



An Integrated Three-Long Non-coding RNA Signature Predicts Prognosis in Colorectal Cancer Patients

Yuhang Liu^{1,2}, Bingxin Liu^{1,2}, Guoying Jin^{1,2}, Jia Zhang^{1,2}, Xue Wang², Yuyang Feng², Zehua Bian^{1,2}, Bojian Fei³, Yuan Yin^{1,2*} and Zhaohui Huang^{1,2*}

¹ Wuxi Cancer Institute, Affiliated Hospital of Jiangnan University, Wuxi, China, ² Laboratory of Cancer Epigenetics, Wuxi School of Medicine, Jiangnan University, Wuxi, China, ³ Department of Surgical Oncology, Affiliated Hospital of Jiangnan University, Wuxi, China

OPEN ACCESS

Edited by:

Andrew Zloza,
Rush University Medical Center,
United States

Reviewed by:

Praveen Bommarreddy,
Rutgers, The State University of New
Jersey, United States
Shanmugarajan Krishnan,
Massachusetts General Hospital,
United States

*Correspondence:

Zhaohui Huang
hzhwxsy@126.com
Yuan Yin
yinyuandiana@163.com

Specialty section:

This article was submitted to
Cancer Molecular Targets and
Therapeutics,
a section of the journal
Frontiers in Oncology

Received: 28 July 2019

Accepted: 04 November 2019

Published: 22 November 2019

Citation:

Liu Y, Liu B, Jin G, Zhang J, Wang X,
Feng Y, Bian Z, Fei B, Yin Y and
Huang Z (2019) An Integrated
Three-Long Non-coding RNA
Signature Predicts Prognosis in
Colorectal Cancer Patients.
Front. Oncol. 9:1269.
doi: 10.3389/fonc.2019.01269

Colorectal cancer (CRC) is one of the most common cancers worldwide, whose morbidity and mortality gradually increased. Here, we aimed to identify and access prognostic long non-coding RNAs (lncRNAs) associated with overall survival (OS) in CRC. Firstly, RNA expression profiles were obtained from The Cancer Genome Atlas (TCGA) database, and 439 CRC patients were enrolled as a training set. Univariate Cox analysis and the least absolute shrinkage and selection operator analysis (LASSO) were performed to identify the prognostic lncRNAs. Multivariable Cox regression analysis was used to establish a prognostic risk formula including three lncRNAs (AP003555.2, AP006284.1, and LINC01602). The low-risk group had a better OS than the high-risk group ($P < 0.0001$), and the areas under the receiver operating characteristic curve (AUCs) of 3- and 5-year OS were 0.712 and 0.674, respectively. Then, we evaluated the signature in a clinical validation set which were collected from the Affiliated Hospital of Jiangnan University. Compared with the low-risk group, patients' OS were found to be significantly worse in the high-risk group ($P = 0.0057$). The AUCs of 3- and 5-year OS were 0.701 and 0.694, respectively. Finally, we constructed an lncRNA–microRNA (miRNA)–messenger RNA (mRNA) competing endogenous RNA (ceRNA) network to explore the potential function of three differentially expressed lncRNAs (DElncRNAs). The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis indicated that these DElncRNAs were involved with several cancer-related pathways. In summary, our data provide evidence that the three-lncRNA signature could serve as an independent biomarker to predict prognosis in CRC. This study will also suggest that these three lncRNAs potentially participate in the progression of CRC.

Keywords: long non-coding RNA, overall survival, prognosis, biomarkers, training set, clinical validation set

INTRODUCTION

Colorectal cancer (CRC) is a worldwide public health problem with the fourth highest incidence and second highest mortality in malignant tumors (1). The occurrence and progression of CRC were associated with multiple and complex factors, including lifestyle (2, 3), living surroundings (4), and genetic and epigenetic alterations (4, 5). Although therapeutic methods for CRC have largely

improved in recent years, prognosis remains unsatisfactory due to individual differences (6). Thus, it is necessary to find sensitive and precise biomarkers to better predict survival and prognosis of CRC patients.

Many studies showed that genes are no longer a piece of DNA that encodes a protein but also transcribes RNA that is not translated into protein (7–9). With the development of sequencing technology, abundant novel non-coding RNAs have been found (10). Long non-coding RNAs (lncRNAs), a kind of non-coding RNAs with length greater than 200 nucleotides (11, 12), were initially considered as the “noise” of genome transcription, a by-product of RNA polymerase II transcription, and had no biological function. Nevertheless, accumulating researches manifested that lncRNA expression or dysfunction was associated with a variety of major human diseases, including cancers (13–17). Rigoutsos et al. found that lncRNA N-BLR could promote cell migration and facilitate CRC invasion (18). Our previous study identified SNHG6 as an oncogene and predicted poor prognosis in CRC (19); we also found that the lncRNA LINC00152 could promote CRC cell proliferation and metastasis (20). Another study from our group suggested that lncRNA-FEZ1-AS1 promoted CRC cell proliferation and metastasis (21). Ozawa et al. identified lncRNAs CCAT1 and CCAT2, regarded as prognostic biomarkers in CRC (22). However, the biological function of lncRNAs in CRC is still not well-known, and a systematic study incorporating multiple lncRNAs’ expression is still needed for improving the predictive accuracy of CRC prognosis.

Since a competing endogenous RNA (ceRNA) hypothesis, microRNAs (miRNAs) and lncRNAs could interact with each other through miRNA response elements (MREs), was proposed in 2011 (23), more studies have suggested that an lncRNA–miRNA–messenger RNA (mRNA) network associated with the pathomechanism of multiple cancers (24–29). lncRNAs could act as sponges or ceRNAs of miRNAs to regulate tumorigenesis and metastasis of CRC (30–37). Many studies have suggested that ceRNAs participate in prognosis of different cancers. What’s more, studies on the mechanism of ceRNAs have potential clinical diagnostic and prognostic value. They could be potential prognostic biomarkers or novel therapeutic targets.

In this study, we identified three potential prognostic lncRNAs in CRC and confirmed the integrated three-lncRNA signature as an independent prognostic biomarker in the clinical validation set. Furthermore, we constructed an lncRNA–miRNA–mRNA network to investigate possible biological functions of lncRNAs in CRC.

MATERIALS AND METHODS

Data Treating

The data of RNA expression profiles and clinical information for CRC were downloaded from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/repository>), including 482 CRC tissues and 42 non-tumor tissues. lncRNAs were annotated by human gene annotation files (GRCh38.p12), which was downloaded from the Ensembl database (<https://asia.ensembl.org/index.html>). The expression data of lncRNAs were

analyzed by the R/Bioconductor package edgeR. Differentially expressed lncRNAs (DELncRNAs) were distinguished according to a $|\log_2 \text{fold change}| > 1.5$ and a $P\text{-value} < 0.05$. Then hierarchical clustering analysis was performed by the R package pheatmap. A volcano plot was drawn using the R software.

Characterization of lncRNAs Associated With Overall Survival

We excluded CRC patients from TCGA dataset according to the following criteria: (1) patients without information of survival status and survival time; (2) patients without complete lncRNA expression data. Finally, 439 CRC patients were selected in our study. Among these patients’ tissues, 42 CRC tissues and their paired adjacent non-cancerous tissues were used to identify DELncRNAs. And all the 439 patients were grouped into the training set.

