Identification of a Hypoxia-Related LncRNA Signature for the Prognosis of Colorectal Cancer

Hua Huang
Changshu Hospital, Nanjing University of Chinese Medicine
Mingjia Gu
Changshu Hospital, Nanjing University of Chinese Medicine
Shanshan Xu
Nanjing University of Chinese Medicine
Youran Li
Affiliated Hospital of Nanjing University of Chinese Medicine
Yunfei Gu
Affiliated Hospital of Nanjing University of Chinese Medicine
Lijiang Ji (Ji512@163.com)
Changshu Hospital, Nanjing University of Chinese Medicine

Research Article

Keywords: Colorectal cancer (CRC), hypoxia-related lncRNA, prognosis

Posted Date: December 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-757504/v2

Abstract

Background: Colorectal cancer (CRC), the commonly seen malignancy, ranks 3rd place among the causes of cancer-associated mortality. As suggested by more and more studies, long non-coding RNAs (lncRNAs)have been considered as prognostic biomarkers for CRC. But the significance of hypoxic lncRNAs in predicting CRC prognosis remains unclear.

Methods: The gene expressed profiles for CRC cases were obtained based on the Cancer Genome Atlas (TCGA) and applied to estimate the hypoxia score using a single-sample gene set enrichment analysis (ssGSEA) algorithm. Overall survival (OS) of the high- and low-hypoxia score group was analyzed by the Kaplan–Meier (KM) plot. To identify differentially expressed lncRNAs (DELs) between two hypoxia score groups, this study carried out differential expression analysis, and then further integrated with the DELs between controls and CRC patients to generate the hypoxia-related lncRNAs for CRC. Besides, prognostic lncRNAs were screened by the univariate Cox regression, which was later utilized for constructing the prognosis nomogram for CRC by adopting the least absolute shrinkage and selection operator (LASSO) algorithm. In addition, both accuracy and specificity of the constructed prognostic signature were detected through the receiver operating characteristic (ROC) analysis. Moreover, our constructed prognosis signature also was validated in the internal testing test. This study operated gene set enrichment analysis (GSEA) for exploring potential biological functions associated with the prognostic signature. Finally, the ceRNA network of the prognostic lncRNAs was constructed.

Results: Among 2299 hypoxia-related lncRNAs of CRC in total, LINC00327, LINC00163, LINC00174, SYNPR-AS1, and MIR31HG were identified as prognostic lncRNAs by the univariate Cox regression and adopted for constructing the prognosis signature for CRC. ROC analysis showed the predictive power and accuracy of the prognostic signature. Additionally, the GSEA revealed that ECM-receptor interaction, PI3K-Akt pathway, phagosome, and Hippo pathway were mostly associated with the high-risk group. 352 miRNAs-mRNAs pairs and 177 lncRNAs-miRNAs were predicted.

Conclusion: To conclude, we identified 5 hypoxia-related lncRNAs to establish an accurate prognostic signature for CRC, providing important prognostic markers and therapeutic targets.

Background

Colorectal cancer (CRC), a malignancy, ranks 3nd place among the causes leading to cancer-associated mortality worldwide, with a mortality rate as high as 9.2% [1]. Surgery remains the efficient method to treat CRC. However, approximately 80% of patients relapse within 3 years [2]. Due to its high recurrence rate and late-stage metastasis to distant organs, the mortality rate of CRC patients remains extremely high. Therefore, accurate assessment of the prognosis of CRC patients is clinically significant.

Long non-coding RNAs (lncRNAs), the RNA molecules that contain at least 200 nucleotides, can not encode proteins. As high-throughput sequencing (HTS) technology develops, an increasing number of studies have shown that lncRNAs have a key function in cancer genesis and progression [3, 4], which can

be used to diagnose and predict the prognosis of pancreatic cancer [5], hepatocellular carcinoma (HCC) [6], lung cancer (LC) [7], colorectal cancer (CRC) [8] and gastric cancer (GC) [9]. Due to the accuracy of lncRNA predicting prognosis, researches on constructing prognostic risk characteristics based on lncRNAs have attracted considerable attention [10-12]. Hypoxia is one of the main tumor microenvironment (TME) characteristics, which predicts the dismal survival of many tumors. Hypoxia can promote cancer cell growth, metastasis, invasion, new vessel formation, and resistance to treatment [13]. According to the previous study, lncRNA NORAD can promote the mesenchymal transition of pancreatic cancer under hypoxic conditions [5]. lncRNA HAS2-AS1 level increases within the hypoxic oral squamous cell carcinoma (OSCC) tissues, thereby promoting tumor invasiveness [14]. However, there are no available prognostic biomarkers for CRC explored on the basis of the expression levels of hypoxic lncRNAs.

Therefore, this study systematically explored the prognostic role of hypoxia-associated lncRNAs in CRC by bioinformatic methods, developed a hypoxia-related lncRNA prognostic signature, and constructed a ceRNA regulatory network based on prognostic lncRNAs. This study may provide a primary theoretical basis for elucidating the mechanisms by which lncRNAs regulate the development of CRC.

Results

Identification of hypoxia-related lncRNAs for CRC

We generated a K-M curve to assess the OS of high- versus low-hypoxia score groups upon the hypoxia score threshold. The result demonstrated that for CRC patients, those showing high hypoxia scores exhibited shorter OS relative to low-hypoxia score patients (**Fig. 1A**). As shown in **Fig.1B**, some clinical traits (T and M) of CRC patients were found to be mostly related to the hypoxia score ($P < 0.05$). Additionally, we identified 2889 DELs between the two hypoxia score groups, including 656 upregulated and 2233 downregulated lncRNAs (**Fig. 1C-D**). Between the controls and CRC patients, a total of 9544 DELs were also screened with 3373 being upregulated and 6171 being downregulated (**Fig. 1E-F**). The 2889 DELs of the hypoxia groups were overlapped with the 9544 DELs among the controls and patients to obtain 2299 lncRNAs mainly associated with CRC for further study through the Venn analysis (**Fig. 1G**).

Establishment of the prognostic risk signature for CRC

Next, effects of those lncRNAs on prognosis of CRC were detected. The univariate Cox regression analysis manifested that LINC00327 ($P = 0.017$, HR = 1.4, 95% CI: 1.1-1.8), LINC00163 ($P = 0.043$, HR = 1.5, 95% CI: 1-2.1), LINC00174 ($P = 0.049$, HR = 1.5, 95% CI: 1-2.1), SYNPR-AS1 ($P = 0.027$, HR = 0.81, 95% CI: 0.67-0.98), and MIR31HG ($P = 0.023$, HR = 1.2, 95% CI: 1-1.4) presented powerful effects on the survival of CRC, which were further used to construct a prognostic signature for CRC by the LASSO algorithm (Risk score=LINC00327*0.2569555+LINC00163*0.2121025+LINC00174*0.4270289-SYNPR-AS*0.1209041+MIR31HG*0.1299143), for the sake of classifying high- or low-risk patients (**Fig. 2A-B**). The high-risk CRC cases showed a dismal outcome ($P < 0.001$, **Fig. 2C**). The 1-, 3-, and 5-year AUC values

of ROC curve were 0.691, 0.701, and 0.688, separately, suggesting the better performance of our **prognostic risk signature** for CRC (**Fig. 2D**). Besides, with the increase in risk score, more training set samples died. Levels of prognostic lncRNAs are also displayed in the heatmap plot (**Fig. 2E**).

