Technology in Cancer Research & Treatment
Volume 19: 1-6
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1533033820962120
journals.sagepub.com/home/tct
SAGE

Identification of lnc RNAs Related to Prognosis of Patients With Colorectal Cancer

Yuqi Sun, BS[1], Peng Peng, MM[2], Lanlan He, MM[2], and Xueren Gao, PhD[1]

Abstract

The purpose of this study was to identify long noncoding RNAs (lncRNAs) related to prognosis of patients with colorectal cancer (CRC) and develop a prognostic prediction model for CRC. Transcriptome data and survival information of CRC patients were downloaded from The Cancer Genome Atlas. The differentially expressed lncRNAs (DElncRNAs) between CRC and normal colorectal tissues were identified by the edgeR package. The association of DElncRNAs expression with prognosis of CRC patients was analyzed by the survival package. A nomogram predicting 3- and 5- year overall survival of CRC patients was drawn by the rms package. A total of 1046 DElncRNAs were identified, including 271 down-regulated and 775 up-regulated lncRNAs in CRC. Multivariate Cox regression analysis showed 10 lncRNAs related to the prognosis of CRC patients. Thereinto high expression of AC004009.1, LHX1-DT, ELFN1-AS1, AL136307.1, AC087379.2, RBAKDN and AC078820.1 was associated with poorer prognosis of CRC patients. High expression of LINC01055, AL590483.1 and AC008514.1 was associated with better prognosis of CRC patients. Furthermore, the risk score model developed based on the 10 lncRNAs could effectively predict overall survival of CRC patients. In conclusion, 10 prognostic biomarkers for CRC were identified, which would be helpful to understand the role of lncRNAs in CRC progression.

Keywords

lncRNA, prognosis, colorectal cancer, biomarker, bioinformation

Abbreviations

lncRNAs, long noncoding RNAs; CRC, colorectal cancer; DElncRNAs, differentially expressed lncRNAs; C-index, concordance index; FDR, false discovery rate; OS, overall survival; AUC, area under curve.

Received: May 27, 2020; Revised: August 7, 2020; Accepted: August 21, 2020.

Introduction

Long noncoding RNAs (lncRNAs) represent a large family of RNA transcripts longer than 200 nucleotides and having no protein-coding potential. The most recent statistics from GENCODE show that human genome contains 17904 lncRNA genes that encode more than 48000 distinct lncRNA transcripts. In the last decade, human lncRNAs have received considerable attention partly due to their abundant presence in the human genome and tissue-specific expression patterns. Meanwhile, biological function and clinical significance of lncRNAs are increasingly reported in some human diseases.[1-3] For colorectal cancer (CRC), several lncRNAs have been reported to be involved in tumorigenesis and development. For instance, Xu et al. found that an immune-related lncRNA, MIR17HG, could promote tumorigenesis and metastasis of CRC cells and might serve as a promising therapeutic target.[4] Liu et al. demonstrated that lncRNA GAS5 could inhibit migration and invasion of CRC cells and promote autophagy by targeting miR-222-3p via the GAS5/PTEN-signaling pathway.[5] Zhao et al. revealed that LINC02418 significantly overexpressed in CRC could promote proliferation of CRC cells and inhibit apoptosis of CRC cells through the miR-1273g-3p-MELK axis.[6] Furthermore,

[1] School of Pharmacy, Yancheng Teachers' University, Yancheng, Jiangsu, China
[2] Nanjing Health School, Jiangsu Union Technical Institute, Nanjing, Jiangsu, China

Corresponding Author:
Xueren Gao, School of Pharmacy, Yancheng Teachers' University, No. 2, South hope Avenue, Yancheng, Jiangsu 224007, China.
Email: gaoxr@yctu.edu.cn

exosomal LINC02418 could effectively distinguish CRC patients from healthy controls, suggesting that exosomal LINC02418 might be a promising diagnostic biomarker for CRC.[6] Considering that most CRC-related lncRNAs have not yet been identified, we explored prognosis-related lncRNAs by mining high-throughput RNA sequencing data of CRC patients in the current study. In addition, biological processes and pathways closely linked to prognosis-related lncRNAs were investigated by gene co-expression analysis. A prognostic prediction model for CRC was developed based on prognosis-related lncRNAs. In a word, the study would not only contribute to understanding the role of lncRNAs in CRC progression, but also provide evidence of developing effective prognostic biomarkers for CRC.