Univariate Cox analysis was performed to assess the association between the expression levels of DELncRNAs and the overall survival (OS) of patients, and lncRNAs with a $P\text{-value} < 0.05$ were selected for further analysis. Then the least absolute shrinkage and selection operator (LASSO) model was used to find vital lncRNAs from the prognostic DELncRNAs. The LASSO method was utilized by the package “glmnet” in the R (version 3.5.1) software. Then Kaplan–Meier analysis was performed to single out lncRNAs significantly associated with the OS from DELncRNAs, which were selected by the LASSO method.

Establishment of Prognostic Risk Score Formula

In light of the expression level of lncRNAs and regression coefficients, a prognostic risk formula was established by multivariable Cox regression analysis. The random forest plot was performed using the R package survminer. The risk scores of each patient were calculated by the formula as mentioned above. Finally, all patients were divided into a high-risk group and a low-risk group by utilizing the median risk score as the cutoff value.

Assessing the Prognostic Risk Score Model

Next, a Kaplan–Meier survival curve was used to evaluate the prognosis between the low-risk group and high-risk group. And a time-dependent receiver operating characteristic (ROC) curve was performed to assess the diagnostic accuracy based on the risk score for 3- and 5-year OS probability. A $P\text{-value} < 0.05$ was recognized as statistically significant. A concordance index (C-index) was calculated to estimate the value of prognostic risk formula by the R package survcomp. The expression profiles of key lncRNAs in the high-risk and low-risk groups were plotted by risk heatmap. The above analyses were employed using R (version 3.5.1).

Clinical Samples and Quantitative RT-PCR Assay

In the clinical validation set, we chose CRC patients based on the following criteria: (1) patients treated in the Affiliated Hospital

of Jiangnan University; (2) patients with complete follow-up information, including survival status and survival time; (3) patients who did not receive treatment before surgery. Eighty-five CRC patients were gained from the Affiliated Hospital of Jiangnan University as the clinical validation set. The clinical information of all patients was listed in **Table 1**. All patients signed the informed consent about using their tumor tissues, and this study was ratified by the clinical research ethics committees of the participating institutions.

Total RNA of tumor tissues was isolated using RNAiso Plus (Takara, Japan). NanoDrop 2000 (Thermo, USA) was used to measure RNA concentrations. Total RNA was reverse transcribed to complementary DNA (cDNA) using the PrimeScript II 1st Strand Synthesis Kit (TaKaRa). The expression levels of lncRNAs were quantitated using the UltraSYBR Mixture (CWBIO, China) by the ViiA7 real-time PCR system. The lncRNA expression level was calculated as follows: $\Delta CT = CT(\text{lncRNA}) - CT(\beta\text{-actin})$. The sequences of used primers are listed in **Supplementary Table 1**.

Construction of the CeRNA Network

The miRNAs which potentially interacted with three DElncRNAs were predicted using LncBOOK (<http://bigd.big.ac.cn/lncbook>), a curated information database of human lncRNAs. Finally,

a total of 1,039 miRNAs were identified. Next, according to the score of interaction between lncRNAs and miRNAs, we selected the top 30 closely related miRNAs of each lncRNAs for further study. Then the target mRNAs of miRNA were predicted by utilizing miRDB (<http://www.mirdb.org/>), TargetScan (<http://www.targetscan.org/>), and miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw/php/index.php>). Finally, we established the lncRNA–miRNA–mRNA network by using the Cytoscape software (version 3.6.1).

Functional Prediction

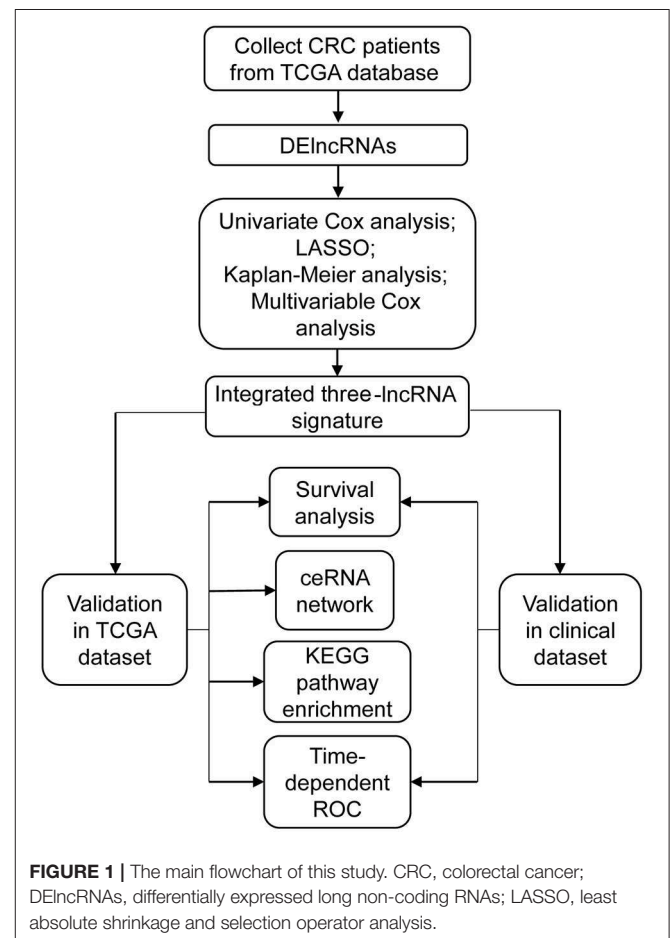
The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed to describe the potential function of differentially expressed mRNAs (DEmRNAs) by the package “clusterProfiler” (<http://bioconductor.org/packages/clusterProfiler/>) in the R (version 3.5.1) software.

Statistical Analysis

Both the expression profiles of RNAs and clinical information data were excavated from TCGA database by the Perl (version 5.24.3) software and R (version 3.5.1) software. All statistical analyses were carried out using SPSS23.0 (SPSS, Chicago, IL, United States) or R software. A *P*-value < 0.05 was regarded as statistically significant.

TABLE 1 | Clinicopathological characteristics of 85 patients with colorectal cancer (CRC) in the validation set.

Characteristics	CRC patients (N = 85)	
	n	%
Age (years)		
≥60	47	55.29
<60	38	44.71
Gender		
Male	47	55.29
Female	38	44.71
AJCC stage		
I	23	27.05
II	36	42.35
III	26	30.60
IV	0	
AJCC-T		
T1	4	4.70
T2	22	25.88
T3	39	45.88
T4	20	23.54
AJCC-N		
N0	59	69.41
N1	20	23.53
N2	6	7.06
N3	0	0
AJCC-M		
M0	84	98.82
M1	1	1.18



RESULTS

DElncRNAs in CRC

The overview of the screening strategy used in this study is displayed in **Figure 1**. In the 439 tissues of CRC patients, 42 CRC tissues and their paired adjacent non-cancerous tissues were utilized to screen DElncRNAs. Then 580 DElncRNAs, with a critical point of $|\logFC > 1.5|$ and P -value < 0.05 , were gained by using the R/Bioconductor package edgeR. Among the DElncRNAs, 319 DElncRNAs are upregulated and 261 DElncRNAs are downregulated (**Supplementary Table 2**). The heatmap and volcano plot were shown respectively in **Figures 2A,B**.