Verification of our constructed risk model

We also validated our previously constructed risk **model** for its accuracy in the testing set. KM analysis revealed that high-risk group patients showed obviously worse OS vs low-risk patients (P = 0.042, **Fig. 3A**). 1-, 3- and 5-year AUC of that gene signature reached 0.765, 0.706, and 0.701, separately, by the ROC analysis, confirming the accuracy of the prognostic signature (**Fig. 3B**). **Fig. 3C** illustrates the distributions of risk scores, expression of risk genes, and survival status. Consistent with those above results, findings in the whole test also presented a predictive accuracy of the risk signature (**Fig. 3D-F**).

Association of risk score with clinical traits for CRC

Subsequently, we examined the association of prognosis **signature** with other clinical traits for CRC. As displayed in **Fig. 4A-B**, clinical traits of CRC patients (stage, T, N, M) showed positive correlation with the risk score (P < 0.05). Univariate Cox analysis identified N (P = 2.7e-05, HR = 2.7, 95% CI: 1.7-4.3), M (P = 1.4e-07, HR = 3.5, 95% CI: 2.2-5.5), T (P = 0.0054, HR = 5.2, 95% CI: 1.6-16), stage (P = 3.3e-06, HR = 3.1, 95% CI: 1.9-4.9) and risk score (P = 1.6e-05, HR = 2.3, 95% CI: 1.6-3.4) as the candidate factors to independently predict the prognosis of CRC (**Fig. 4C**). Moreover, multivariable Cox analysis identified M (P = 0.00061, HR = 2.4, 95% CI: 1.4-3.9), risk score (P = 0.00079, HR = 2, 95% CI: 1.3-2.9) and stage (P = 0.029, HR = 1.8, 95% CI: 1.1-3) as the factors to independently predict prognosis for further study (**Fig. 4D**).

We further included all independent prognostic factors for CRC (risk score, stage, M) in constructing the nomogram for predicting 1-, 3-, and 5-year survival (**Fig. 4E**). According to **Fig. 4F**, calibration plots for 1-year OS probabilities of constructed nomogram presented an optimal agreement with the ideal observation by calibration curves analysis.

Functional enrichment on the prognosis signature

GSEA result determined that several important signaling pathways were remarkably related to high-risk patients, including ECM-receptor interaction, Hippo signaling pathway, phagosome, endoplasmic reticulum protein processing, PI3K-Akt pathway, and Rap1 signaling pathway, indicating that these pathways might contribute to the pathology development of high-risk CRC cases (**Fig. 5**).

Construction of a ceRNA network

For investigating the prognostic lncRNAs related mechanism, this study constructed a ceRNA network. We obtained 6407 risk score-related DEGs that satisfied $|\log_2 FC| \geq 1$ and FDR ≤ 0.05 between high- and low-risk by R package DEseq2 **(Fig. 6A)**. Subsequently, 352 miRNAs targeting the above DEGs were predicted by the miRWalk database, while the miRcode database predicted the relationship pairs of 5

prognostic lncRNAs and 177 miRNAs. After excluding the positive correlation miRNAs, an mRNA-miRNA-lncRNA network was visualized in **Fig. 6B**.

Discussion

The incidence and mortality of CRC have remained high for numerous years. The establishment of a prognostic model is helpful for the early detection and treatment of CRC patients. Current studies have confirmed that lncRNA is the largest subtype of non-coding RNA, which can encode microproteins to participate in epigenetic regulation and cell cycle control [15], apoptosis [16], immune response [17], tumor cell proliferation [18], migration and invasion [19] as well as many other important physiological functions.

This study found 5 potential hypoxia-related lncRNAs as follows: LINC00327, LINC00163, LINC00174, SYNPR-AS1, and MIR31HG, which can be used as prognostic markers for CRC. The high expression of LINC00163 in human papillary thyroid carcinoma (PTC) promotes cancer cell growth and metastasis through regulating epithelial-mesenchymal transition (EMT) [20]. When treating GC, LINC00163 can be used as a ceRNA to enhance the anti-cancer effect of AKAP12 by inhibiting the expression of miR-183 [9]. In addition, LINC00163 is effective in the treatment and prediction the prognosis for bladder cancer and lung cancers [21-22]. In this study the overexpression of LINC00174 is markedly related to higher mortality and tumor grade in survival analysis. Jin and colleagues [8] discovered that LINC00174 up-regulation predicted the unfavorable survival outcomes of colon adenocarcinoma (COAD). In addition, LINC00174 may serve as the molecular target to diagnose and treat glioma [23]. Down-regulation of its expression significantly suppresses tumor cell growth in vivo and in vitro [24-25]. Recent studies have shown that SYNPR-AS1, an immune-related lncRNA, has a key function within the immune microenvironment of lung adenocarcinoma (LUAD), which is also involved in cancer genesis and progression [7]. According to Wu and colleagues [26], SYNPR-AS1 was considered a protective biomarker in LUAD. At the same time, the SYNPR-AS1 level was in direct proportion to the survival of non-small cell lung cancer (NSCLC) [27]. MIR31HG, a HIF-1α co-activator, promotes tumor occurrence and development [28-29]. At present, MIR31HG overexpression predicts a good prognosis for gastrointestinal tumors [30-32, 10] while the opposite results have indeed been obtained in lung cancer, breast cancer, glioblastoma, and other cancers [33 -36]. Whether the microenvironment of gastrointestinal tumors or certain specific immune mechanisms affect the clinical significance of MIR31HG requires further clinical and experimental studies.

The present work established the ceRNA network based on 4 lncRNAs, 253 mRNAs, and their corresponding 12 targeted miRNA, and it helped to better elucidate the CRC molecular mechanism. Studies have revealed that miR-30b overexpression inhibits CRC cells LS-174t growth [37]. Based on Fan and colleagues [38], miR-30b-5p may exert a tumor metastasis inhibitory effect by targeting the regulatory gene Rap1b of cell adhesion and migration. MiR-24-3p is the main regulator of the miR-23a-24-27 gene cluster. Overexpression of this factor suppresses CRC cell growth, invasion, and migration [39]. miR-24-3p overexpression indicates the dismal survival of CRC cases, but it has nothing to do with the

clinical-pathological indicators currently used to judge the prognosis of CRC [40]. MiR-107 is a tumor-related gene, which can influence the occurrence, development, recurrence, and metastasis of various tumors [41]. However, miR-107 has diverse functions within different tumor tissues. Its up-regulation in vitro promotes human CRC cell invasion and growth [42]. Meanwhile, miR-107 expression was in indirect proportion to CRC survival, and direct proportion to distant metastasis (DM) and lymph node metastasis (LNM) of CRC. [43]. Thus, miR-107 can be adopted for evaluating chemotherapy effects and prognosis for CRC. Regarding the mechanism of action, Liu et al. [44] proved that miR-107 significantly enhances LoVo and SW480 CRC cell proliferation through targeting Par4. Zhang et al. [45] proved that lncRNA UASR1 overexpression reduced miR-107's inhibition on CRC as well as CDK8 level. However, no relevant data on the interaction between lncRNAs and miRNAs in the ceRNA network have been found in previous literature, but also proves that CRC has a large space for exploration in the mechanism of the hypoxic tumor microenvironment.

