prediction on our miRNAs and got 19571 relationship pairs. Among the three databases, 5311 genes and 780 miRNAs were predicted as target genes at the same time.

Based on the ceRNA hypothesis, lncRNAs can competitively bind miRNAs to prevent them from binding to mRNA in the model of ceRNA. Therefore, the expression level of lncRNA that can theoretically exert ceRNA function is inversely related to the expression level of miRNA. The mRNA expression level in the ceRNA relationship is also negatively correlated with the miRNA expression. We calculated the Pearson correlation coefficients of lncRNAs and miRNAs in lncRNA-miRNA relationship pairs, and screened out 240 pairs with significantly negative correlations ($P < 0.05$). Similarly, we calculated the Pearson correlation coefficients and their P-values of miRNAs and mRNAs in miRNA-mRNA relationship pairs, and screened out 2883 relationship pairs with significant negative correlations ($P < 0.05$).

The ceRNA network is a lncRNAs and mRNAs network using miRNA as a medium. In this network, each miRNA must be linked to lncRNA and mRNA simultaneously. That is to say, in our relationship files, only the lncRNA-miRNA relationship files or the miRNA-mRNA relationship files exist. We further filtered out the incompliant relationships in the two relationship pairs, combined with filtered lncRNA-miRNA and miRNA-mRNA pairs, finally we got 227 lncRNA-miRNA relationship pairs and 171 miRNA-mRNA relationship pairs. Some of the relationships are shown in Table 1–2.

Table 1
Several of the lncRNA–miRNA pairs obtained from this analysis.

lncRNA	miRNA
FAM182A	hsa-miR-139-5p
FAM182A	hsa-miR-17-5p
FAM182A	hsa-miR-20b-5p
FAM182A	hsa-miR-107
PART1	hsa-miR-22-3p
PART1	hsa-miR-24-3p
PART1	hsa-miR-125a-5p
PART1	hsa-miR-129-5p
SMAD5-AS1	hsa-miR-455-5p
AC068020.1	hsa-miR-22-3p

Table 2
Several of the miRNA–mRNA pairs
obtained from this analysis.

miRNA	mRNA
hsa-miR-24-3p	TAOK1
hsa-miR-139-5p	STAMBP
hsa-miR-17-5p	PAFAH1B1
hsa-miR-301b-3p	CDADC1
hsa-miR-24-3p	PDXK
hsa-miR-24-3p	CNNM3
hsa-miR-17-5p	LDLR
hsa-miR-125a-5p	ESRRA
hsa-miR-17-5p	RAPGEF4
hsa-miR-129-5p	ETV6

CeRNA network construction and visualization

Combined with the 227 pairs of lncRNA-miRNA and 171 pairs of mRNA-miRNA obtained above, we used Cytoscape software to visualize the network and, and finally got a ceRNA network diagram. There were 398 edges and 326 nodes in this network, including 146 lncRNAs, 20 miRNAs, and 160 mRNAs (Fig. 1).

Topology analysis and stability analysis of ceRNA networks

For the topology analysis of the constructed ceRNA network, Fig. 2A showed the distribution of node degrees. The degrees of most nodes are very low, and only a small part of the nodes is high, which suggests that those nodes with higher degrees act as hubs for the entire network.

The closeness (Centrality, CC) of a node can calculate the number of connection steps between a given node and other nodes. A more concentrated node will get a higher score, so CC means the shortest path. As shown in Fig. 2B, nodes with the relatively lower number of connections are relatively concentrated, and some points with higher degrees of connectivity are relatively scattered.

The path can reflect the combination of all nodes in the network. Figure 2C showed the shortest path length distribution of the network. It can be seen in the figure that the distribution of the path length is more concentrated, with fewer extremes at both ends, an upper limit of 10, and a lower limit of 1, indicating that most of the nodes in the network can be connected by short path correlation.

Figure 2D showed the degree distribution density graph of nodes. We find that as node degrees increase, the density of nodes decreases sharply, suggesting that most of the nodes in the network show isolation.

In the process of disease, it may be due to changes of several key nodes, and then interact with neighboring nodes to induce coexpression, and finally affect the downstream biological processes.

MRNA enrichment analysis (GO and KEGG)

GO (Gene Ontology) analysis shows that most of the enriched biological processes (BP) are all associated with the cell cycle, suggesting that changes in the cell cycle and regulation are involved in the development of breast cancer (Fig. 3A).

MAPK signaling pathway ranked first in the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis. Studies have shown that MAPK signaling pathway mainly increases the expression of cyclin D1, promotes the transition from cell cycle G1 to S phase and accelerates cell division and proliferation(23, 24). It can also inhibit the expression of cell cycle arrest proteins, release cell cycle arrest and promote cell division, which is consistent with the results of our GO analysis. In addition, bladder cancer and pancreatic cancer also appeared in other items, suggesting that mRNAs are closely related to cancer (Fig. 3B).

Network node survival analysis and potential markers of lncRNA

We performed Cox survival analysis on 146 lncRNAs in the network, and finally screened 6 lncRNAs with significant prognostic significance. The result was shown in Table 3. Next, we carried out univariate regression analysis on 6 prognostically differentiated lncRNAs. A forest map was shown in Fig. 4, the high expression of 6 prognostically significant lncRNAs was associated with unfavorable prognosis.