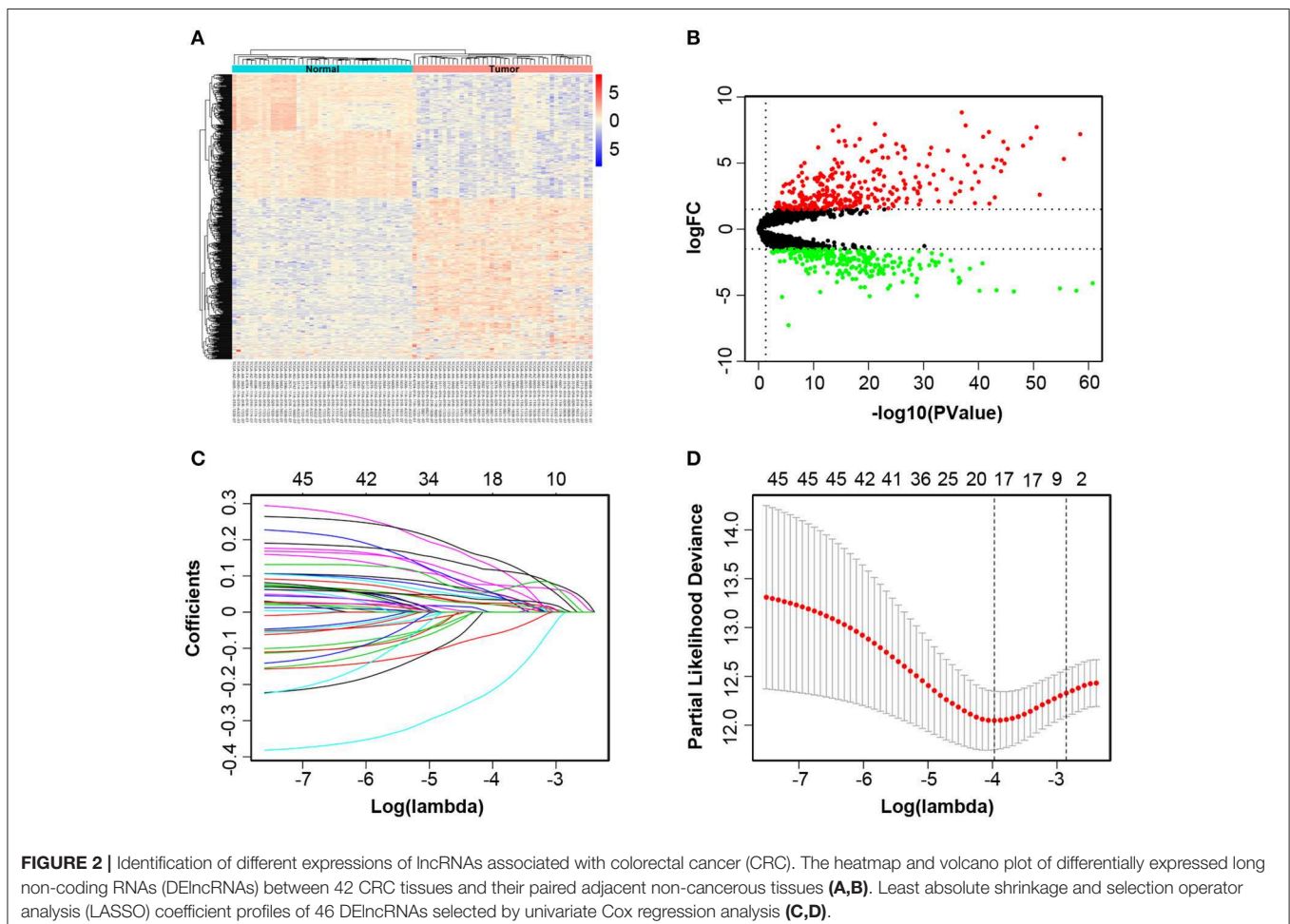
Recognition of Key lncRNAs Correlated With OS

All of 439 CRC patients collected from TCGA database were regarded as the training set. Next, we performed univariate Cox regression analysis to estimate the prognostic relationship between lncRNA expression profiles and patient OS. Among 580 DElncRNAs, 46 lncRNAs with a P -value < 0.05 were selected for the following study. Next, the LASSO method was performed, and the coefficients of 46 lncRNAs were shown in

Figure 2C, and the minimize λ method screened out 17 lncRNAs (**Figure 2D**). We then used Kaplan-Meier survival curves to further analyze the relationship between the 17 lncRNAs and the OS of patients. Finally, we found that 10 lncRNAs were observably related with OS (**Figure 3**), and the other seven lncRNAs were not significant (**Supplementary Figure 1**). Using multivariable Cox regression analysis, three lncRNAs (AP003555.2, AP006284.1, and LINC01602) were identified as integrated prognostic biomarkers for CRC patients (**Figure 4A**). The results showed that all lncRNAs play as hazards with positive coefficients. The relative expression levels of the three lncRNAs between 42 CRC tissues and their paired adjacent non-cancerous tissues were displayed in **Supplementary Figure 2**.

Establishment of Prognostic Risk Score Formula in the Training Set

We then established a prognostic risk score formula in the training set based on the expression profiles of three prognostic lncRNAs and their regression coefficients. The prognostic risk score formula was as follows: risk score = $0.2212 * (\text{the expression level of AP003555.2}) + 0.2081 * (\text{the expression level of AP006284.1}) + 0.1214 * (\text{the expression level of LINC01602})$.