Materials and Methods

Data Sources and Processing

Gene expression profiles and clinical data of CRC patients were obtained from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/). LncRNAs/mRNAs expression profiles and survival information of CRC patients were extracted by the Perl programming language. Finally, LncRNAs/mRNAs expression profiles of 647 CRC tissues and 51 normal colorectal tissues were included in the current study. The edgeR package was used to acquire standardized data and the differentially expressed lncRNAs (DElncRNAs) between CRC tissues and normal colorectal tissues.[7,8] DElncRNAs were selected according to the following criteria: |log2fold change| >2 and false discovery rate (FDR) < 0.05.

Survival Analysis

Univariate Cox regression analysis was used to explore the relationship between each DElncRNA expression and overall survival (OS) of CRC patients. The Benjamini-Hochberg procedure was applied to adjust the false discovery rate. DElncRNAs with $P < 0.05$ were further explored by multivariate Cox regression analysis. Finally, DElncRNAs related to OS of CRC patients were selected according to the following criteria: $P < 0.05$ under multivariate analysis, which was performed by survival package in R software.

Construction of a Prognostic Prediction Model

A prognostic prediction model for CRC was constructed based on multivariate Cox regression analysis. Proportional hazards (PH) hypothesis testing for the model was performed using Cox.zph function in survival package. The rms package was used to draw a nomogram predicting 3- and 5- year OS of CRC patients. The concordance index (C-index) with 95% confidence interval (CI) was calculated to evaluate predictive performance of the model. Meanwhile, risk score of each CRC patient was calculated using the model formula. The time ROC package was used to draw time-dependent ROC curve and estimate area under curve (AUC).[9] CRC patients were grouped into high- or low-risk patients based on the median risk score value. Kaplan-Meier survival curve was drawn to analyze the OS difference between high-risk patients and low-risk patients. $P < 0.05$ was considered significant.

The Correlation Between Prognosis-Related lncRNA Expression and Tumor Stages of CRC Patients

The correlation between prognosis-related lncRNA expression and tumor stages of CRC patients was explored using Spearman's correlation analysis in R software. A total of 599 CRC patients, including 105 stage I patients, 227 stage II patients, 179 stage III patients and 88 stage IV patients, were analyzed. $P < 0.05$ was considered significant.

Identification of Biological Processes and Pathways Closely Linked to Prognosis-Related lncRNAs

To reveal the potential function of each prognosis-related lncRNA, we analyzed the expression correlation between each prognosis-related lncRNA and all protein-coding genes, and performed biological process and pathway enrichment analysis for these protein-coding genes related to lncRNA. $P < 0.05$ and |correlation coefficient| >0.3 were considered relevant. The biological process and pathway with adjusted $P < 0.05$ were considered to be significantly enriched. These analyses were performed by clusterProfiler package in R software.[10]

Results

There were 1046 DElncRNAs between CRC tissues and normal colorectal tissues. Thereinto, the expression levels of 775 lncRNAs, such as AC026336.3, FEZF1-AS1, AC105460.1,

Table 1. Top 10 up-Regulated and Down-Regulated lncRNAs in CRC.