Table 3
Prognosis-related lncRNAs in ceRNA networks.

LncRNA	P value	HR	Low 95%CI	High 95%CI
PSORS1C3	0.0495213	0.10516734	0.01111244	0.99529603
AC010536.2	0.03919404	0.26422563	0.07457465	0.93617851
FGD5-AS1	0.01875322	1.0110719	1.00183077	1.02039826
IQCF5-AS1	0.03728005	0	0	5.41E-54
PART1	0.04086366	1.18655672	1.00713105	1.39794801
AL021707.1	0.02160282	0.22725254	0.06419527	0.80447853

The prognosis model of lncRNA

Further, we used these 6 lncRNA expression profiles for multivariate Cox regression to construct a prognostic model. Risk Score = $-2.67 \times \text{PSORS1C3} - 0.71 \times \text{AC010536.2} + 0.01 \times \text{FGD5-AS1} - 1.46 \times \text{IQCF5-AS1} + 0.18 \times \text{PART1} - 0.89 \times \text{AL021707.1}$. We substituted each lncRNA expression profile into the model to

calculate the risk coefficient of each sample, and classified the samples according to the median risk factor. It can be seen from the figure that the number of deaths in high-risk group was significantly higher than that in the low-risk group, and the higher the risk coefficient, the lower the expression of the lncRNAs in PSORS1C3, AC010536.2, IQCF5-AS1, AL021707.1, FGD5-AS1, and PART1 (Fig. 5).

To observe the prognostic differences between the high-risk and low-risk samples of the model, we performed survival analysis and plotted K-M curves for the high-risk group and the low-risk group (Fig. 6A). There are significant prognostic differences between the two groups of samples. To further observe the heterogeneity and stability of the model, we analyzed the ROC curve of the model (Fig. 6B). The prognostic model constructed by the 6 lncRNAs has a high AUC area and can significantly distinguish the risk of prognosis, suggesting that the 6 lncRNAs can be served as good prognostic markers.

Discussion

Recently, ceRNA hypothesis has been demonstrated to represent a novel model of post-transcriptional regulation working through miRNA competition(25). In the ceRNA network, ceRNA can inhibit multiple miRNAs and regulate the network by changing the type and number of miRNA binding sites, involving various signaling pathways and biological phenomena (22, 25). Further study on the mechanism of ceRNA will help to deepen the understanding of the pathogenesis of the disease, and has a good application prospect in disease treatment.

With the increasing attention to lncRNA genes and related researches, more and more new lncRNAs have been found to be regulators in ceRNA gene expression network. Although a large number of studies have shown that the expression of lncRNAs plays an important role in the development and progression of tumor, including cell growth, apoptosis, cell migration and invasion, as well as cancer cell stemness, the precise molecular mechanism is still not well studied(26–30). Generally, lncRNAs can be identified as both carcinogenic regulators and tumor suppressors, depending on the complicated context of tumor(31–33). One mechanism of lncRNAs has been reported to be a competitive endogenous RNA (ceRNA). It has been reported that lncRNAs can regulate miRNAs expression by sponging miRNAs and further mediate their downstream genes(34–36). For example, lncRNA GACAT3 can predict poor prognosis and promote cell proliferation by miR-497/CCND2 axis in breast cancer(37). lncRNA H19 and HOTAIR sponge miR-200 and miR-20a-5p respectively to promote breast cancer development(38, 39). A functional analysis of the ceRNA networks in four subtypes of breast cancer (basal-like, HER2+, luminal A and luminal B) shows that three common lncRNAs, NEAT1, OPI5-AS1 and AC008124.1, play specific roles in each subtype by competing with diverse mRNAs(40).

This article analyzes the RNA-seq and miRNA-seq data of double hormone receptor positive breast cancer in TCGA public database, constructs a ceRNA network, analyzes the network topology, and analyzes the survival of the intrinsic nodes in the network. Finally, six lncRNA markers that may influence the prognosis, PSORS1C3, AC010536.2, IQCF5-AS1, AL021707.1, FGD5-AS1, and PART1, are found. Several

analyses have revealed FGD5-AS1 as a potential regulator of clear cell renal cell carcinoma progression(41), and another integrated analysis of lncRNA-associated ceRNA networks in periodontitis found that FGD5-AS1 targets both miR-125a-3p and miR-142-3p(42). PART1 expression can effectively predict the early recurrence of hepatocellular carcinoma after surgical resection(43). PART1 is considered as a new target for the treatment of prostate cancer by inhibiting toll-like receptor pathways to promote cell proliferation and apoptosis(44). PART1 may also be a promising biomarker in stage I-III non-small cell lung cancer since its expression is associated with poor prognosis and tumor recurrence(45). Additionally, the expression of PART1 is related to OS in oral squamous cell carcinoma(46). The rest lncRNAs are not well studied yet.