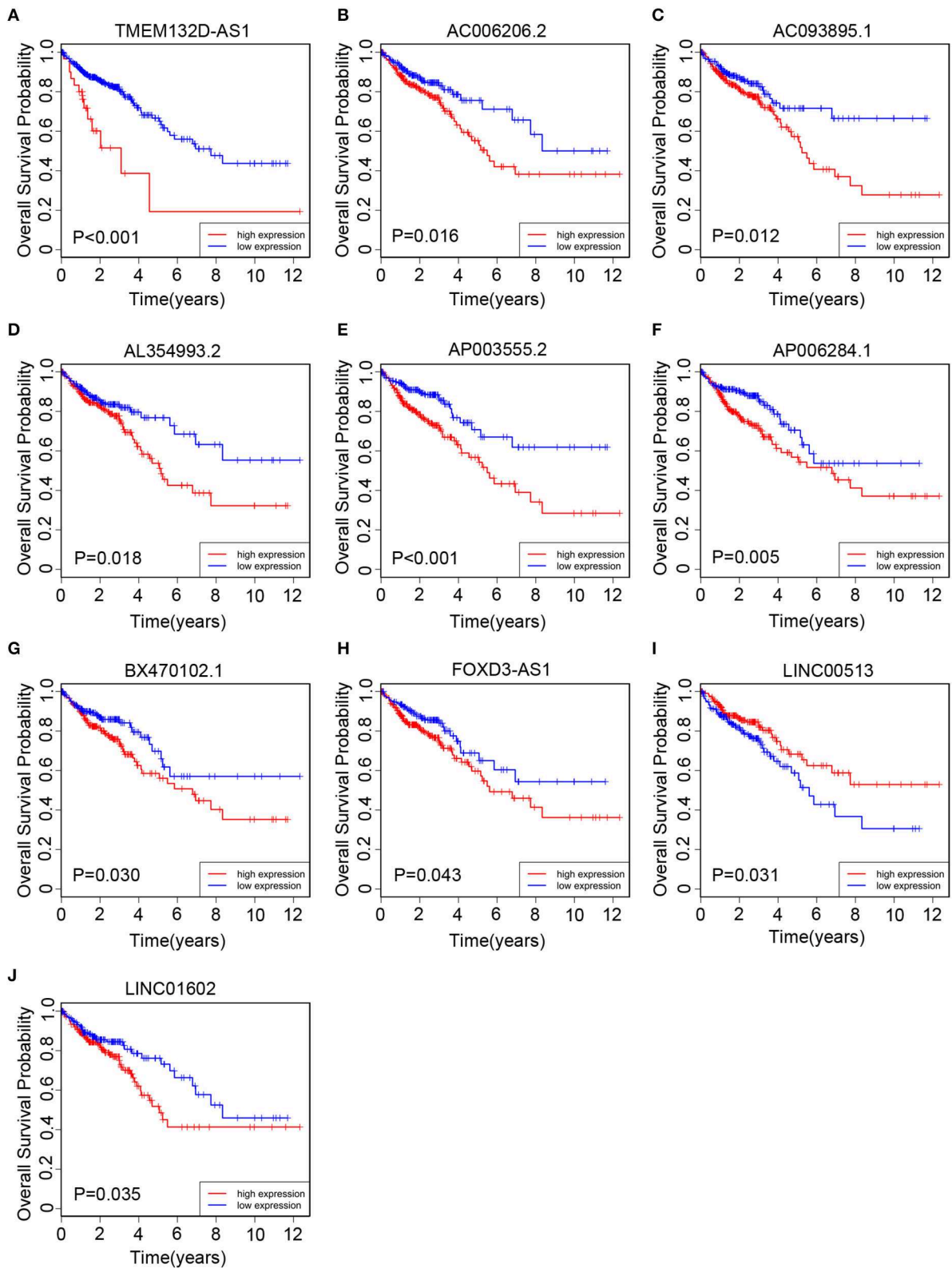
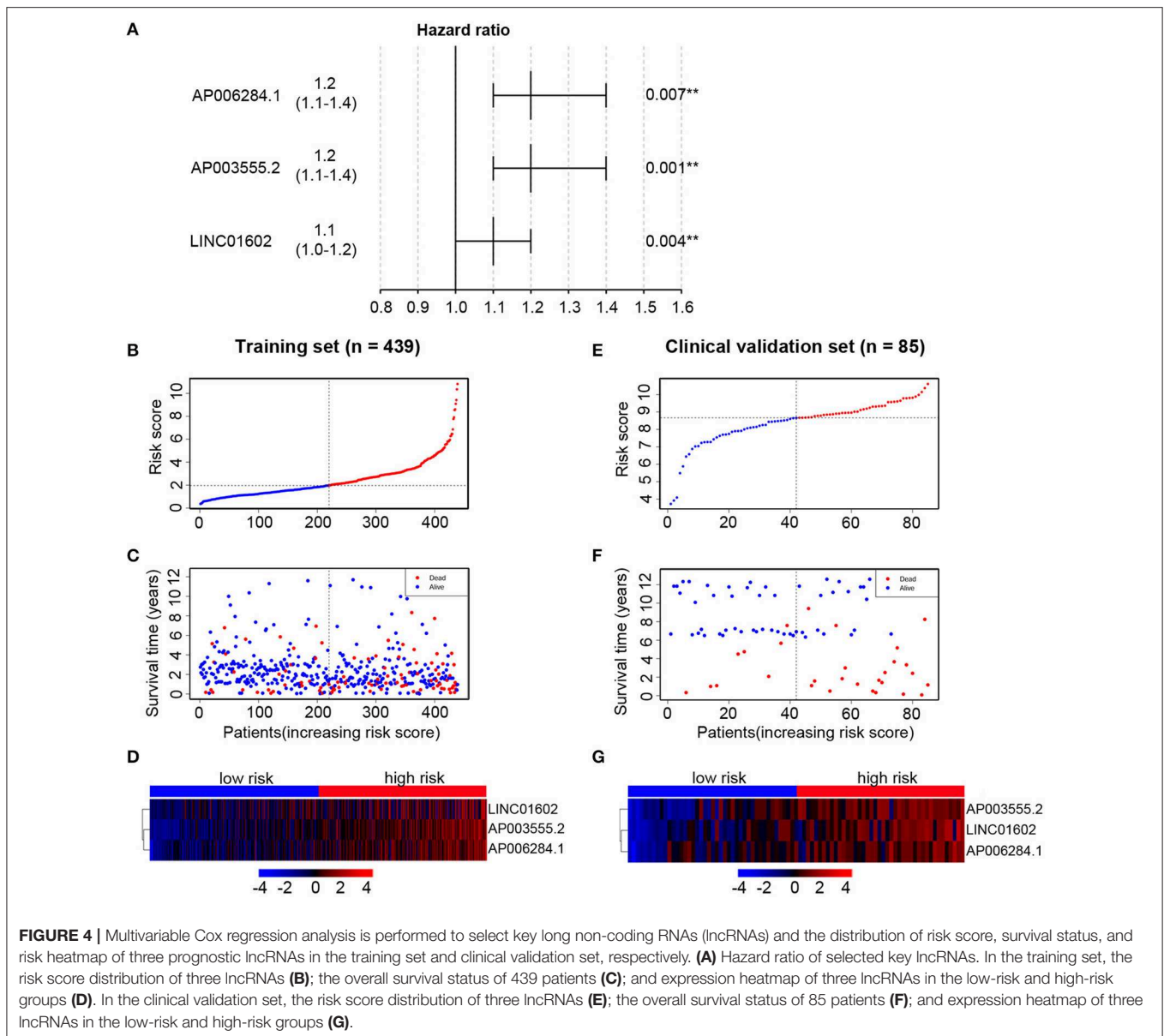


FIGURE 3 | The Kaplan–Meier curve of 10 prognostic lncRNAs in CRC patients collected from The Cancer Genome Atlas (TCGA) cohort. The Kaplan–Meier curve for **(A)** TMEM132D-AS1, **(B)** AC006206.2, **(C)** AC093895.1, **(D)** AL354993.2, **(E)** AP003555.2, **(F)** AP006284.1, **(G)** BX470102.1, **(H)** FOXD3-AS1, **(I)** LINC00513, and **(J)** LINC01602.



The risk scores were calculated for all patients and split patients into a high-risk group ($n = 219$) and a low-risk group ($n = 220$) by using the median risk score as the cutoff value. The distribution of risk scores and survival status of the patients were shown in **Figures 4B,C**. The expression profiles of three lncRNAs were shown using a risk heatmap in 439 ordered patients (**Figure 4D**). These results suggested that patients with higher risk scores currently had worse OS than those with lower risk scores.