LncRNAs	logFC	logCPM	PValue	FDR
AL121974.1	−6.00	3.88	3.13E-52	1.58E-50
PGM5-AS1	−5.23	5.63	1.45E-105	5.19E-103
CDKN2B-AS1	−5.10	9.44	2.02E-157	4.83E-154
HAND2-AS1	−5.04	8.39	4.68E-158	1.68E-154
LINC02490	−4.94	4.54	7.99E-90	1.69E-87
AC110491.1	−4.93	3.89	1.97E-73	2.36E-71
AC087379.1	−4.91	7.09	1.25E-149	1.79E-146
LINC00974	−4.84	4.82	1.31E-108	5.24E-106
LINC00682	−4.74	3.48	9.96E-143	1.02E-139
AC007182.1	−4.65	6.08	1.20E-140	1.08E-137
AC026336.3	8.55	7.96	1.38E-29	2.64E-28
FEZF1-AS1	8.21	9.07	6.83E-46	2.67E-44
AC105460.1	8.01	8.51	1.94E-13	1.27E-12
LINC02418	8.00	11.09	3.09E-54	1.69E-52
AC104823.1	7.84	9.33	3.16E-27	5.00E-26
TMEM132D-AS1	7.76	7.07	2.13E-12	1.26E-11
AC073365.1	7.62	6.40	5.11E-22	5.91E-21
BX322234.2	7.53	6.17	2.93E-11	1.52E-10
LINC02474	7.40	6.78	2.31E-20	2.40E-19
LINC01234	7.32	8.96	2.61E-31	5.48E-30

lncRNA: long noncoding RNAs; logFC: log2 fold change; logCPM: log2 counts per million; FDR: false discovery rate.

Table 2. Identification of lncRNAs Related to Prognosis of CRC Patients.

Univariate analysis				Multivariate analysis			
lncRNAs	HR	95%CI	P	lncRNAs	HR	95%CI	P
AC011840.1	1.22	1.09–1.38	0.023				
AL807761.3	1.35	1.13–1.62	0.023				
LINC01249	1.45	1.15–1.83	0.027				
AC004080.1	1.18	1.06–1.32	0.027				
AC004009.1	1.19	1.06–1.33	0.027	AC004009.1	1.28	1.00–1.64	0.046
LINC00973	1.24	1.07–1.43	0.03				
LINC00461	1.36	1.1–1.69	0.032				
DBET	1.21	1.06–1.39	0.033				
LHX1-DT	1.32	1.08–1.61	0.033	LHX1-DT	1.42	1.04–1.94	0.026
ELFN1-AS1	1.27	1.07–1.52	0.033	ELFN1-AS1	1.32	1.05–1.67	0.019
AC020891.2	1.31	1.07–1.59	0.033				
LINC01351	1.24	1.05–1.47	0.04				
LMO7-AS1	1.31	1.06–1.62	0.04				
LINC01055	0.69	0.52–0.93	0.04	LINC01055	0.65	0.43–0.98	0.038
AC105118.1	1.34	1.06–1.7	0.04				
LINC02577	1.19	1.03–1.38	0.04				
AL136307.1	1.25	1.04–1.5	0.04	AL136307.1	1.72	1.22–2.41	0.002
LINC02241	1.24	1.04–1.49	0.04				
AP005230.1	1.23	1.04–1.45	0.04				
CLMAT3	1.19	1.03–1.37	0.041				
AC079612.1	0.72	0.54–0.95	0.041				
EVX1-AS	1.15	1.02–1.28	0.041				
Z97200.1	1.21	1.03–1.42	0.041				
AL590483.1	0.81	0.67–0.97	0.045	AL590483.1	0.67	0.49–0.91	0.011
AL079303.1	1.21	1.02–1.44	0.045				
LINC01219	1.26	1.03–1.55	0.045				
AC073365.1	1.11	1.01–1.22	0.045				
AC087379.2	1.44	1.04–2.01	0.045	AC087379.2	2.09	1.20–3.64	0.009
AC008649.2	0.70	0.51–0.97	0.045				
AC093895.1	1.15	1.01–1.31	0.045				
AC008514.1	0.81	0.67–0.98	0.045	AC008514.1	0.62	0.46–0.83	0.001
IGFBP7-AS1	1.15	1.01–1.3	0.045				
AL662890.1	1.25	1.02–1.55	0.045				
AC016831.6	1.22	1.01–1.48	0.045				
LINC01980	1.13	1.01–1.26	0.045				
AC092969.1	1.17	1.01–1.36	0.045				
AC092723.1	0.83	0.7–0.99	0.045				
AC022034.1	0.85	0.72–0.99	0.046				
AC247036.1	1.16	1.01–1.34	0.046				
AC048344.4	1.28	1.01–1.63	0.046				
RBAKDN	1.22	1.01–1.48	0.046	RBAKDN	1.47	1.12–1.92	0.005
BX470102.1	1.22	1.01–1.47	0.046				
AC010789.1	1.14	1–1.3	0.047				
AC105219.2	1.17	1–1.38	0.048				
AC078820.1	1.19	1–1.42	0.048	AC078820.1	1.32	1.00–1.73	0.05