In conclusion, our research strategy using a combination of bioinformatics and multiple databases provides a new methodological reference for the study of tumor mechanism. By targeting the ceRNA regulatory network, lncRNA and other key molecules have been identified as novel prognostic markers that will help to select appropriate therapies to improve the survival outcomes of ER + PR + breast cancer patients. However, there are still some shortcomings in this article. Due to the mismatch in the amount of data between normal samples and cancer samples in TCGA database, our results need to be further verified by extended sample and cellular and molecular experiments.

Declarations

Funding

This work was supported by National Natural Science Foundation of China (grant numbers: 82003037) and Natural Science Foundation of Anhui Province of China (grant numbers: 2008085QH379, 2008085QH413).

Competing Interests

The authors declare that there are no conflicts of interest.

Availability of data

The data were downloaded from TCGA database.

Author Contributions

Qi Huang, Literature research, Data analysis, Manuscript preparation. Jingjing Wu, Data extraction. Yao Peng, Manuscript editing. Na Li, Study design.

Ethical approval

The studies involving human participants were reviewed and approved by all data are from public database on the internet.

Consent to participate

All authors agree to submit this article for publication.

Consent for publication

All authors agree with publication.

References

1. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, et al. Breast cancer statistics, 2019. *CA: a cancer journal for clinicians*. 2019;69(6):438-51.
2. Haynes B, Sarma A, Nangia-Makker P, Shekhar MP. Breast cancer complexity: implications of intratumoral heterogeneity in clinical management. *Cancer metastasis reviews*. 2017;36(3):547-55.
3. Januskeviciene I, Petrikaite V. Heterogeneity of breast cancer: The importance of interaction between different tumor cell populations. *Life Sci*. 2019;239:117009.
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*. 2018;68(6):394-424.
5. Feng RM, Zong YN, Cao SM, Xu RH. Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics? *Cancer Commun (Lond)*. 2019;39(1):22.
6. Lamb CA, Vanzulli SI, Lanari C. Hormone receptors in breast cancer: more than estrogen receptors. *Medicina (B Aires)*. 2019;79(Spec 6/1):540-5.
7. Li Y, Yang D, Yin X, Zhang X, Huang J, Wu Y, et al. Clinicopathological Characteristics and Breast Cancer-Specific Survival of Patients With Single Hormone Receptor-Positive Breast Cancer. *JAMA Netw Open*. 2020;3(1):e1918160.
8. Li Z, Tu Y, Wu Q, Wang Z, Li J, Zhang Y, et al. Clinical Characteristics and Outcomes of Single Versus Double Hormone Receptor-Positive Breast Cancer in 2 Large Databases. *Clin Breast Cancer*. 2020;20(2):e151-e63.
9. Lv M, Mao Y, Song Y, Wang Y, Liu X, Wang X, et al. Clinical Features and Survival of Single Hormone Receptor-Positive Breast Cancer: A Population-Based Study of 531,605 Patients. *Clin Breast Cancer*. 2020;20(5):e589-e99.
10. Waks AG, Winer EP. Breast Cancer Treatment: A Review. *JAMA*. 2019;321(3):288-300.
11. Shien T, Iwata H. Adjuvant and neoadjuvant therapy for breast cancer. *Jpn J Clin Oncol*. 2020;50(3):225-9.
12. Odle TG. Precision Medicine in Breast Cancer. *Radiol Technol*. 2017;88(4):401M-21M.
13. Kwon MJ. Emerging immune gene signatures as prognostic or predictive biomarkers in breast cancer. *Arch Pharm Res*. 2019;42(11):947-61.