Prognostic Value of the Three lncRNAs Respectively in the Clinical Validation Set

We collected 85 CRC patients from the Affiliated Hospital of Jiangnan University as the clinical validation set. Next, we quantitated the relative expression levels of the three lncRNAs

in each patient's cancer tissue by quantitative RT-PCR assay and used Kaplan-Meier analysis to evaluate the prognostic capacity. As a result, we found that the high-expression group had a worse OS than the low-expression group based on the median expression level as the cutoff value for each lncRNA (**Figures 5A-C**). The results were in accord with those of TCGA cohort (**Figures 3E,F,J**).

Prognostic Value of the Three-lncRNA Signature in the Training and Clinical Validation Sets

Subsequently, we assessed the prognostic value of the above risk formula in the training set by using Kaplan-Meier analysis. We found that the low-risk group had a better OS than the high-risk group ($P < 0.0001$) (**Figure 6A**).

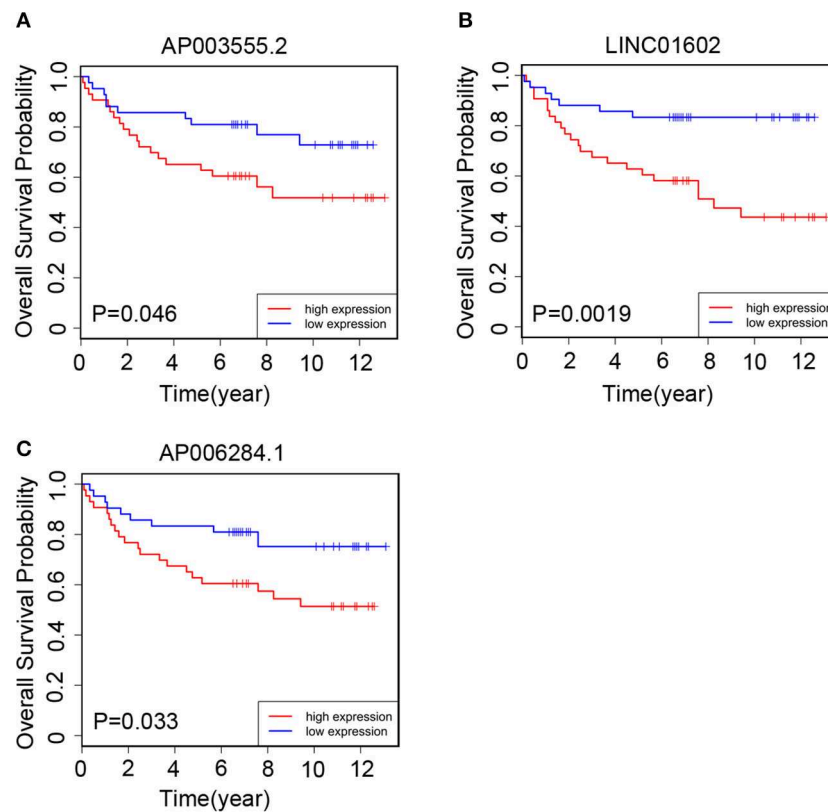


FIGURE 5 | The Kaplan–Meier curve of three prognostic long non-coding RNAs (lncRNAs) respectively in the clinical validation set. In the validation set, the Kaplan–Meier curve for **(A)** AP003555.2, **(B)** LINC01602, and **(C)** AP006284.1.

Moreover, time-dependent ROC analysis was also utilized to evaluate the prognostic capacity of the risk formula. The areas under the ROC curve at 3 and 5 years were 0.712 and 0.674, respectively, which suggested that the integrated three-lncRNA signature had better utility than each single one (**Figure 6B**). The value of the C-index was 0.65 (95% CI: 0.59–0.72), showing a fine prognostic value on predicting patients' survival.

In order to confirm the common prognostic value of the integrated three-lncRNA signature in different patient cohorts, our risk model was assessed in the clinical validation set. Eighty-five patients were divided into a high-risk group ($n = 43$) and a low-risk group ($n = 42$) by using the above three-lncRNA signature.

Compared with that in the low-risk group, patients' OS was found to be significantly worse in the high-risk group. The risk scores were calculated, and the distribution were shown in **Figures 4E,F**. A risk heatmap in 85 ordered patients was shown in **Figure 4G**. As expected, Kaplan-Meier analysis suggested that the low-risk group had a better OS than the high-risk group ($P = 0.0057$) (**Figure 6C**). Time-dependent ROC analysis indicated that the areas under the ROC curve at 3 and 5 years were 0.701 and 0.694, respectively (**Figure 6D**). These results were consistent with those of the training set.

Function of Three lncRNAs

To explore the potential roles of the three lncRNAs in CRC, we constructed an interactional network of lncRNA–miRNA–mRNA. Firstly, utilizing the LncBOOK online analysis tool, we predicted the target relationship among three lncRNAs and miRNAs. We then chose the top 30 target miRNAs of each lncRNA, according to the score of interaction, to predict the downstream target mRNAs by using miRDB, TargetScan, and miRTarBase. In the end, we constructed a ceRNA network including 3 lncRNAs, 90 miRNAs, and 1,791 mRNAs (**Supplementary Figure 3A**).

Next, we employed the KEGG pathway analysis to explore the potential mechanisms involved in the progression of CRC using the package ClusterProfiler in the R software based on the selected mRNAs. The top 15 KEGG pathways were shown in **Supplementary Figure 3B**. Some key cancer-related pathways were found, such as the PI3K–Akt pathway, TGF- β pathway, and CRC and P53 pathways.

DISCUSSION

Currently, a growing number of studies have demonstrated that dysregulated lncRNAs are involved in the progression of CRC and might be potential biomarkers for diagnosis, treatment,