Conclusion

Through data mining and literature summary, we finally obtained lncRNAs related to the prognosis of CRC under hypoxic microenvironment and elaborated the research status of 4 lncRNAs and some miRNAs related to CRC from the ceRNA network. This study provides a theoretical basis for treating and predicting the prognosis for CRC. However, the present work has many limitations, such as small sample size and insufficient literature support. In future work, we need to further collect experimental and clinical data to verify the above results.

Materials And Methods

Data collection

We acquired transcriptome profiling data of 503 samples with CRC based on TCGA database (https://portal.gdc.cancer.gov/), which contained 41 controls and 462 CRC patients. Data on the gene sets in total related to hypoxia were acquired based on the MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb, hypoxia-induced genes: M10508, hypoxia response of cells: M26925) along with 151 hypoxia-related genes were identified for subsequent study.

Differentially expressed analysis

This study adopted the R DEseq2 package for identifying DELs of control versus CRC samples. This study selected lncRNAs meeting the screening criteria of $\log_2$ fold change (FC) > 0.5 and false discovery rate (FDR) < 0.05 as DELs. Moreover, DELs between two groups also screened with identical criteria, which further overlapped with the harvested DELs between controls and patients to obtain hypoxia-related lncRNAs of CRC through the Venn diagram analysis (http://bioinformatics.psb.ugent.be/webtools/Venn/).

Hypoxia score identification for CRC patients

Based on the transcriptome profiling information and hypoxia-related gene expression profiles, the hypoxia scores for all samples were determined through the ssGSEA method using the gsva package in R, showing the hypoxia status for the patients. Then, CRC cases were classified into 2 groups based on the threshold of hypoxia score, namely, the high- and low-hypoxia score groups. Moreover, we conducted KM analysis for assessing the OS of patients. Besides, the associations of hypoxia score with clinical traits (M and T) were also detected by Pearson analysis.

Construction of prognostic risk signature

At a ratio of 7:3, all CRC patients were classified as training or test set.

In the training set, we conducted Univariate Cox regression analysis among those hypoxic lncRNAs, for the sake of identifying the most relevant prognostic lncRNAs for CRCs upon $P<0.05$. Subsequently, a prognostic risk signature that included the most relevant lncRNAs for CRC was generated with the LASSO algorithm using the R glmnet package. Risk scores for CRC cases were estimated below:

$$\text{Risk score} = e^{\text{sum(expression of every gene} \times \text{relevant coefficient)}} / e^{\text{sum(mean expression of every gene} \times \text{relevant coefficient)}}$$

We separated cases into the high- or low-risk group according to the median risk score. Thereafter, this study carried out a KM curve analysis for detecting survival differences between 2 groups using the R survival package. The accuracy of our constructed prognosis nomogram for CRC was also assessed through the ROC curves using the survivalROC package in R. The above results were also needed to validate in the testing set to confirm the predictive accuracy of the signature.

The nomogram construction

The nomogram was constructed for predicting OS for CRC cases by integrating risk score with other crucial clinical traits using the R package rms. Calibration plots analysis was conducted to further assess the accuracy of our model in predicting prognosis.

GSEA

The different enriched pathways between the two risk groups were analyzed by the GSEA (http://www.broadinstitute.org/gsea) method using the gsea package in R. The P-value < 0.05 stood for significance.

ceRNA network construction

DEGs between two groups were identified using the R DEseq2 package. Then, we utilized the miRWalk (http://mirwalk.umm.uni-heidelberg) database for predicting those targeted miRNAs of DEGs. The miRNA-prognostic lncRNA interactions were predicted by the miRcode database (http://www.mircode.org/). Based on the miRNA-mRNA and miRNA-lncRNA interaction pairs, we constructed and visualized a ceRNA network by Cytoscape software (https://www.cytoscape.org/).

Statistical analysis

This study adopted R pheatmap package to draw the heatmap plot. The R programming language was adopted for statistical analyses. Unless the specified noted, $P<0.05$ stood for statistical significance.

Abbreviations

CRC: Colorectal cancer, lncRNAs: long coding RNAs, TCGA: The Cancer Genome Atlas, ssGSEA: single-sample gene set enrichment analysis, OS: Overall survival, KM: Kaplan–Meier, DELs: differentially expressed lncRNAs, LASSO: least absolute shrinkage and selection operator, ROC: operating characteristic, GSEA: gene set enrichment analysis, HTS: high-throughput sequencing, HCC: hepatocellular carcinoma, LC: lung cancer, GC: gastric cancer, TME: tumor microenvironment, OSCC: oral squamous cell carcinoma, PTC: papillary thyroid carcinoma, EMT: epithelial-mesenchymal transition, COAD: colon adenocarcinoma, LUAD: lung adenocarcinoma (LUAD).

Declarations

Acknowledgments

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Funding information: None

Declaration of conflicts: No conflict of interests existed in this study.

Author's contribution: Lijiang Ji and Yunfei Gu carried out the studies. Hua Huang and Mingjia Gu participated in collecting the data and drafted the manuscript. Shanshan Xu and Youran Li analyzed the data. Hua Huang and Mingjia Gu directed the writing. Lijiang Ji helped to revise the manuscript. The authors read and approved the final manuscript.