14. Mueller C, Haymond A, Davis JB, Williams A, Espina V. Protein biomarkers for subtyping breast cancer and implications for future research. *Expert Rev Proteomics*. 2018;15(2):131-52.
15. Bhan A, Soleimani M, Mandal SS. Long Noncoding RNA and Cancer: A New Paradigm. *Cancer research*. 2017;77(15):3965-81.
16. Beermann J, Piccoli MT, Viereck J, Thum T. Non-coding RNAs in Development and Disease: Background, Mechanisms, and Therapeutic Approaches. *Physiol Rev*. 2016;96(4):1297-325.
17. Lu TX, Rothenberg ME. MicroRNA. *The Journal of allergy and clinical immunology*. 2018;141(4):1202-7.
18. Peng WX, Koirala P, Mo YY. LncRNA-mediated regulation of cell signaling in cancer. *Oncogene*. 2017;36(41):5661-7.
19. Chi Y, Wang D, Wang J, Yu W, Yang J. Long Non-Coding RNA in the Pathogenesis of Cancers. *Cells*. 2019;8(9).
20. Sanchez Calle A, Kawamura Y, Yamamoto Y, Takeshita F, Ochiya T. Emerging roles of long non-coding RNA in cancer. *Cancer Sci*. 2018;109(7):2093-100.
21. Chan JJ, Tay Y. Noncoding RNA:RNA Regulatory Networks in Cancer. *International journal of molecular sciences*. 2018;19(5).
22. Lin W, Liu H, Tang Y, Wei Y, Wei W, Zhang L, et al. The development and controversy of competitive endogenous RNA hypothesis in non-coding genes. *Molecular and cellular biochemistry*. 2021;476(1):109-23.
23. Fu H, Gao H, Qi X, Zhao L, Wu D, Bai Y, et al. Aldolase A promotes proliferation and G1/S transition via the EGFR/MAPK pathway in non-small cell lung cancer. *Cancer Commun (Lond)*. 2018;38(1):18.
24. Wang S, Wang X, Gao Y, Peng Y, Dong N, Xie Q, et al. RN181 is a tumour suppressor in gastric cancer by regulation of the ERK/MAPK-cyclin D1/CDK4 pathway. *The Journal of pathology*. 2019;248(2):204-16.
25. Smillie CL, Sirey T, Ponting CP. Complexities of post-transcriptional regulation and the modeling of ceRNA crosstalk. *Crit Rev Biochem Mol Biol*. 2018;53(3):231-45.
26. Tu C, Yang K, Wan L, He J, Qi L, Wang W, et al. The crosstalk between lncRNAs and the Hippo signalling pathway in cancer progression. *Cell Prolif*. 2020;53(9):e12887.
27. Tang Y, He Y, Zhang P, Wang J, Fan C, Yang L, et al. LncRNAs regulate the cytoskeleton and related Rho/ROCK signaling in cancer metastasis. *Molecular cancer*. 2018;17(1):77.
28. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. *Cancer cell*. 2016;29(4):452-63.
29. Castro-Oropeza R, Melendez-Zajgla J, Maldonado V, Vazquez-Santillan K. The emerging role of lncRNAs in the regulation of cancer stem cells. *Cell Oncol (Dordr)*. 2018;41(6):585-603.
30. Huo X, Han S, Wu G, Latchoumanin O, Zhou G, Hebbard L, et al. Dysregulated long noncoding RNAs (lncRNAs) in hepatocellular carcinoma: implications for tumorigenesis, disease progression, and liver cancer stem cells. *Molecular cancer*. 2017;16(1):165.

31. Chen Q, Zhu C, Jin Y. The Oncogenic and Tumor Suppressive Functions of the Long Noncoding RNA MALAT1: An Emerging Controversy. *Frontiers in genetics*. 2020;11:93.
32. Guzel E, Okyay TM, Yalcinkaya B, Karacaoglu S, Gocmen M, Akcakuyu MH. Tumor suppressor and oncogenic role of long non-coding RNAs in cancer. *North Clin Istanb*. 2020;7(1):81-6.
33. Soudyab M, Iranpour M, Ghafouri-Fard S. The Role of Long Non-Coding RNAs in Breast Cancer. *Arch Iran Med*. 2016;19(7):508-17.
34. Luan X, Wang Y. LncRNA XLOC_006390 facilitates cervical cancer tumorigenesis and metastasis as a ceRNA against miR-331-3p and miR-338-3p. *J Gynecol Oncol*. 2018;29(6):e95.
35. Kong X, Duan Y, Sang Y, Li Y, Zhang H, Liang Y, et al. LncRNA-CDC6 promotes breast cancer progression and function as ceRNA to target CDC6 by sponging microRNA-215. *Journal of cellular physiology*. 2019;234(6):9105-17.
36. Chen J, Huang X, Wang W, Xie H, Li J, Hu Z, et al. LncRNA CDKN2BAS predicts poor prognosis in patients with hepatocellular carcinoma and promotes metastasis via the miR-153-5p/ARHGAP18 signaling axis. *Aging (Albany NY)*. 2018;10(11):3371-81.
37. Zhong H, Yang J, Zhang B, Wang X, Pei L, Zhang L, et al. LncRNA GACAT3 predicts poor prognosis and promotes cell proliferation in breast cancer through regulation of miR-497/CCND2. *Cancer biomarkers : section A of Disease markers*. 2018;22(4):787-97.
38. Zhou W, Ye XL, Xu J, Cao MG, Fang ZY, Li LY, et al. The lncRNA H19 mediates breast cancer cell plasticity during EMT and MET plasticity by differentially sponging miR-200b/c and let-7b. *Sci Signal*. 2017;10(483).
39. Zhao W, Geng D, Li S, Chen Z, Sun M. LncRNA HOTAIR influences cell growth, migration, invasion, and apoptosis via the miR-20a-5p/HMGA2 axis in breast cancer. *Cancer Med*. 2018;7(3):842-55.
40. Zhou S, Wang L, Yang Q, Liu H, Meng Q, Jiang L, et al. Systematical analysis of lncRNA-mRNA competing endogenous RNA network in breast cancer subtypes. *Breast cancer research and treatment*. 2018;169(2):267-75.
41. Hamilton MJ, Girke T, Martinez E. Global isoform-specific transcript alterations and deregulated networks in clear cell renal cell carcinoma. *Oncotarget*. 2018;9(34):23670-80.
42. Li S, Liu X, Li H, Pan H, Acharya A, Deng Y, et al. Integrated analysis of long noncoding RNA-associated competing endogenous RNA network in periodontitis. *J Periodontal Res*. 2018;53(4):495-505.
43. Lv Y, Wei W, Huang Z, Chen Z, Fang Y, Pan L, et al. Long non-coding RNA expression profile can predict early recurrence in hepatocellular carcinoma after curative resection. *Hepatol Res*. 2018;48(13):1140-8.
44. Sun M, Geng D, Li S, Chen Z, Zhao W. LncRNA PART1 modulates toll-like receptor pathways to influence cell proliferation and apoptosis in prostate cancer cells. *Biol Chem*. 2018;399(4):387-95.
45. Li M, Zhang W, Zhang S, Wang C, Lin Y. PART1 expression is associated with poor prognosis and tumor recurrence in stage I-III non-small cell lung cancer. *J Cancer*. 2017;8(10):1795-800.