lncRNA: long noncoding RNAs; HR: Hazard Ratio; CI: Confidence Interval.

LINC02418, AC104823.1, TMEM132D-AS1, AC073365.1, BX322234.2, LINC02474 and LINC01234, were up-regulated in CRC (Table 1). The expression levels of 271 lncRNAs, such as AL121974.1, PGM5-AS1, CDKN2B-AS1, HAND2-AS1, LINC02490, AC110491.1, AC087379.1, LINC00974, LINC00682 and AC007182.1, were down-regulated in CRC (Table 1).

Univariate analysis showed 45 lncRNAs related to prognosis of CRC patients (Table 2). Thereinto, high expression of 7 lncRNAs (LINC01055, AC008649.2, AC079612.1, AL590483.1, AC008514.1, AC092723.1, AC022034.1) was associated with better prognosis. High expression of 38 lncRNAs (AC073365.1, LINC01980, AC010789.1, EVX1-AS, AC093895.1, IGFBP7-AS1, AC247036.1, AC092969.1, AC105219.2, AC004080.1, AC004009.1, LINC02577, CLMAT3, AC078820.1, DBET, Z97200.1, AL079303.1, AC011840.1, AC016831.6, RBAKDN, BX470102.1, AP005230.1, LINC00973, LINC01351, LINC02241, AL136307.1, AL662

Table 3. The Basic Information of 10 lncRNAs Predicting the Prognosis of Colorectal Cancer.

	Genomic (GRCh38/hg38)			
lncRNA name	Location	Size (bases)	Orientation	Subcellular locations
AC008514.1	chr5:170,747,047-170,788,650	41,604	Minus strand	NA
LINC01055	chr13:45,680,184-45,701,683	21,500	Minus strand	NA
AL590483.1	chr1:244,068,820-244,093,026	24,207	Minus strand	NA
AC004009.1	chr7:27,361,591-27,410,358	48,768	Plus strand	NA
ELFN1-AS1	chr7:1,727,354 -1,742,310	14,957	Minus strand	Nucleus
AC078820.1	chr12:75,694,010-75,698,816	4,807	Minus strand	NA
LHX1-DT	chr17:36,861,674-36,936,723	75,050	Minus strand	NA
RBAKDN	chr7:5,072,060-5,073,223	1,164	Plus strand	NA
AL136307.1	chr6:5,851,506-5,870,220	18,715	Plus strand	NA
AC087379.2	chr11:15,605,484-15,705,376	99,893	Plus strand	NA

lncRNA: long noncoding RNAs.