References

1. 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 Cancer J Clin. 2018 Nov,68(6):394-424.
2. Engstrand J, Nilsson H, Strömberg C, Jonas E, Freedman J. Colorectal cancer liver metastases - a population-based study on incidence, management and survival. BMC Cancer. 2018 Jan 15,18(1):78.
3. Wang W, Min L, Qiu X, Wu X, Liu C, Ma J, Zhang D, Zhu L. Biological Function of Long Non-coding RNA (LncRNA) Xist. Front Cell Dev Biol. 2021 Jun 10,9:645647.
4. Liu SJ, Dang HX, Lim DA, Feng FY, Maher CA. Long noncoding RNAs in cancer metastasis. Nat Rev Cancer. 2021 Jul,21(7):446-460.
5. Li H, Wang X, Wen C, Huo Z, Wang W, Zhan Q, Cheng D, Chen H, Deng X, Peng C, Shen B. Long noncoding RNA NORAD, a novel competing endogenous RNA, enhances the hypoxia-induced epithelial-mesenchymal transition to promote metastasis in pancreatic cancer. Mol Cancer. 2017 Nov 9,16(1):169.
6. Yuan P, Qi X, Song A, Ma M, Zhang X, Lu C, Bian M, Lian N, He J, Zheng S, Jin H. LncRNA MAYA promotes iron overload and hepatocyte senescence through inhibition of YAP in non-alcoholic fatty liver disease. J Cell Mol Med. 2021 Jun 30.
7. Li JP, Li R, Liu X, Huo C, Liu TT, Yao J, Qu YQ. A Seven Immune-Related lncRNAs Model to Increase the Predicted Value of Lung Adenocarcinoma. Front Oncol. 2020 Oct 14,10:560779.
8. Jin L, Li C, Liu T, Wang L. A potential prognostic prediction model of colon adenocarcinoma with recurrence based on prognostic lncRNA signatures. Hum Genomics. 2020 Jun 10,14(1):24.
9. Zhang J, Piao HY, Guo S, Wang Y, Zhang T, Zheng ZC, Zhao Y. LINC00163 inhibits the invasion and metastasis of gastric cancer cells as a ceRNA by sponging miR-183 to regulate the expression of AKAP12. Int J Clin Oncol. 2020 Apr,25(4):570-583.
10. Tu C, Ren X, He J, Li S, Qi L, Duan Z, Wang W, Li Z. The predictive value of lncRNA MIR31HG expression on clinical outcomes in patients with solid malignant tumors. Cancer Cell Int. 2020 Apr 7,20:115.
11. Zhang L, Wang H, Xu M, Chen F, Li W, Hu H, Yuan Q, Su Y, Liu X, Wuri J, Yan T. Long noncoding RNA HAS2-AS1 promotes tumor progression in glioblastoma via functioning as a competing endogenous RNA. J Cell Biochem. 2020 Jan,121(1):661-671.
12. Yin T, Wu J, Hu Y, Zhang M, He J. Long non-coding RNA HULC stimulates the epithelial-mesenchymal transition process and vasculogenic mimicry in human glioblastoma. Cancer Med. 2021 Jul 2.
13. Kim MH, Green SD, Lin CC, Konig H. Engineering Tools for Regulating Hypoxia in Tumour Models. J Cell Mol Med. 2021 Jul 2.
14. Zhu G, Wang S, Chen J, Wang Z, Liang X, Wang X, Jiang J, Lang J, Li L. Long noncoding RNA HAS2-AS1 mediates hypoxia-induced invasiveness of oral squamous cell carcinoma. Mol Carcinog. 2017 Oct,56(10):2210-2222.
15. Seborova K, Vaclavikova R, Rob L, Soucek P, Vodicka P. Non-Coding RNAs as Biomarkers of Tumor Progression and Metastatic Spread in Epithelial Ovarian Cancer. Cancers (Basel). 2021 Apr

12,13(8):1839.

16. Thakur KK, Kumar A, Banik K, Verma E, Khatoon E, Harsha C, Sethi G, Gupta SC, Kunnumakkara AB. Long noncoding RNAs in triple-negative breast cancer: A new frontier in the regulation of tumorigenesis. J Cell Physiol. 2021 Jun 8.
17. Di Martino MT, Riillo C, Scionti F, Grillone K, Polerà N, Caracciolo D, Arbitrio M, Tagliaferri P, Tassone P. miRNAs and lncRNAs as Novel Therapeutic Targets to Improve Cancer Immunotherapy. Cancers (Basel). 2021 Mar 30,13(7):1587.
18. Ostovarpour M, Khalaj-Kondori M, Ghasemi T. Correlation between expression levels of lncRNA FER1L4 and RB1 in patients with colorectal cancer. Mol Biol Rep. 2021 Jun 16.
19. Lee CP, Ko AM, Nithiyanantham S, Lai CH, Ko YC. Long noncoding RNA HAR1A regulates oral cancer progression through the alpha-kinase 1, bromodomain 7, and myosin IIA axis. J Mol Med (Berl). 2021 Jun 7.
20. Gao P, Jiao H, Zhe L, Cui J. High expression of LINC0163 promotes progression of papillary thyroid cancer by regulating epithelial-mesenchymal transition MITF. Eur Rev Med Pharmacol Sci. 2020 May,24(10):5504-5511.
21. Xu Z, Wang C, Xiang X, Li J, Huang J. Characterization of mRNA Expression and Endogenous RNA Profiles in Bladder Cancer Based on The Cancer Genome Atlas (TCGA) Database. Med Sci Monit. 2019 Apr 25,25:3041-3060.
22. Guo X, Wei Y, Wang Z, Liu W, Yang Y, Yu X, He J. LncRNA LINC00163 upregulation suppresses lung cancer development though transcriptionally increasing TCF21 expression. Am J Cancer Res. 2018 Dec 1,8(12):2494-2506.
23. Wang Z, Wang Q, Bao Z, Guo L, Chen H, Lv T, Xu T, Zhang X, Zhou C, Sun L. LINC00174 is a favorable prognostic biomarker in glioblastoma via promoting proliferative phenotype. Cancer Biomark. 2020,28(4):421-427.
24. Li B, Zhao H, Song J, Wang F, Chen M. LINC00174 down-regulation decreases chemoresistance to temozolomide in human glioma cells by regulating miR-138-5p/SOX9 axis. Hum Cell. 2020 Jan,33(1):159-174.
25. Shi J, Zhang Y, Qin B, Wang Y, Zhu X. Long non-coding RNA LINC00174 promotes glycolysis and tumor progression by regulating miR-152-3p/SLC2A1 axis in glioma. J Exp Clin Cancer Res. 2019 Sep 6,38(1):395.
26. Wu X, Sui Z, Zhang H, Wang Y, Yu Z. Integrated Analysis of lncRNA-Mediated ceRNA Network in Lung Adenocarcinoma. Front Oncol. 2020 Sep 15,10:554759.]showed that SYNPR-AS
27. Wang XW, Guo QQ, Wei Y, Ren KM, Zheng FS, Tang J, Zhang HY, Zhao JG. Construction of a competing endogenous RNA network using differentially expressed lncRNAs, miRNAs and mRNAs in non-small cell lung cancer. Oncol Rep. 2019 Dec,42(6):2402-2415.
28. Shih JW, Chiang WF, Wu ATH, Wu MH, Wang LY, Yu YL, Hung YW, Wang WC, Chu CY, Hung CL, Changou CA, Yen Y, Kung HJ. Long noncoding RNA LncHIFCAR/MIR31HG is a HIF-1α co-activator driving oral cancer progression. Nat Commun. 2017 Jun 22,8:15874.

