Abstract. Competing endogenous RNAs (ceRNAs) render the functions of long non-coding RNAs (lncRNAs) more complicated during cancer processes. Potential lncRNA biomarkers and their roles as ceRNAs have not been clearly described for rectal adenocarcinoma (READ). In the present study, we extracted data from The Cancer Genome Atlas (TCGA) including data from 167 tumor samples and 10 adjacent non-tumor samples. A total of 202 lncRNAs, 190 microRNAs (miRNAs) and 1,530 mRNAs were identified as READ-specific RNAs [log₂(fold-change)>2, FDR<0.01]. The Gene Ontology (GO) biological processes and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways were analysed for 1,530 specific mRNAs. Among 202 READ-specific lncRNAs, 7 lncRNAs were identified as being associated with overall survival of READ patients. Then, a ceRNA network was constructed with 34 key lncRNAs, 25 miRNAs and 65 mRNAs. A total of 7 lncRNAs from the network were revealed to be linked to clinical features. The results of qRT-PCR ascertained that our analysis was credible. Overall, this research provides a novel perspective from which to study the lncRNA-related ceRNA network in READ and assists in the identification of new potential biomarkers to be used for diagnostic and prognostic purposes.

Introduction

Colorectal cancer (CRC) is a cancer type that has high incidence and mortality worldwide. Rectal adenocarcinoma (READ) is a type of CRC (1). Many studies have focused on READ. However, its mortality remains high due to a lack of efficient biomarkers. Previous studies have demonstrated that various types of RNAs play important roles in cancer development and progression by acting in multiple ways (2,3). To further investigate the relationship among various types of RNAs and to obtain more efficient biomarkers, we identified cancer-specific RNAs and developed a competing endogenous RNA (ceRNA) network based on three types of RNAs, including long non-coding RNAs (lncRNAs), microRNAs (miRNAs) and mRNAs, that are differentially expressed in READ.

Non-coding RNA (ncRNA) is a type of RNA molecule that exists ubiquitously in organisms but lacks protein-coding ability (4). lncRNAs, once viewed as transcriptional ‘noise’, are one subtype of ncRNA and are identified as ncRNAs with >200 bp. Previous studies have revealed that lncRNAs use MREs to compete for miRNA-binding sites, further affecting the expression of miRNAs and the competition between mRNAs, lncRNAs and pseudogene transcripts and playing a crucial role in tumor processes (9,10).

Some studies regarding lncRNA profiles in CRC have been performed, and the functions of some lncRNAs in CRC have already been demonstrated (11-14), however studies with large sample sizes and high throughput detection methods on specific READ lncRNAs are still lacking. In addition, few studies have focused on dysregulated lncRNAs that are associated with sex, TNM, survival or other clinical features, and even fewer studies have been designed to identify the potential ceRNA network in READ. To provide answers to the questions mentioned, we used bioinformatic tools and analysis data from The Cancer Genome Atlas (TCGA), which is a public database with expression data of lncRNAs, miRNAs and mRNAs of READ, TCGA.
contains RNA sequencing data of a total of 167 READ tumor tissues and 10 adjacent non-tumor tissue samples. To the best of our knowledge, the present study is the first to identify lncRNAs that can be potential biomarkers and to further identify the ceRNA network in READ. To ascertain the credibility of our results, we randomly selected several lncRNAs from the ceRNA network and confirmed their profiles by qRT-PCR. The relationship between the expression pattern and the clinical features of seven lncRNAs were also confirmed by qRT-PCR. This approach aided in revealing the functions of IncRNAs and constructed a ceRNA network for READ.

Materials and methods

Data of patients and samples. RNA expression data with clinical data such as pathological stage, sex, and TNM information were all downloaded from TCGA database. We set exclusion criteria as follows: i) histological diagnosis revealing that the tissue was not READ; ii) suffering malignancies other than READ; iii) samples without enough data for analysis; and iv) patients who underwent preoperative radiotherapy and chemotheraphy. Ultimately, 167 tumor tissues and 10 adjacent non-tumor tissues were used in the study. The number of tumor tissues in tumor stages I, II, III, and IV were 30, 51, 51, and 24, respectively, based on the pathological stage. The study followed the guidelines of TCGA, thus, the approval of an ethics committee was not required.

The READ specimens and their paired adjacent non-tumor tissues of 90 patients were selected from the First Affiliated Hospital of Nanjing Medical University (Jiangsu, China) for qRT-PCR analysis. Their ages ranged from 45-80 years, and they were diagnosed as having rectal cancer based on histopathology and clinical history. The tissues were stored in RNAlater (Ambion; Thermo Fisher Scientific, Inc., Austin, TX, USA) at -80˚C until RNA extraction and further analyses were performed. The clinical features of 60 patients were also obtained.

RNA sequence data and further analysis. RNA expression pattern data (level 3) from patients with READ was obtained from TCGA database (September 2017), which provides normalized data from RNA sequencing by the RNASeqV2 platform (Illumina, Inc., Hayward, CA, USA) to obtain STAD data package in R {edgeR, R version 3.4.1} [absolute log(fold-change)>2.0, FDR<0.01]. Then, we chose an intersection of differentially expressed lncRNAs and mRNAs. Cytoscape v3.0 was used to construct the lncRNA/miRNA/mRNA ceRNA network. The flowchart of the ceRNA network is presented in Fig. 2.

Clinical feature analysis of key IncRNAs. We extracted clinical information from TCGA database. Based on the established ceRNA network in this study, we selected IncRNAs from the network for further clinical analysis, and clinical features such as sex, age, tumor staging, TNM staging and lymphatic metastasis were chosen to analyze the correlation between these features and key IncRNAs. We also investigated the association between specific READ IncRNAs and the overall survival time of patients.