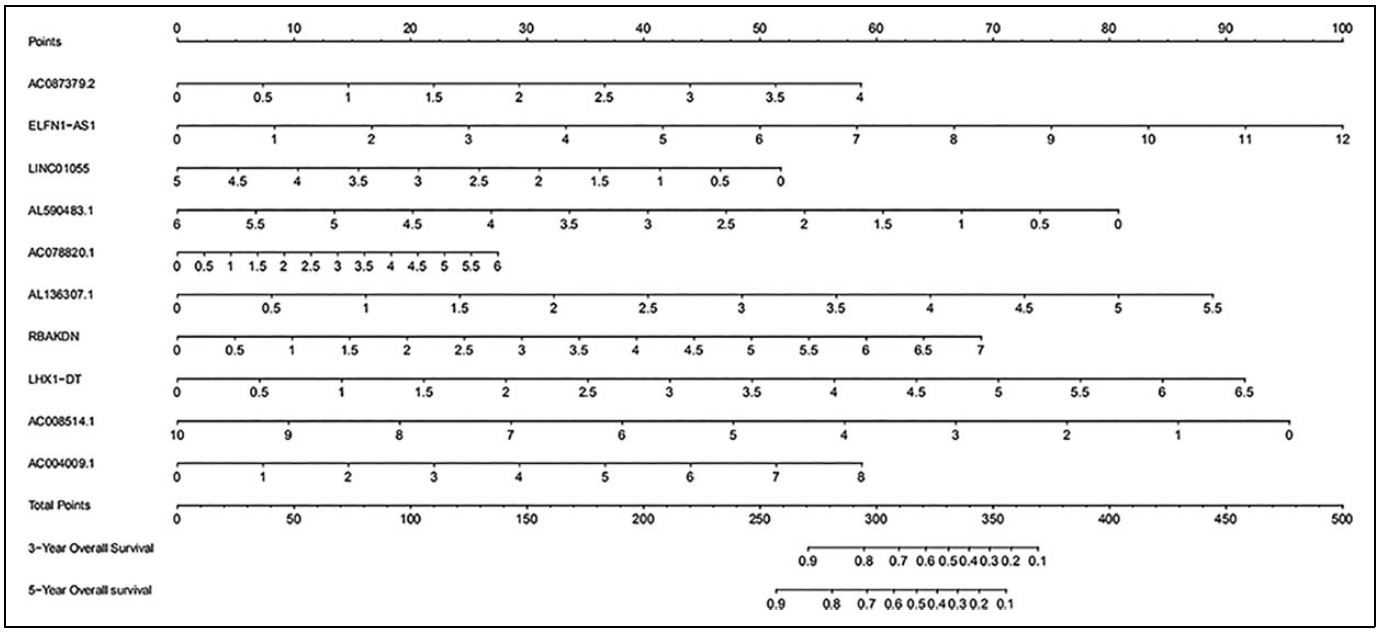

Figure 1. A nomogram predicting 3- and 5- year OS of CRC patients based on 10-lncRNA signature.

890.1, LINC01219, ELFN1-AS1, AC048344.4, AC020891.2, LMO7-AS1, LHX1-DT, AC105118.1, AL807761.3, LINC0 0461, AC087379.2, LINC01249) was associated with poorer prognosis. Further multivariate analysis showed that the expression levels of 10 lncRNAs were related to prognosis of CRC patients (Table 3). Thereinto, high expression of 3 lncRNAs (AC008514.1, LINC01055, AL590483.1) was associated with better prognosis. High expression of 7 lncRNAs (AC004009.1, ELFN1-AS1, AC078820.1, LHX1-DT, RBAKDN, AL136307.1, AC087379.2) was associated with poorer prognosis.

A prognostic prediction model for CRC, including the following 10 lncRNAs: AC008514.1, LINC01055, AL590483.1, AC004009.1, ELFN1-AS1, AC078820.1, LHX1-DT, RBAKDN, AL136307.1 and AC087379.2, was constructed (Figure 1). PH hypothesis testing showed that the model satisfied the PH test ($P = 0.35$). The C-index with 95% CI was 0.80 (0.74-0.86), suggesting that the model had an effective prediction performance. In addition, ROC curves showed that AUC of 3- and 5-year OS were 0.725 and 0.803, respectively, suggesting that the model had a moderate sensitivity and specificity (Figure 2). The scatter diagram showed that mortality of high-risk patients was significantly higher than that of low-risk patients (16.8% vs 4.2%, $P < 0.001$, Figure 3). Kaplan-Meier plot showed that the OS of high-risk patients was significantly poorer than that of low-risk patients ($P = 2.7e-07$, Figure 4).