29. Feng L, Wang R, Lian M, Ma H, He N, Liu H, Wang H, Fang J. Integrated Analysis of Long Noncoding RNA and mRNA Expression Profile in Advanced Laryngeal Squamous Cell Carcinoma. PLoS One. 2016 Dec 29,11(12):e0169232.
30. Qin F, Xu H, Wei G, Ji Y, Yu J, Hu C, Yuan C, Ma Y, Qian J, Li L, Huo J. A Prognostic Model Based on the Immune-Related lncRNAs in Colorectal Cancer. Front Genet. 2021 Apr 20,12:658736.
31. Wang XJ, Zeng B, Lin S, Chen M, Chi P. An Integrated miRNA-lncRNA Signature Predicts the Survival of Stage II Colon Cancer. Ann Clin Lab Sci. 2019 Nov,49(6):730-739.
32. Zhang H, Wang Z, Wu J, Ma R, Feng J. Long noncoding RNAs predict the survival of patients with colorectal cancer as revealed by constructing an endogenous RNA network using bioinformation analysis. Cancer Med. 2019 Mar,8(3):863-873.
33. Zheng S, Zhang X, Wang X, Li J. MIR31HG promotes cell proliferation and invasion by activating the Wnt/β-catenin signaling pathway in non-small cell lung cancer. Oncol Lett. 2019 Jan,17(1):221-229.
34. Dandan W, Jianliang C, Haiyan H, Hang M, Xuedong L. Long noncoding RNA MIR31HG is activated by SP1 and promotes cell migration and invasion by sponging miR-214 in NSCLC. Gene. 2019 Apr 15,692:223-230.
35. Zheng S, Zhang X, Wang X, Li J. MIR31HG promotes cell proliferation and invasion by activating the Wnt/β-catenin signaling pathway in non-small cell lung cancer. Oncol Lett. 2019 Jan,17(1):221-229.
36. Zhang R, Wu D, Wang Y, Wu L, Gao G, Shan D. LncRNA MIR31HG is activated by STAT1 and facilitates glioblastoma cell growth via Wnt/β-catenin signaling pathway. Neurosci Res. 2021 Apr 30:S0168-0102(21)00092-4.
37. Li L, Shao J, Wang J, Liu Y, Zhang Y, Zhang M, Zhang J, Ren X, Su S, Li Y, Cao J, Zang W. MiR-30b-5p attenuates oxaliplatin-induced peripheral neuropathic pain through the voltage-gated sodium channel Nav1.6 in rats. Neuropharmacology. 2019 Jul 15,153:111-120.
38. Fan M, Ma X, Wang F, Zhou Z, Zhang J, Zhou D, Hong Y, Wang Y, Wang G, Dong Q. MicroRNA-30b-5p functions as a metastasis suppressor in colorectal cancer by targeting Rap1b. Cancer Lett. 2020 May 1,477:144-156.
39. Gao Y, Liu Y, Du L, Li J, Qu A, Zhang X, Wang L, Wang C. Down-regulation of miR-24-3p in colorectal cancer is associated with malignant behavior. Med Oncol. 2015 Jan,32(1):362.
40. Kerimis D, Kontos CK, Christodoulou S, Papadopoulos IN, Scorilas A. Elevated expression of miR-24-3p is a potentially adverse prognostic factor in colorectal adenocarcinoma. Clin Biochem. 2017 Apr,50(6):285-292.
41. Peng Y, Croce CM. The role of MicroRNAs in human cancer. Signal Transduct Target Ther. 2016 Jan 28,1:15004.
42. Liu H, Xie H. Overexpression of miR-107 promotes the proliferation and tumorigenic ability of HT29 cells]. Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi. 2018 Oct,34(10):908-913. Chinese.
43. Zhang J, Sun J, Liu J, Gu D, Shi X. Correlation between microRNA-107 expression level and prognosis in patients with colorectal cancer. Int J Clin Exp Pathol. 2020 Sep 1,13(9):2342-2347.

44. Liu F, Liu S, Ai F, Zhang D, Xiao Z, Nie X, Fu Y. miR-107 Promotes Proliferation and Inhibits Apoptosis of Colon Cancer Cells by Targeting Prostate Apoptosis Response-4 (Par4). Oncol Res. 2017 Jul 5,25(6):967-974.
45. Zhang Q, Chen Z. lncRNA UASR1 sponges miR-107 in colorectal cancer to upregulate oncogenic CDK8 and promote cell proliferation. Oncol Lett. 2020 Dec,20(6):305.