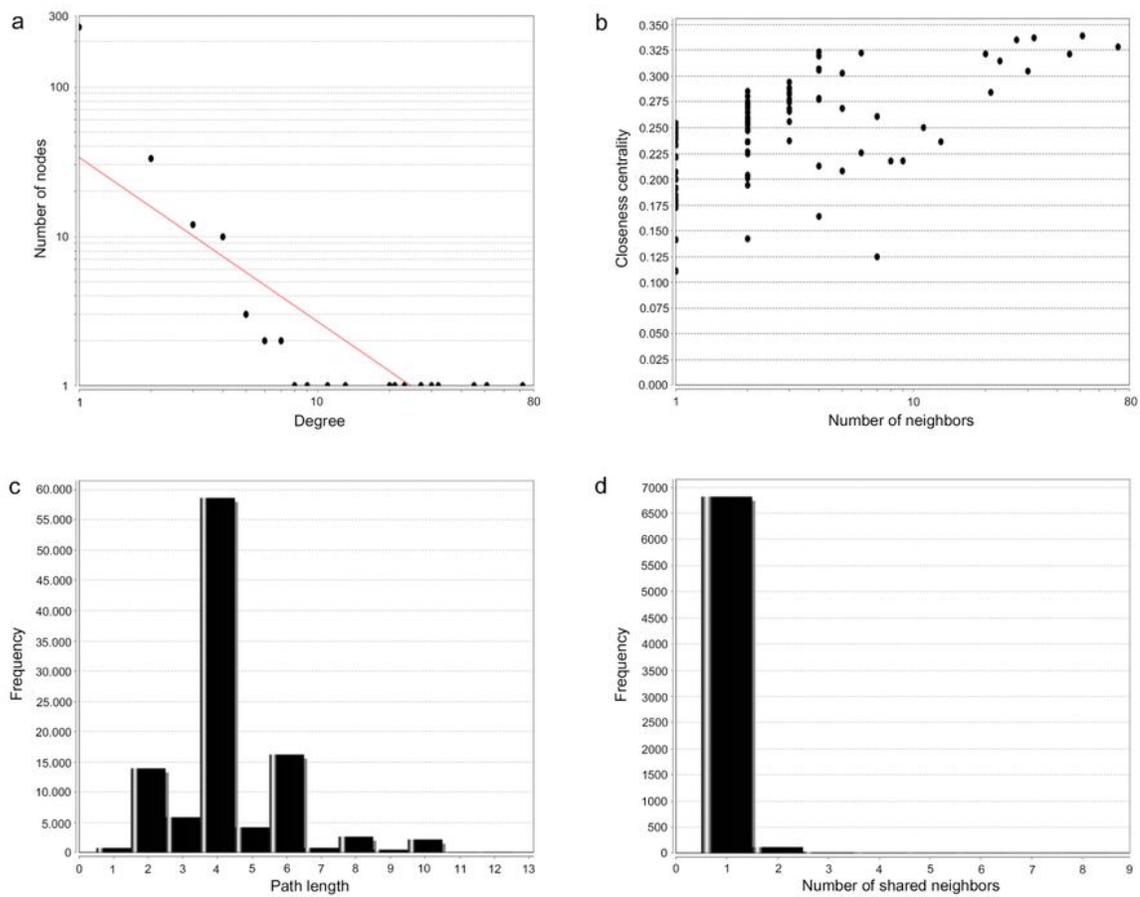


Figure 2

CeRNA network topology analysis and stability analysis. a: Node degree distribution; b: Proximity central distribution graph; c: Shortest path distribution graph; d: Degree distribution density graph of nodes.