Correlation analysis was use to explore the correlation between prognosis-related lncRNA expression and tumor stages of CRC patients. The results showed a relatively weak correlation between RBAKDN expression and tumor stages of CRC patients ($P = 0.001$, $r_s = 0.13$, Figure 5).

To explore the potential biological functions of prognosis-related lncRNAs, we performed biological process and

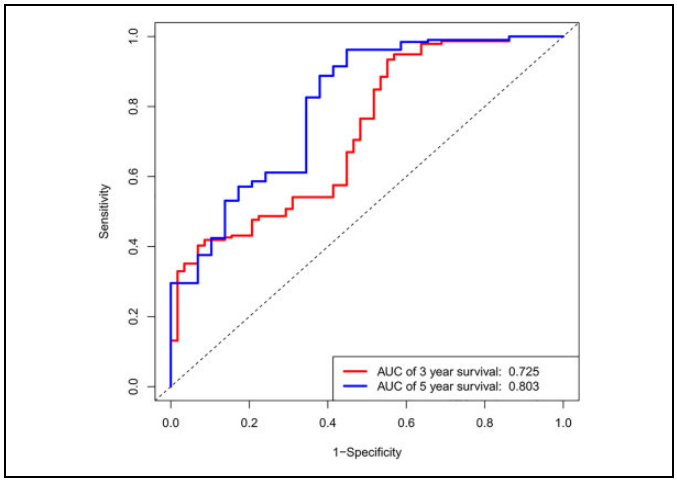

Figure 2. ROC curve analysis of prognostic prediction model for CRC.

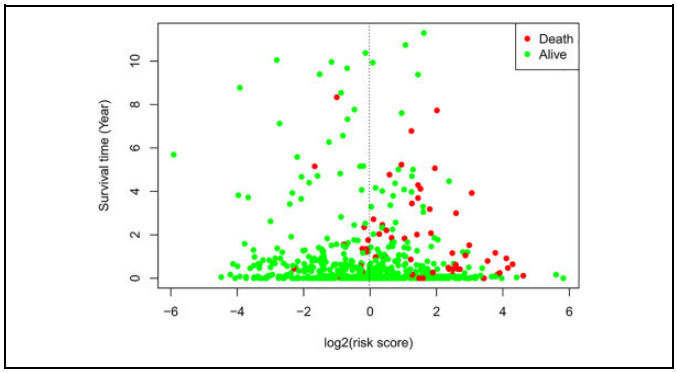

Figure 3. The relationship between overall survival and risk score in CRC patients.

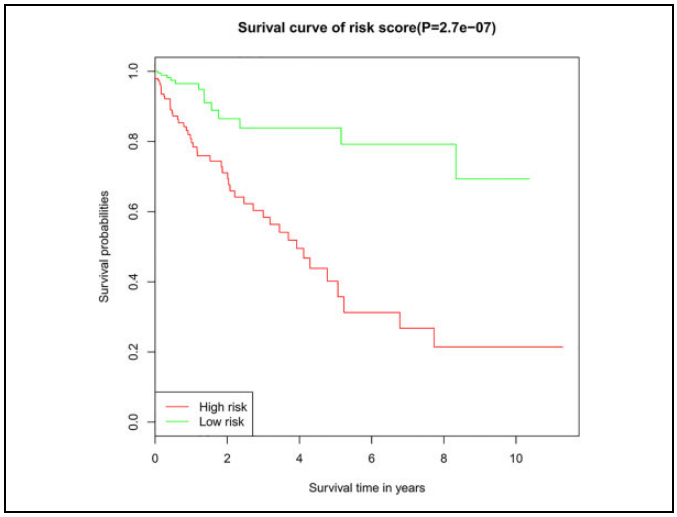

Figure 4. Kaplan-Meier survival analysis of CRC patients with high-risk and low-risk score.