Figures

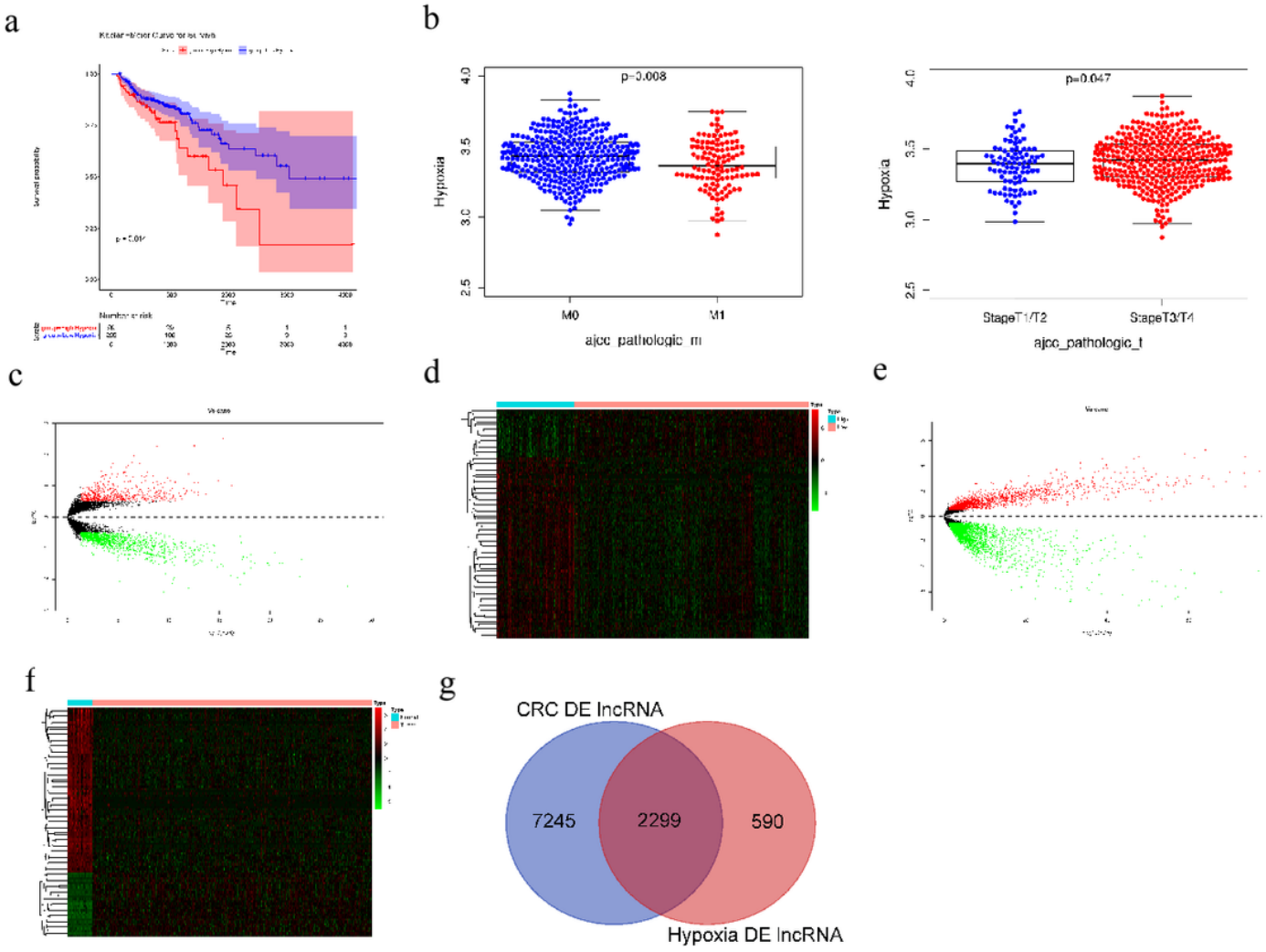

Figure 1

Evaluation and expresstion of hypoxia-related lncRNAs for CRC, (A) Prognosis analysis of high- and low-hypoxia score group, (B) Distribution of hypoxia score of patients with different M (left) and T (right) staging, (C) Volcano plot of differentially expressed lncRNAs between high- and low-hypoxia score groups. Red dots represent up-regulation of expression, while blue represents down-regulation, and black is for genes with insignificant changes, (D) Heatmap of the Top 100 differentially expressed lncRNAs of high-hypoxia score group vs. low-hypoxia score group. Red: up-regulation, green: down-regulation, (E)

Differential lncrna volcanic map between normal and tumor tissues . All dots stand for genes. Red: up-regulation, green: down-regulation, (F) Calorimetric map of differential gene expression between normal and tumor tissues (top 100) The squares stand for genes, with the colors representing their expression levels. A darker color indicates larger gene expression. Red: up-regulation, green: down-regulation, (G) Venn diagram showing the number of common genes in hypoxia score-related lncRNAs and CRC-related lncRNAs.

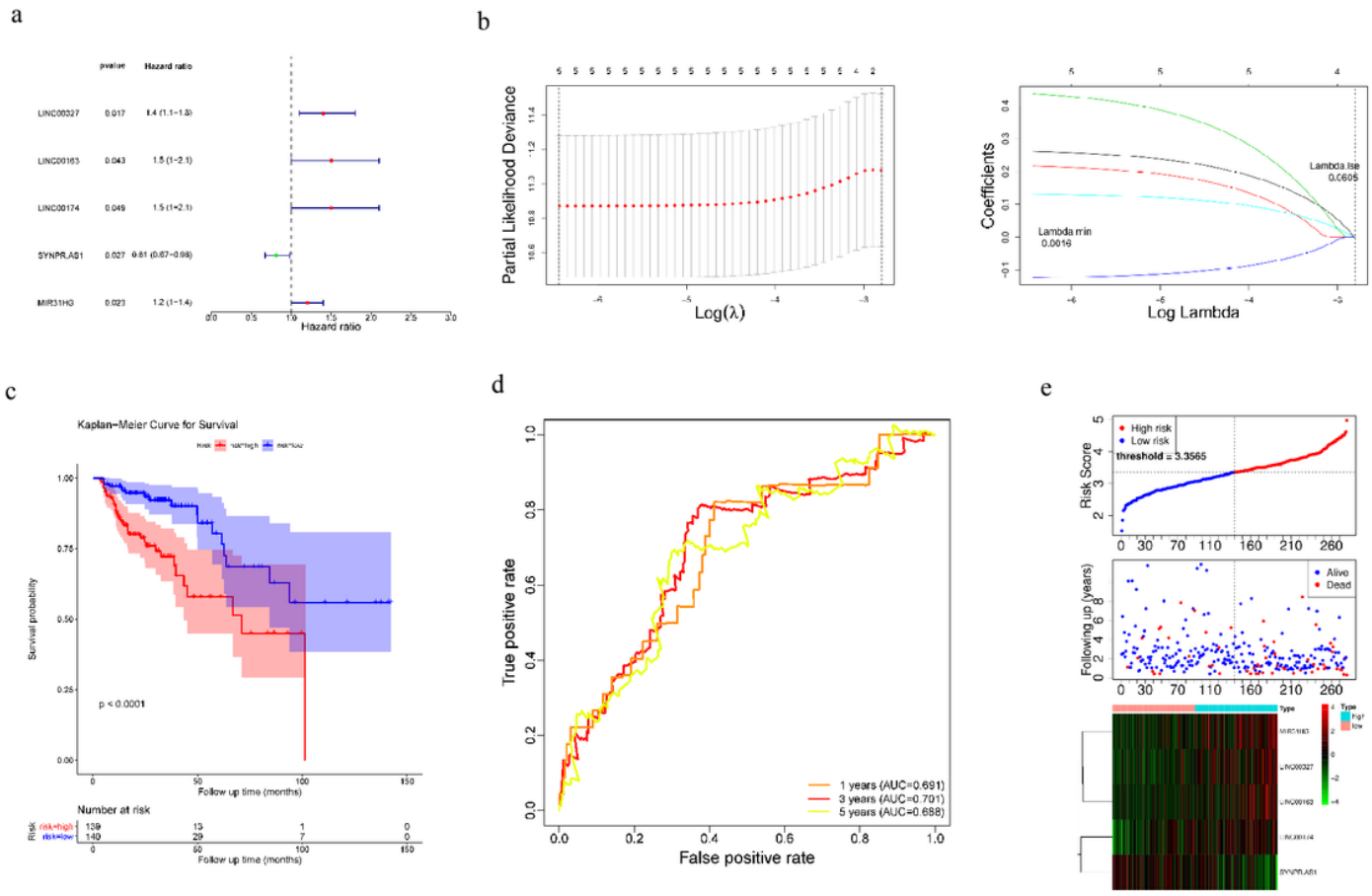

Figure 2

Construction of CRC hypoxia related lncRNAs risk model, (A) Forest plots showing the univariate Cox regression analyses results of the differentially expressed hypoxia-related lncRNAs with OS, (B) Lasso coefficient values of 5 differentially expressed hypoxia-related lncRNAs in CRC. The vertical dashed lines are at the optimal log (lambda) value (left). The changing trajectory of each independent variable. The horizontal axis represents the log value of the independent variable lambda, and the vertical axis represents the coefficient of the independent variable (right), (C) The K-M survival curve of risk score based on five hypoxia-related lncRNAs. Patients in the high-risk group (red) exhibited worse overall

survival (OS) than those in the low-risk group (blue), (D) The areas under the ROC curve about 1 year, 3 years, and 5 years, (E) The risk curve of each sample was reordered by risk score (top). The scatter plot of the sample survival overview (middle). The blue and red dots represent survival and death, respectively. Heatmap showed the expression profiles of the signature in the low-risk groups and high-risk groups. The pink bar represented the low-risk group, and the bright blue bar represents the high-risk group. The -2 to 4 level of lncRNA expression was represented by the evolution from green to red.