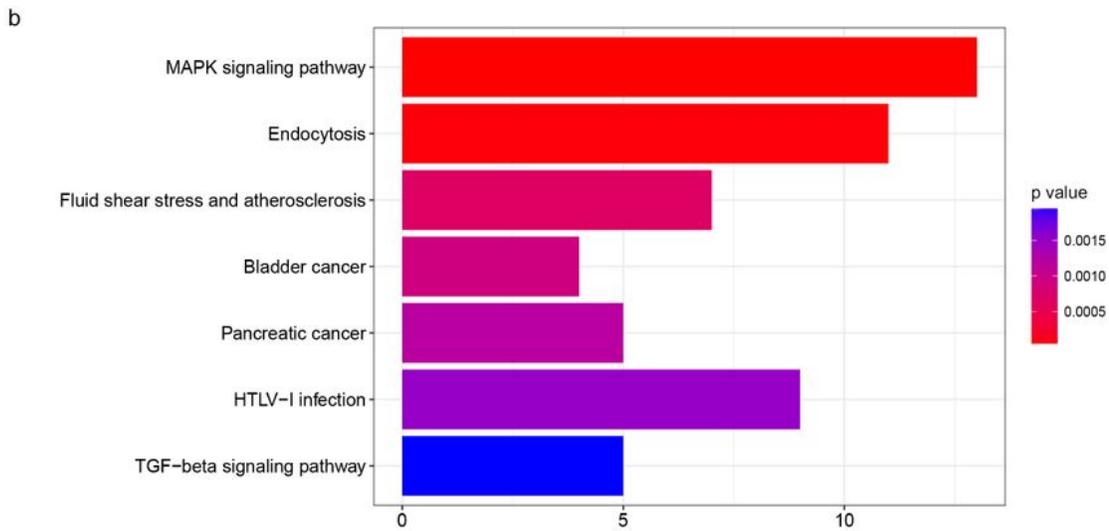
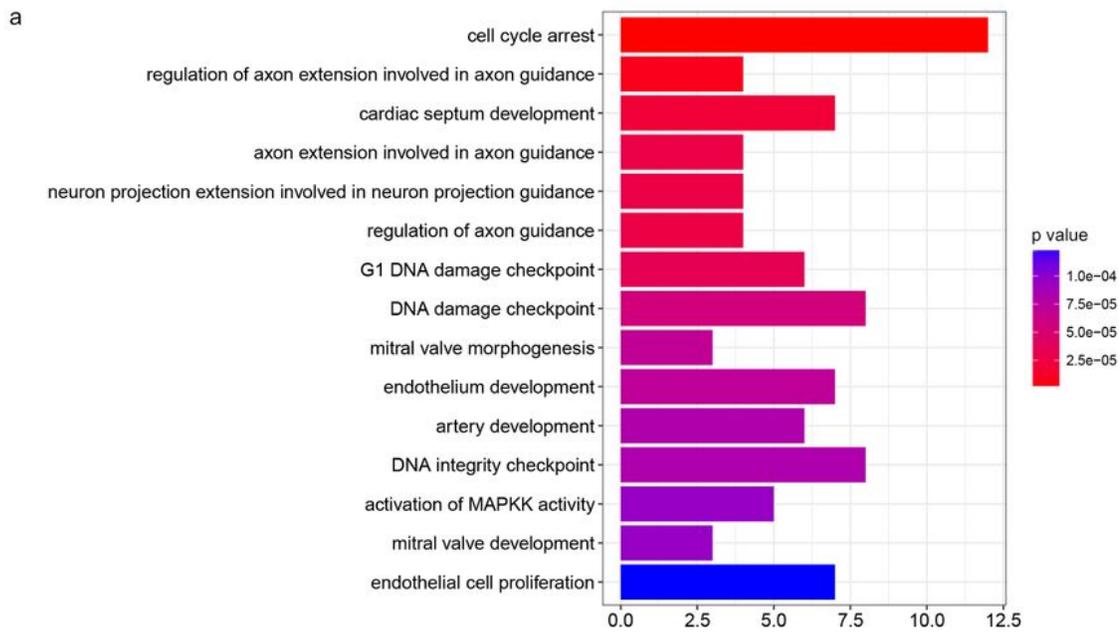


Figure 3

Enrichment analysis. a: GO enrichment analysis results; b: KEGG enrichment analysis results.