pathway enrichment analysis for protein-coding genes related to each prognosis-related lncRNA. The results showed that the protein-coding genes related to AC087379.2 were enriched in

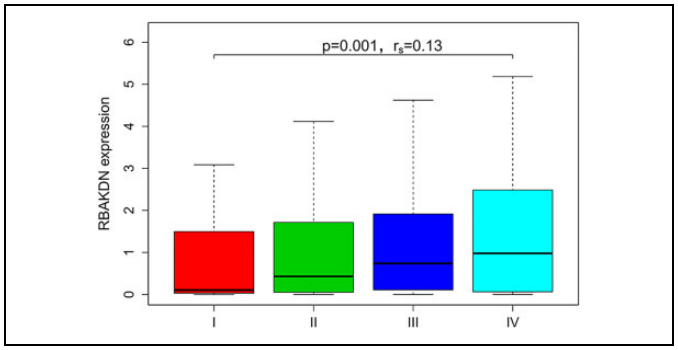

Figure 5. The correlation between RBAKDN expression and tumor stages of CRC patients.

336 biological processes and 27 pathways. The protein-coding genes related to ELFN1-AS1 were enriched in 225 biological processes and 13 pathways. The protein-coding genes related to LHX1-DT were enriched in 341 biological processes and 12 pathways. The protein-coding genes related to LINC01055 were enriched in 289 biological processes and 4 pathways. The protein-coding genes related to AC078820.1 were enriched in 41 biological processes and 2 pathways. Thereinto some enriched biological processes and pathways, such as PPAR signaling pathway, ABC transporters, oxidative phosphorylation, immune response-regulating signaling pathway and regulation of telomere maintenance, were involved in the occurrence and development of cancer.

Discussion

CRC was one of the leading causes of cancer-related death worldwide. Although a growing number of therapies were available for early and advanced CRC, the OS of CRC patients remained unsatisfactory. Therefore, it was of great importance to explore reliable biomarkers to predict the prognosis and optimize the clinical therapy decision.

In the present study, we analyzed transcriptome data of CRC and normal colorectal tissues, and identified 1046 CRC-related lncRNAs. Thereinto, the expression levels of 775 and 271 lncRNAs were up-regulated and down-regulated in CRC tissues, respectively. Further survival analysis showed that the expression levels of 10 lncRNAs were associated to the prognosis of CRC patients, and might be served as independent prognostic markers. Thereinto, high expression of AC008514.1, LINC01055 and AL590483.1 was associated with better prognosis. High expression of AC004009.1, ELFN1-AS1, AC078820.1, LHX1-DT, RBAKDN, AL136307.1 and AC087379.2 was associated with poorer prognosis.

To further explore the biological function and and clinical significance of these lncRNAs in CRC, we searched the literature and found that ELFN1-AS1 could accelerate the proliferation and migration of CRC via regulation of miR-4644/TRIM44 axis.[11] ELFN1-AS1 up-regulation was correlated with poor prognosis of CRC patients.[11] In addition, enrichment analysis suggested that multiple biological processes and pathways

related to cancer, such as PPAR signaling pathway,[12] oxidative phosphorylation,[13-15] immune response-regulating signaling pathway and regulation of telomere maintenance,[16-18] were closely linked to these prognosis-related lncRNAs. The correlation analysis showed that high expression of RBAKDN was related to advanced tumor stages of CRC, suggesting that RBAKDN could act as an oncogene participating in the occurrence and development of CRC. Based on the 10 prognosis-related lncRNAs, a risk score model effectively predicting 3- and 5- year OS of CRC patients was constructed. Further subgroup analysis based on risk score showed that the prognosis of high-risk patients was significantly poorer than that of low-risk patients. Thus, high-risk patients who required targeted treatment interventions might be identified by the 10-lncRNA signature.