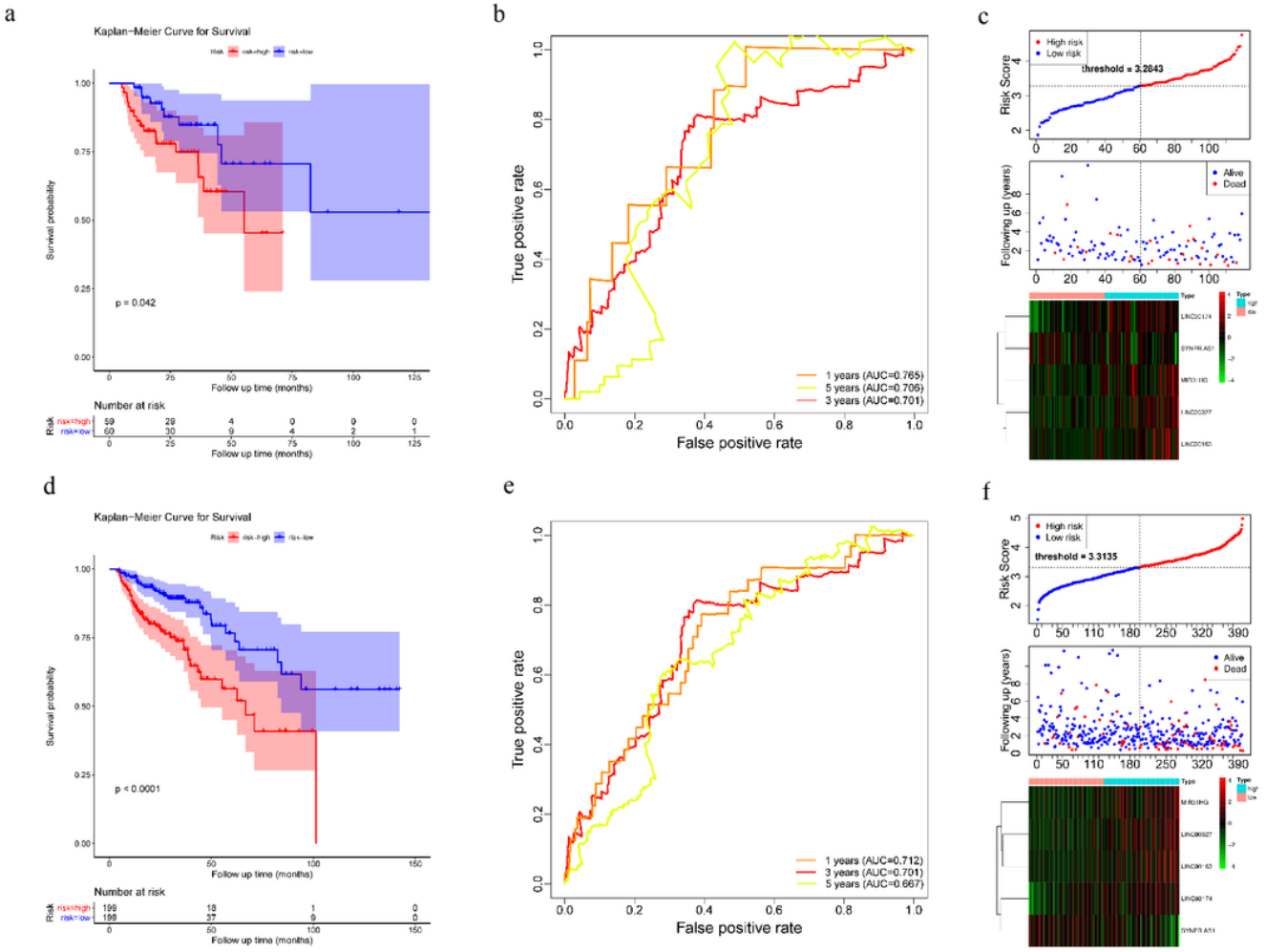

Figure 3

Validation of CRC hypoxia related lncRNAs risk model, (A) According to survival curve of training set, patients in the high-risk group (red) exhibited worse overall survival (OS) than those in the low-risk group (blue), (B) The areas under the ROC curve of training set about 1 year, 3 years, and 5 years, (C) The validation set risk curve of each sample was reordered by risk score (top). The scatter plot of the sample survival overview (middle). The blue and red dots represent survival and death, respectively. The verification set risk heatmap showed the expression profiles of the signature in the low-risk groups and high-risk groups. The pink bar represented the low-risk group, and the bright blue bar represents the high-

risk group. The -2 to 4 level of lncRNA expression was represented by the evolution from green to red (D) Overall survival curve shows that patients in the high-risk group (red) exhibited worse overall survival (OS) than those in the low-risk group (blue), (E) The areas under the ROC curve of overall patients about 1 year, 3 years, and 5 years, (F) The overall risk curve of each sample was reordered by risk score (top). The scatter plot of the sample survival overview (middle). The blue and red dots represent survival and death, respectively. The Overall risk heat map showed the expression profiles of the signature in the low-risk groups and high-risk groups. The pink bar represented the low-risk group, and the bright blue bar represents the high-risk group. The -2 to 4 level of lncRNA expression was represented by the evolution from green to red.

Figure 4

Independent prognostic role of the risk score, (A) The heatmap illustrates the distribution of patients stratified by clinical characteristics between high- and low-risk groups, (B) Distribution of risk score of patients with the different stage (top left), T stage (top right), N stage (bottom left), and M stage (bottom right). Chi-square test showed the correlation between risk score and tumor stage and TNM stage. * $P < 0.05$, ** $P < 0.01$ Univariate (C) and multivariate (D) Cox regression model indicated that the risk score was an independent prognostic predictor for overall survival, (E) Construction of a nomogram combining the risk score and clinical features for prediction of OS, (F) Calibration curve for nomogram-predicting 1-year, 3-year, and 5-year overall survival. The X-axis is nomogram-predicted survival probability and the Y-axis is observed survival probability respectively. The solid lines represent the performance of the nomogram relative to the 45-degree line.