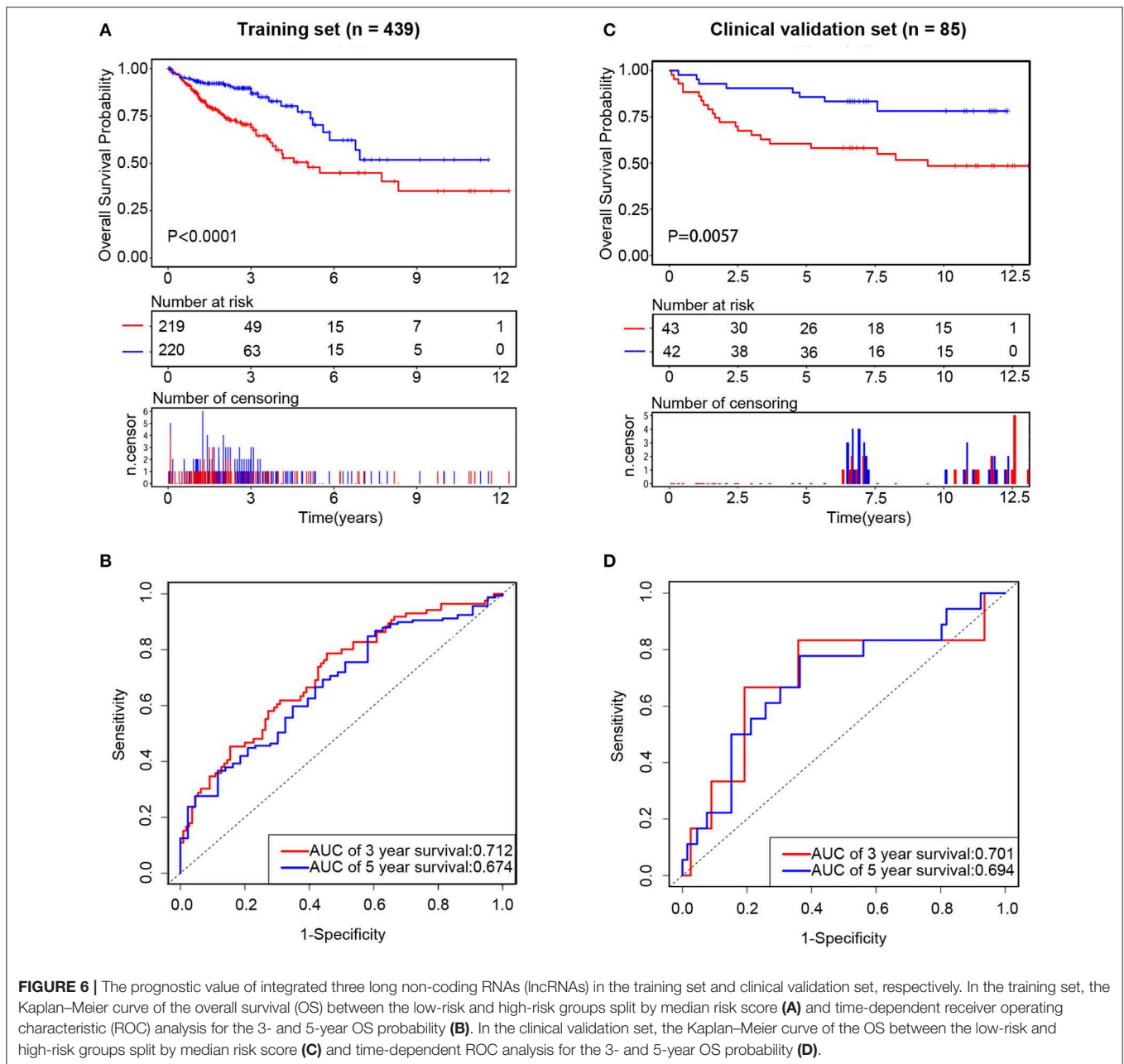


FIGURE 6 | The prognostic value of integrated three long non-coding RNAs (lncRNAs) in the training set and clinical validation set, respectively. In the training set, the Kaplan–Meier curve of the overall survival (OS) between the low-risk and high-risk groups split by median risk score (A) and time-dependent receiver operating characteristic (ROC) analysis for the 3- and 5-year OS probability (B). In the clinical validation set, the Kaplan–Meier curve of the OS between the low-risk and high-risk groups split by median risk score (C) and time-dependent ROC analysis for the 3- and 5-year OS probability (D).

and prognostic judgment of the disease. Compared with single biomarkers, a system of multiple integrated biomarkers can improve predictive accuracy. In our present study, a novel three-lncRNA signature was constructed for predicting OS of CRC patients. Using multiple appropriate statistical methods, three lncRNAs (AP003555.2, AP006284.1, and LINC01602) were identified as independent prognostic indicators. Then, an integrated three-lncRNA signature was constructed based on their expression profiles and regression coefficients. The value of prognosis was confirmed by two independent patients' cohort. In TCGA training set, patients were divided into low-risk and high-risk groups based on the prognostic score calculated by the

above lncRNA signature. Survival curves indicated significant differences in the patients' OS between the two groups. Time-dependent ROC analysis indicated that the lncRNA signature had high predictive accuracy in predicting OS. A clinical validation set was used to verify the universal applicability of the model.

All the three lncRNAs were already annotated by the Ensembl database. However, little was known about their roles in tumorigenesis and progression. Zhang et al. found that a decreased expression of LINC01602 was associated with worse survival in patients with rectal adenocarcinoma (38). No public reports were found for the other two lncRNAs (AP003555.2 and AP006284.1) according to a PubMed search. All three

lncRNAs were upregulated in CRC tissues compared with normal tissues. Meanwhile, a high expression of each DELncRNA was significantly associated with poor prognosis in CRC. Importantly, multivariable Cox regression analysis confirmed their prognostic role for CRC. Our results suggest that the three DELncRNAs may play crucial roles in the pathomechanism of CRC and act as potential prognostic biomarkers.

In contrast with previous studies, the integrated three-lncRNA signature reported here was distinctly different. First of all, we used 42 CRC tissues and their paired adjacent non-cancerous tissues selected from TCGA database as self-control to analyze DELncRNAs. In this way, it is effective to enhance the balance between groups, control confounders, and improve the accuracy of research. The most important point was that we collected CRC patients from the Affiliated Hospital of Jiangnan University as the clinical validation set to assess the three-lncRNA signature. Previous studies usually constructed a prognostic model as follows (39–43): (1) using one GEO database to build a prognostic model and using another GEO database or TCGA database as a validation set to evaluate the constructed model; (2) using TCGA database to establish a training model and using a GEO database to validate; and (3) dividing TCGA database randomly into two data sets equally or unequally, one as a training set and the other as a validation set. Generally, most studies were based on the online data from publicly available databases without validating an additional clinical patient cohort. In other words, there was a lack of evidence that the results of bioinformatic analysis, such as RNA-Seq, were well-validated in external patients' cohort by quantitative RT-PCR assay. In our study, we used TCGA data as the training set to establish an lncRNA model associated with OS. We then collected 85 CRC patients as the clinical validation set to evaluate whether the model was also appropriate in clinical samples. Interestingly, the results showed that the model constructed from TCGA dataset had a common prognostic value in the clinical patients' cohort.

CeRNA hypothesis revealed that lncRNAs could competitively bind miRNAs, as a "sponge" of miRNAs, and then indirectly regulate the downstream target genes of miRNAs (23). The lncRNA SNHG6 could sponge miR-26a/b and miR-214 to regulate EZH2 expression and thus promote CRC cell growth, migration, and invasion (30). An oncogenic lncRNA, NEAT1, was a "sponge" in CRC to competitively bind miR-193a-3p and thus upregulate the expression of the downstream gene *KRAS* (31). Liu et al. found that linc01296 functions as a sponge of miR-26a to regulate the expression of its target gene *MUC1* and regulate the activity of the PI3K/AKT signal pathway (35). In order to explore the potential biological function of the three lncRNAs, we constructed an lncRNA–miRNA–mRNA ceRNA network. We predicted the target miRNAs of each lncRNA and used the intersection of three databases (miRDB, TargetScan, and miRTarBase) to predict target mRNAs. KEGG pathway analysis indicated that the three lncRNAs' functions were potentially associated with several cancer-related pathways, such as the "PI3K-Akt signaling pathway," "TGF- β signaling pathway," "Colorectal cancer," and "P53 signaling pathway." Current studies have shown that the PI3K-Akt pathway is activated in many types of cancers. Many studies have indicated

that the PI3K-Akt pathway plays a vital role in cell proliferation, migration, and invasion in CRC (35, 44, 45). Meanwhile, several studies have indicated that the TGF- β signaling pathway was also associated with lncRNAs; for example, the lncRNA SNHG6 could accelerate cell growth, invasion, and migration via activating the TGF- β signaling pathway in CRC (46). Here, we constructed an interactional network of lncRNA–miRNA–mRNA which may be involved in the progression of CRC, but the potential mechanisms still need to be investigated in future studies.