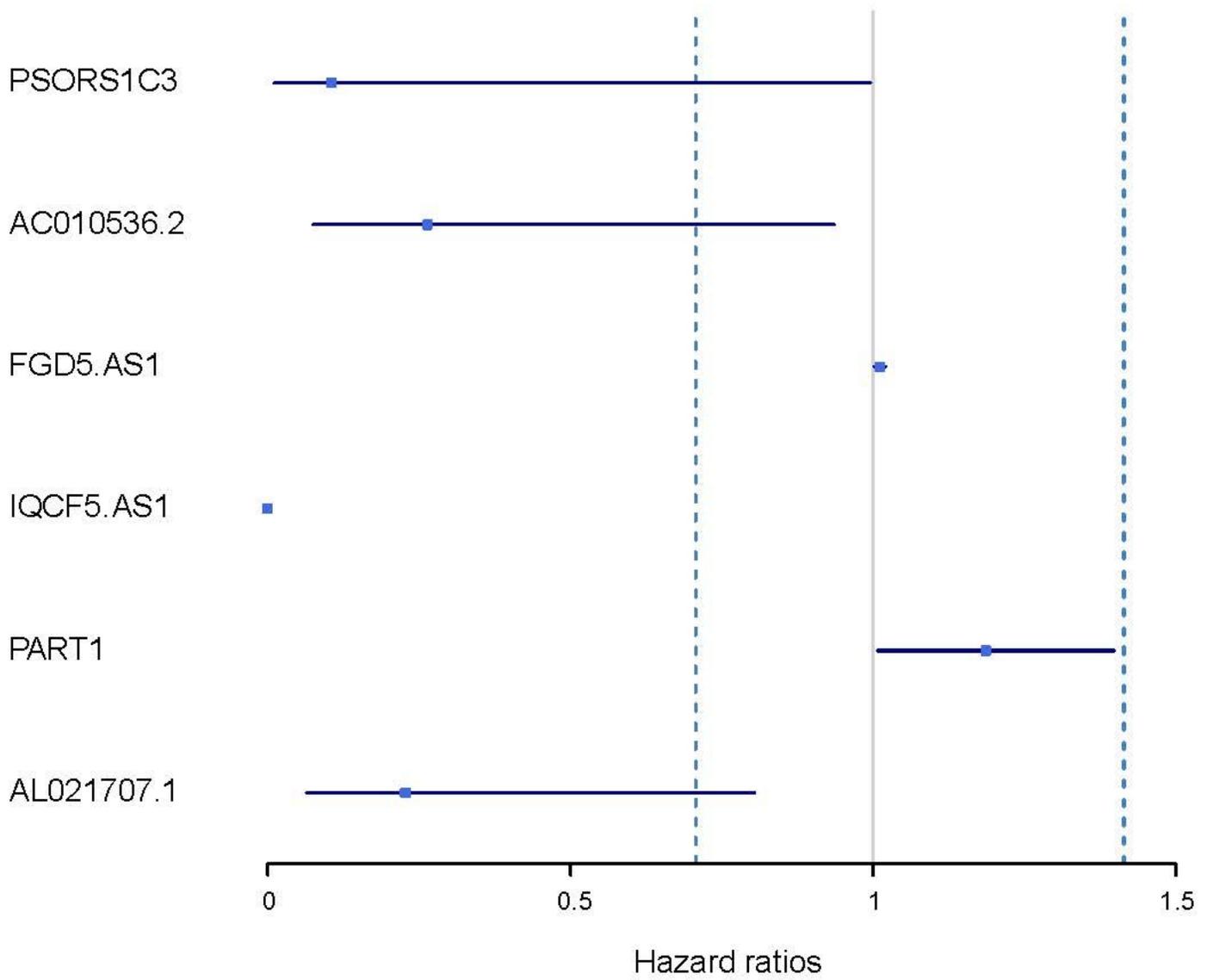


Figure 4

Forest plots of 6 prognosis related lncRNAs. A forest map of 6 prognostically differentiated lncRNAs.