In conclusion, the study identified the 10-lncRNA signature, which might act as an effective tool determining the prognosis of CRC patients, and also indicate potential therapeutic targets for CRC.

Author Contribution

Yuqi Sun and Peng Peng Equivalent contribution author.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics Statement

This is not applicable because all data are from public database.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was supported by the fourth issue of curriculum reform in Jiangsu Province (No: ZYB84).

ORCID iD

Xueren Gao https://orcid.org/0000-0003-3753-1684

Supplemental Material

Supplemental material for this article is available online.

References

1. Huang T, Wang J, Zhou Y, Zhao Y, Hang D, Cao Y. LncRNA CASC2 is up-regulated in osteoarthritis and participates in the regulation of IL-17 expression and chondrocyte proliferation and apoptosis. *Biosci Rep*. 2019;39(5):BSR20182454.
2. Liang Z, Zhu B, Meng D, et al. Down-regulation of lncRNA-NEF indicates poor prognosis in intrahepatic cholangiocarcinoma. *Biosci Rep*. 2019;39(5):BSR20181573.
3. Yang T, Zhang W, Wang L, et al. Long intergenic noncoding RNA-p21 inhibits apoptosis by decreasing PUMA expression in non-small cell lung cancer. *J Int Med Res*. 2019;47(1):481-493.
4. Xu J, Meng Q, Li X, et al. Long noncoding RNA MIR17HG promotes colorectal cancer progression via miR-17-5p. *Cancer Res*. 2019;79(19):4882-4895.
5. Liu L, Wang HJ, Meng T, et al. lncRNA GAS5 inhibits cell migration and invasion and promotes autophagy by targeting miR-222-3p via the GAS5/PTEN-signaling pathway in CRC. *Mol Ther Nucleic Acids*. 2019;17(6):644-656.
6. Zhao Y, Du T, Du L, et al. Long noncoding RNA LINC02418 regulates MELK expression by acting as a ceRNA and may serve as a diagnostic marker for colorectal cancer. *Cell Death Dis*. 2019;10:568.
7. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140.
8. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*. 2012;40(10):4288-4297.
9. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. 2013;32(30):5381-5397.
10. Yu G, Wang LG, Han Y, He QY. Cluster profiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284-287.
11. Lei R, Feng L, Hong D. ELFN1-AS1 accelerates the proliferation and migration of colorectal cancer via regulation of miR-4644/TRIM44 axis. *Cancer Biomark*. 2020;27(4):433-443.
12. Zhang X, Yao J, Shi H, Gao B, Zhang L. LncRNA TINCR/microRNA-107/CD36 regulates cell proliferation and apoptosis in colorectal cancer via PPAR signaling pathway based on bioinformatics analysis. *Biol Chem*. 2019;400(5):663-675.
13. Cheng MH, Huang HL, Lin YY, et al. BA6 induces apoptosis via stimulation of reactive oxygen species and inhibition of oxidative phosphorylation in human lung cancer cells. *Oxid Med Cell Longe*. 2019;2019:6342104.
14. Rao S, Mondragón L, Pranjic B, et al. AIF-regulated oxidative phosphorylation supports lung cancer development. *Cell Res*. 2019;29(7):579-591.
15. Huang Q, Chen Z, Cheng P, et al. LYRM2 directly regulates complex I activity to support tumor growth in colorectal cancer by oxidative phosphorylation. *Cancer Lett*. 2019;455:36-47.
16. Reeves E, James E. Antigen processing and immune regulation in the response to tumours. *Immunology*. 2017;150(1):16-24.
17. Jafri MA, Ansari SA, Alqahtani MH, Shay JW. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med*. 2016;8(1):69.
18. De Vitis M, Berardinelli F, Sgura A.Telomere length maintenance in cancer: at the crossroad between telomerase and alternative lengthening of telomeres (ALT). *Int J Mol Sci*. 2018;19(2):E606.