Although a three-lncRNA signature was constructed and appears to be a potential prognostic biomarker in clinical application, there are also some limitations. First, the sample size in our clinical validation set was not large enough. The second limitation was that the prognostic value of the three-lncRNA panel is not very satisfactory. Third, the binding affinities between lncRNAs and miRNAs, miRNAs, and mRNAs were predicted only by online databases and should be further experimentally investigated. Fourth, the biological functions of the three DELncRNAs were not assessed in this study.

In conclusion, we constructed an integrated three-lncRNA signature that was significantly associated with OS in CRC patients which could accurately identify patients with low prognostic risk from those with high prognostic risk. Furthermore, we evaluated the accuracy and reliability of the above signature in our own clinical validation set. These results suggested that the integrated three-lncRNA signature could potentially act as a prognostic model in CRC.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://portal.gdc.cancer.gov>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Affiliated Hospital of Jiangnan University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZH, YY, and YL designed and conceived the study. YL and BL performed majority of analyses for the study. GJ, JZ, and XW downloaded the data from an online database. YF and ZB performed the experiments. YL and BF analyzed the data. YL and YY wrote the manuscript.

FUNDING

This study was partially supported by grants from the National Natural Science Foundation of China (81972220, 81672328, and 81772636), Medical Key Professionals Program of Jiangsu Province (AF052141), Medical Youth Professionals Program of Jiangsu Province (QNRC2016162), Medical Innovation Team Program of Wuxi (CXTP003),

Fundamental Research Funds for the Central Universities (NOJUSRP51619B), National First-class Discipline Program of Food Science and Technology (JUFSTR20180101), and Project of the Wuxi Health and Family planning Commission (Z201806).

REFERENCES

- 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) 68:394–424. doi: 10.3322/caac.21492
- Roncucci L, Mariani F. Prevention of colorectal cancer: how many tools do we have in our basket? *Eur J Intern Med.* (2015) 26:752–6. doi: 10.1016/j.ejim.2015.08.019
- Bode AM, Dong Z, Wang H. Cancer prevention and control: alarming challenges in China. *Natl Sci Rev.* (2016) 3:117–27. doi: 10.1093/nsr/nwv054
- Okugawa Y, Grady WM, Goel A. Epigenetic alterations in colorectal cancer: emerging biomarkers. *Gastroenterology.* (2015) 149:1204–25.e12. doi: 10.1053/j.gastro.2015.07.011
- van Engeland M, Derks S, Smits KM, Meijer GA, Herman JG. Colorectal cancer epigenetics: complex simplicity. *J Clin Oncol.* (2011) 29:1382–91. doi: 10.1200/jco.2010.28.2319
- Li Y, He M, Zhou Y, Yang C, Wei S, Bian X, et al. The prognostic and clinicopathological roles of PD-L1 expression in colorectal cancer: a systematic review and meta-analysis. *Front Pharmacol.* (2019) 10:139. doi: 10.3389/fphar.2019.00139
- Lin L, Shen S, Tye A, Cai JJ, Jiang P, Davidson BL, et al. Diverse splicing patterns of exonized Alu elements in human tissues. *PLoS Genet.* (2008) 4:e1000225. doi: 10.1371/journal.pgen.1000225
- Khurana E, Fu Y, Colonna V, Mu XJ, Kang HM, Lappalainen T, et al. Integrative annotation of variants from 1092 humans: application to cancer genomics. *Science.* (2013) 342:1235587. doi: 10.1126/science.1235587
- Stojic L, Niemczyk M, Orjalo A, Ito Y, Ruijter AE, Uribe-Lewis S, et al. Transcriptional silencing of long noncoding RNA GNG12-AS1 uncouples its transcriptional and product-related functions. *Nat Commun.* (2016) 7:10406. doi: 10.1038/ncomms10406
- Kietrys AM, Velema WA. Fingerprints of modified RNA bases from deep sequencing profiles. *J Am Chem Soc.* (2017) 139:17074–81. doi: 10.1021/jacs.7b07914
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* (2007) 129:1311–23. doi: 10.1016/j.cell.2007.05.022
- St. Laurent G, Wahlestedt C, Kapranov P. The landscape of long noncoding RNA classification. *Trends Genet.* (2015) 31:239–51. doi: 10.1016/j.tig.2015.03.007
- Schmitt AM, Chang HY. Long noncoding RNAs in cancer pathways. *Cancer Cell.* (2016) 29:452–63. doi: 10.1016/j.ccell.2016.03.010
- Zhang Y, Pitchiaya S, Cieslik M, Niknafs YS, Tien JC, Hosono Y, et al. Analysis of the androgen receptor-regulated lncRNA landscape identifies a role for ARLNC1 in prostate cancer progression. *Nat Genet.* (2018) 50:814–24. doi: 10.1038/s41588-018-0120-1
- Kim J, Piao HL, Kim BJ, Yao F, Han Z, Wang Y, et al. Long noncoding RNA MALAT1 suppresses breast cancer metastasis. *Nat Genet.* (2018) 50:1705–15. doi: 10.1038/s41588-018-0252-3
- Zhuo W, Liu Y, Li S, Guo D, Sun Q, Jin J, et al. Long noncoding RNA GMAN, up-regulated in gastric cancer tissues, is associated with metastasis in patients and promotes translation of Ephrin A1 by competitively binding GMAN-AS. *Gastroenterology.* (2019) 156:676–91.e11. doi: 10.1053/j.gastro.2018.10.054
- Zhang J, Bian Z, Jin G, Liu Y, Li M, Yao S, et al. Long non-coding RNA IQCJ-SCHIP1 antisense RNA 1 is downregulated in colorectal cancer and inhibits cell proliferation. *Ann Transl Med.* (2019) 7:198. doi: 10.21037/atm.2019.04.21
- Rigoutsos I, Lee SK, Nam SY, Anfossi S, Pasculli B, Pichler M, et al. N-BLR, a primate-specific non-coding transcript leads to colorectal cancer invasion and migration. *Genome Biol.* (2017) 18:98. doi: 10.1186/s13059-017-1224-0
- Li M, Bian Z, Yao S, Zhang J, Jin G, Wang X, et al. Up-regulated expression of SNHG6 predicts poor prognosis in colorectal cancer. *Pathol Res Pract.* (2018) 214:784–9. doi: 10.1016/j.prp.2017.12.014
- Bian Z, Zhang J, Li M, Feng Y, Yao S, Song M, et al. Long non-coding RNA LINC00152 promotes cell proliferation, metastasis, and confers 5-FU resistance in colorectal cancer by inhibiting miR-139-5p. *Oncogenesis.* (2017) 6:395. doi: 10.1038/s41389-017-0008-4
- Bian Z, Zhang J, Li M, Feng Y, Wang X, Zhang J, et al. LncRNA-FEZF1-AS1 promotes tumor proliferation and metastasis in colorectal cancer by regulating PKM2 signaling. *Clin Cancer Res.* (2018) 24:4808–19. doi: 10.1158/1078-0432.ccr-17-2967
- Ozawa T, Matsuyama T, Toiyama Y, Takahashi N, Ishikawa T, Uetake H, et al. CCAT1 and CCAT2 long noncoding RNAs, located within the 8q.24.21 'gene desert', serve as important prognostic biomarkers in colorectal cancer. *Ann Oncol.* (2017) 28:1882–8. doi: 10.1093/annonc/mdx248
- Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* (2011) 146:353–8. doi: 10.1016/j.cell.2011.07.014
- Wang Y, Yang L, Chen T, Liu X, Guo Y, Zhu Q, et al. A novel lncRNA MCM3AP-AS1 promotes the growth of hepatocellular carcinoma by targeting miR-194-5p/FOXA1 axis. *Mol Cancer.* (2019) 18:28. doi: 10.1186/s12943-019-0957-7
- Wang M, Mao C, Ouyang L, Liu Y, Lai W, Liu N, et al. Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. *Cell Death Differ.* (2019) 26:2329–43. doi: 10.1038/s41418-019-0304-y
- Dong S, Wang R, Wang H, Ding Q, Zhou X, Wang J, et al. HOXD-AS1 promotes the epithelial to mesenchymal transition of ovarian cancer cells by regulating miR-186-5p and PIK3R3. *J Exp Clin Cancer Res.* (2019) 38:110. doi: 10.1186/s13046-019-1103-5
- Gao R, Zhang N, Yang J, Zhu Y, Zhang Z, Wang J, et al. Long non-coding RNA ZEB1-AS1 regulates miR-200b/FSCN1 signaling and enhances migration and invasion induced by TGF-beta1 in bladder cancer cells. *J Exp Clin Cancer Res.* (2019) 38:111. doi: 10.1186/s13046-019-1102-6
- Chen X, Chen Z, Yu S, Nie F, Yan S, Ma P, et al. Long noncoding RNA LINC01234 functions as a competing endogenous RNA to regulate CBFβ expression by sponging miR-204-5p in gastric cancer. *Clin Cancer Res.* (2018) 24:2002–14. doi: 10.1158/1078-0432.ccr-17-2376
- Zhou G, Chen X. Emerging role of extracellular microRNAs and lncRNAs. *ExRNA.* (2019) 1:10. doi: 10.1186/s41544-019-0012-2
- Xu M, Chen X, Lin K, Zeng K, Liu X, Xu X, et al. LncRNA SNHG6 regulates EZH2 expression by sponging miR-26a/b and miR-214 in colorectal cancer. *J Hematol Oncol.* (2019) 12:3. doi: 10.1186/s13045-018-0690-5
- Zhu Z, Du S, Yin K, Ai S, Yu M, Liu Y, et al. Knockdown long noncoding RNA nuclear paraspeckle assembly transcript 1 suppresses colorectal cancer through modulating miR-193a-3p/KRAS. *Cancer Med.* (2019) 8:261–75. doi: 10.1002/cam4.1798
- Wang L, Jiang F, Xia X, Zhang B. LncRNA FAL1 promotes carcinogenesis by regulation of miR-637/NUPR1 pathway in colorectal cancer. *Int J Biochem Cell Biol.* (2019) 106:46–56. doi: 10.1016/j.biocel.2018.09.015
- Xiong WC, Han N, Wu N, Zhao KL, Han C, Wang HX, et al. Interplay between long noncoding RNA ZEB1-AS1 and miR-101/ZEB1 axis regulates proliferation and migration of colorectal cancer cells. *Am J Transl Res.* (2018) 10:605–17.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.01269/full#supplementary-material>