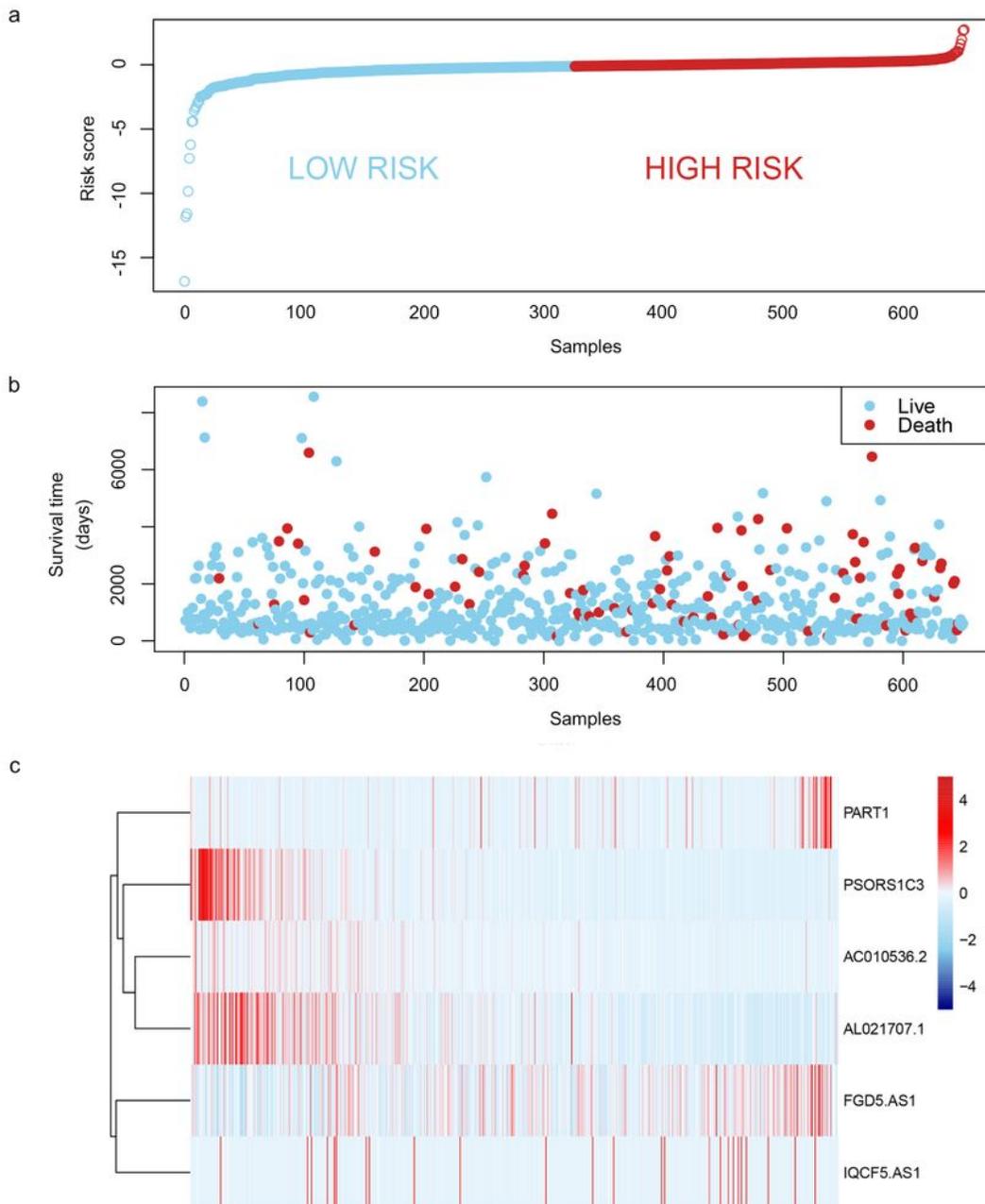


Figure 5

The relationship between the expression of 6 lncRNAs and risk scores. The horizontal axis represents the sample's risk coefficient and the score increases from left to right. a: The lncRNA risk score distribution of ER+PR+ breast cancer samples in TCGA database. b: Survival status and time of all ER+PR+ breast cancer patients across lncRNA risk score. c: Heatmap and hierarchical clustering of lncRNAs from the prognostic risk model are generated.

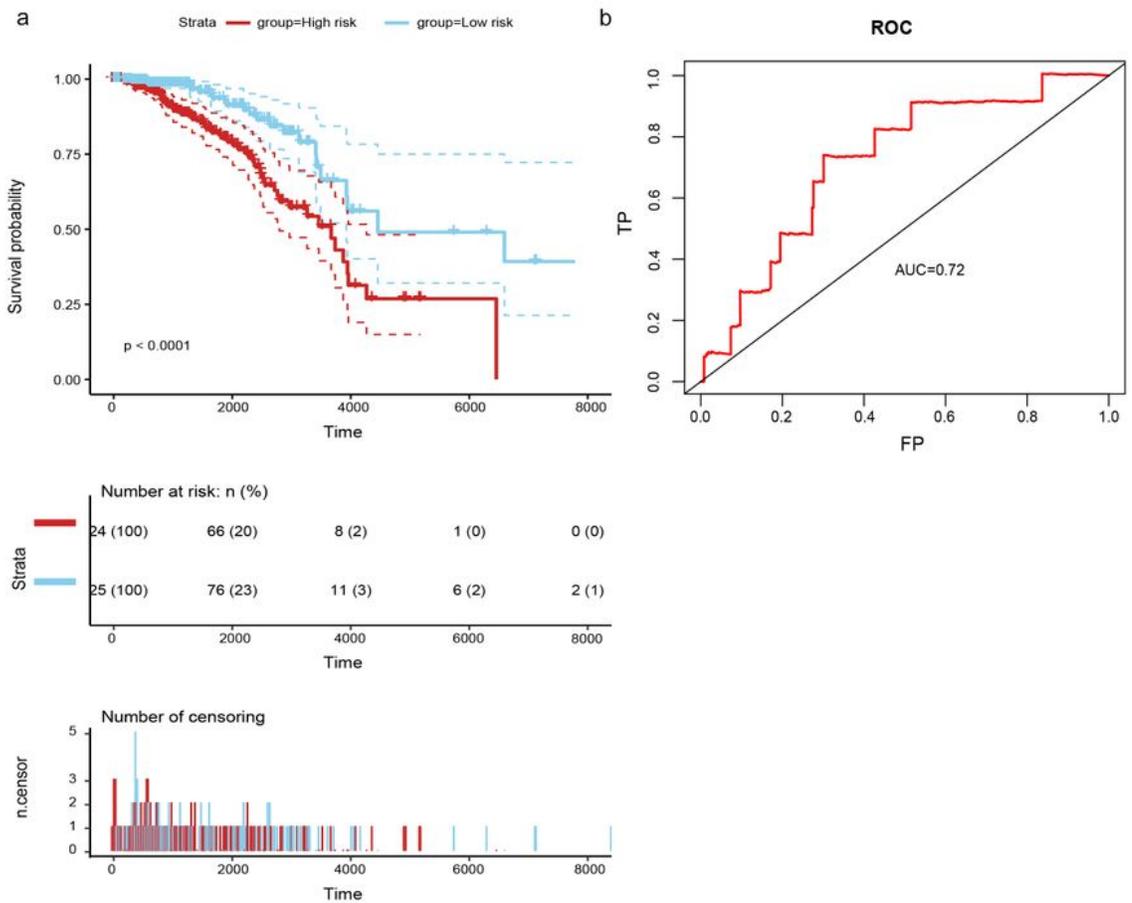


Figure 6

6 lncRNA model prognostic analysis. a: 6 lncRNA multivariate regression models for prognostic analysis and K-M curves. b: ROC curves of 6 lncRNA multivariate regression models.