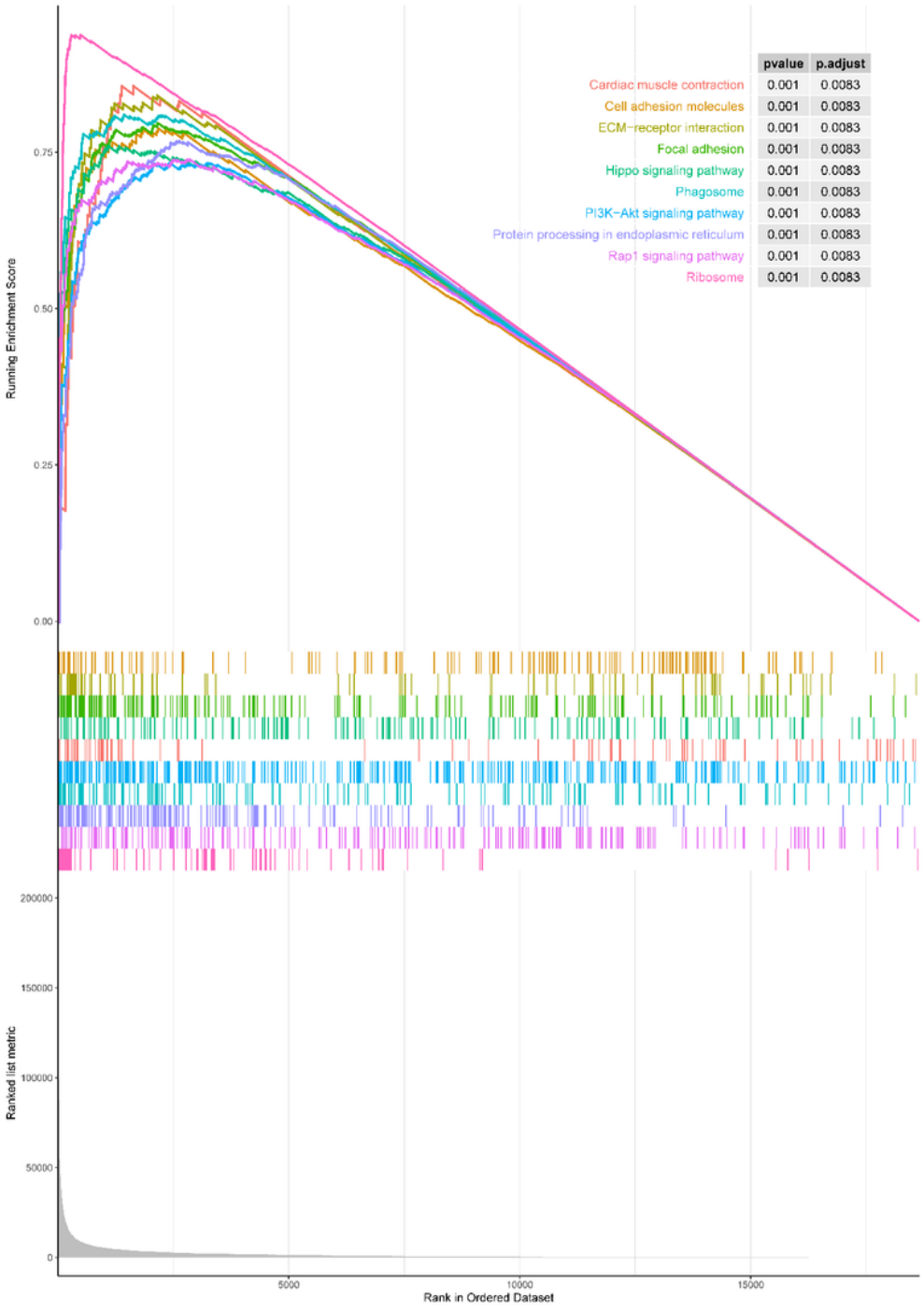

Figure 5

GSEA for samples with high- and low-risk groups. GSEA enrichment analysis of risk score using the TCGA database. Each line represents one particular gene set with unique color. Only gene sets with P < 0.05 were considered significant. And only several leading gene sets (top 10) were displayed in the plot.

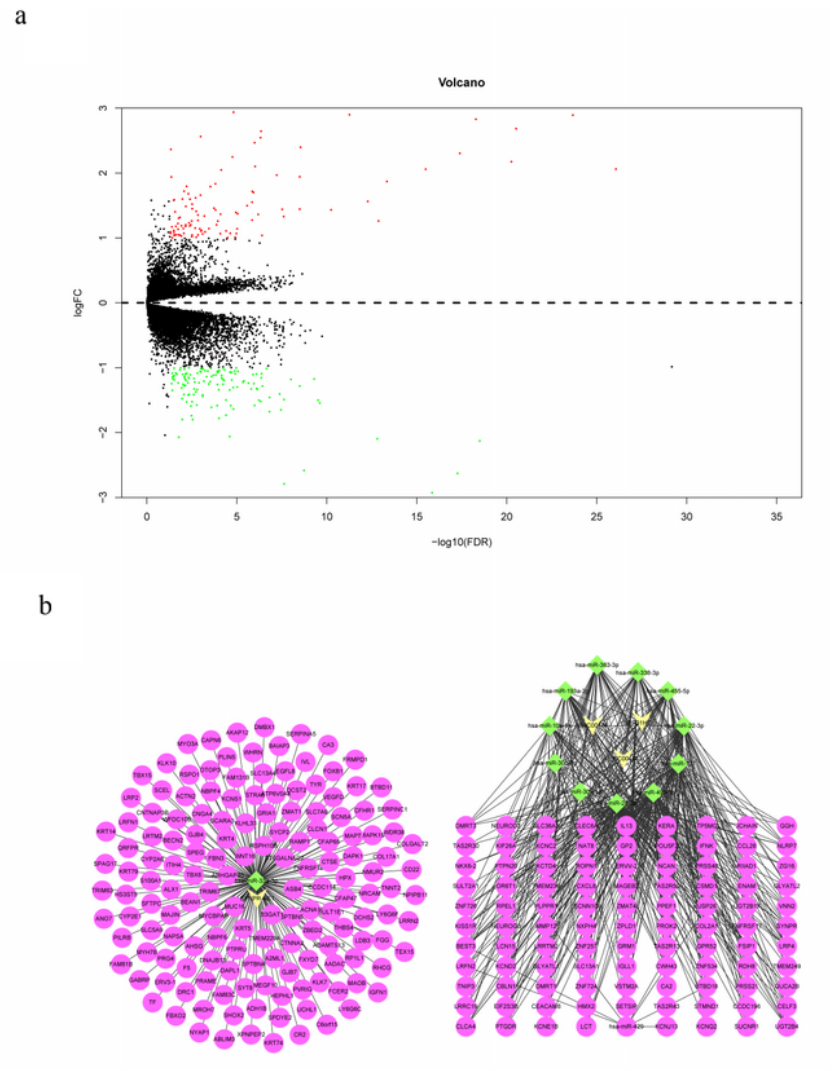

Figure 6

The ceRNA network of lncRNA-miRNA-mRNA in CRC. (A) The volcano plot showed that a total of 6407 DEGs were screened out between high- and low-risk groups. Green and red represent downregulated and upregulated genes, respectively. (B) LncRNA–miRNA–mRNA interactions in CRC. The fishtail triangle represents lncRNA, the diamond represents miRNA, and the circle represents mRNA.