34. Yu HM, Wang C, Yuan Z, Chen GL, Ye T. LncRNA NEAT1 promotes the tumorigenesis of colorectal cancer by sponging miR-193a-3p. *Cell Prolif.* (2019) 52:e12526. doi: 10.1111/cpr.12526
35. Liu B, Pan S, Xiao Y, Liu Q, Xu J, Jia L. LINC01296/miR-26a/GALNT3 axis contributes to colorectal cancer progression by regulating O-glycosylated MUC1 via PI3K/AKT pathway. *J Exp Clin Cancer Res.* (2018) 37:316. doi: 10.1186/s13046-018-0994-x
36. Bian Z, Jin L, Zhang J, Yin Y, Quan C, Hu Y, et al. LncRNA-UCA1 enhances cell proliferation and 5-fluorouracil resistance in colorectal cancer by inhibiting miR-204-5p. *Sci Rep.* (2016) 6:23892. doi: 10.1038/srep23892
37. Li M, Bian Z, Jin G, Zhang J, Yao S, Feng Y, et al. LncRNA-SNHG15 enhances cell proliferation in colorectal cancer by inhibiting miR-338-3p. *Cancer Med.* (2019) 8:2404–13. doi: 10.1002/cam4.2105
38. Zhang Z, Wang S, Ji D, Qian W, Wang Q, Li J, et al. Construction of a ceRNA network reveals potential lncRNA biomarkers in rectal adenocarcinoma. *Oncol Rep.* (2018) 39:2101–13. doi: 10.3892/or.2018.6296
39. Cheng P. A prognostic 3-long noncoding RNA signature for patients with gastric cancer. *J Cell Biochem.* (2018) 119:9261–9. doi: 10.1002/jcb.27195
40. Zhang L, Chen S, Wang B, Su Y, Li S, Liu G, et al. An eight-long noncoding RNA expression signature for colorectal cancer patients' prognosis. *J Cell Biochem.* (2019) 120:5636–43. doi: 10.1002/jcb.27847
41. Miao R, Ge C, Zhang X, He Y, Ma X, Xiang X, et al. Combined eight-long noncoding RNA signature: a new risk score predicting prognosis in elderly non-small cell lung cancer patients. *Aging.* (2019) 11:467–79. doi: 10.18632/aging.101752
42. Dai W, Feng Y, Mo S, Xiang W, Li Q, Wang R, et al. Transcriptome profiling reveals an integrated mRNA-lncRNA signature with predictive value of early relapse in colon cancer. *Carcinogenesis.* (2018) 39:1235–44. doi: 10.1093/carcin/bgy087
43. Gu JX, Zhang X, Miao RC, Xiang XH, Fu YN, Zhang JY, et al. Six-long non-coding RNA signature predicts recurrence-free survival in hepatocellular carcinoma. *World J Gastroenterol.* (2019) 25:220–32. doi: 10.3748/wjg.v25.i2.220
44. Hu J, Shan Y, Ma J, Pan Y, Zhou H, Jiang L, et al. LncRNA ST3Gal6-AS1/ST3Gal6 axis mediates colorectal cancer progression by regulating alpha-2,3 sialylation via PI3K/Akt signaling. *Int J Cancer.* (2019) 145:450–60. doi: 10.1002/ijc.32103
45. Yu S, Wang D, Shao Y, Zhang T, Xie H, Jiang X, et al. SP1-induced lncRNA TINCR overexpression contributes to colorectal cancer progression by sponging miR-7-5p. *Aging.* (2019) 11:1389–403. doi: 10.18632/aging.101839
46. Wang X, Lai Q, He J, Li Q, Ding J, Lan Z, et al. LncRNA SNHG6 promotes proliferation, invasion and migration in colorectal cancer cells by activating TGF-beta/Smad signaling pathway via targeting UPF1 and inducing EMT via regulation of ZEB1. *Int J Med Sci.* (2019) 16:51–9. doi: 10.7150/ijms.27359

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

Copyright © 2019 Liu, Liu, Jin, Zhang, Wang, Feng, Bian, Fei, Yin and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.