Table I. Primer sequences used for qRT-PCR.

| Primer   | Sequence                        |
|----------|---------------------------------|
| HULC     | F 5'-ATCTGCAAGGCGGAAGAAGTC-3'   |
|          | R 5'-CTTTGGCTTATGGCTTTGCTTG-3'  |
| CRNDE    | F 5'-TGGATGCTGTCAAGCTAATTGC-3'  |
|          | R 5'-TTCCAGTGGGATACCTCCCTTTC-3'|
| PVT1     | F 5'-TGAGAACCTGTCTCTAGGAC-3'    |
|          | R 5'-AGAGCACCAAGACTGGCTCT-3'    |
| ADAMTS9-AS2 | F 5'-TAAAGACCCACAGAAGACACAG-3' |
|          | R 5'-CGTATCGTCCGGCCCTTGTC-3'    |
| GAPDH    | F 5'-ACAGTCAGCGCCATCTTCTT-3'    |
|          | R 5'-GACAAGCCTTCCTGGTTCTCAG-3'  |

bioinformatics resources (https://david.ncifcrf.gov/) was used for the functional enrichment analysis, and we only researched the Gene Ontology (GO) biological processes and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways. The criteria were set as P<0.05 and an enrichment score >1.5.

Construction of a ceRNA network. According to the theory that lncRNAs can affect miRNAs and can act as miRNA sponges to further regulate mRNAs, we constructed a ceRNA network. The miRcode (http://www.mircode.org/) was used to predict the lncRNA/miRNA interactions based on specific READ miRNAs. We predicted miRNA-targeted mRNA by using TargetScan (http://www.targetscan.org/), miRDB (http://www.mirdb.org/) and miRanda (http://www.microrna.org/microrna/home.do). We retained the intersection with the differentially expressed IncRNAs and mRNAs. Cytoscape v3.0 was used to construct the IncRNA/miRNA/mRNA ceRNA network. The flowchart of the ceRNA network is presented in Fig. 2.
Figure 1. Flowchart of bioinformatics analysis. READ, rectal adenocarcinoma; lncRNAs, long non-coding RNAs; miRNAs, microRNAs; GO, Gene Ontology; KEGG, Kyoto Encyclopaedia of Genes and Genomes; ceRNAs, competing endogenous RNAs.

Figure 2. Flowchart of the construction of a ceRNA network. READ, rectal adenocarcinoma; lncRNAs, long non-coding RNAs; miRNAs, microRNAs; ceRNAs, competing endogenous RNA.
Statistical analysis. Statistical analysis was performed using R Studio (R version 3.4.1), Statistical Programme for Social Sciences 20.0 (SPSS, Inc., Chicago, IL, USA) and GraphPad Prism 5.0 software (GraphPad Software, Inc., La Jolla, CA, USA). The lncRNA data set and the overall survival information were profiled using the univariate Cox proportional hazards regression model. The results were presented as Kaplan-Meier survival curves and multivariate Cox regression analysis was applied for further study. Paired t-tests were used to compare the differences in the qRT-PCR results. P<0.05 was considered to indicate a statistically significant difference.

Results

Specific READ lncRNAs in patients. In total, 272 lncRNAs were identified that were differentially expressed between READ tissues and adjacent non-tumor tissues from TCGA database [absolute log2(fold-change)>2, FDR<0.01], of which 143 lncRNAs were upregulated, and 129 lncRNAs were downregulated. Further analysis of the differences between tumor tissues and adjacent non-tumor tissues in patients with stage I, II, III and IV cancer was performed. Then, 337, 323, 307 and 330 differentially expressed lncRNAs were identified.
between adjacent non-tumor tissues and READ stage I, II, III and IV tumor tissues, respectively. To enhance the credibility of the data, we used 202 differentially expressed lncRNAs from the intersection of the aforementioned five groups for further analysis (Fig. 3A). These 202 lncRNAs were named READ-specific lncRNAs. Finally, 34 lncRNAs were used to construct the lncRNA/miRNA/mRNA ceRNA network (Table II).

**GO enrichment and KEGG pathway analyses of differentially expressed genes.** For further study of the functions of the differentially expressed genes, we selected intersecting mRNAs across all READ stages from the differentially expressed mRNAs for analysis. We first identified 2,000 differentially expressed mRNAs between the tumor tissues of READ patients and the adjacent non-tumor tissues [absolute \(\log_2(\text{fold-change})>2, \text{FDR}<0.01\)]. We also identified 2,363; 2,124; 2,113; and 2,189 differentially expressed mRNAs between adjacent non-tumor tissues and READ tumor tissues (stage I, II, III and IV, respectively). Then, we chose the mRNAs that were differentially expressed in all five comparison groups, finally obtaining 1,530 differentially expressed mRNAs that were used for further analysis (Fig. 3C). These mRNAs were identified as READ-specific mRNAs.

The 1,530 differentially expressed mRNAs were analyzed with DAVID bioinformatics resources. We chose to show the top 15 GO biological processes and 15 KEGG pathways of the differentially expressed genes based on the P-values (Figs. 4 and 5). Among these pathways, the PI3K-Akt, Wnt, AMPK, and cGMP-PKG signaling pathways and the cell adhesion molecules (CAMs) were confirmed as CRC-associated pathways.
Prediction of miRNA targets and construction of a ceRNA network. The 31 miRNAs that were differentially expressed between READ tumor tissues and adjacent non-tumor tissues were identified. For further investigation, we then divided the tumor tissues into four groups (stages I, II, III and IV) according to the pathological stage, and then compared each group with adjacent non-tumor tissues in turn. Subsequently, we obtained 190 specific miRNAs by selecting the intersection of miRNAs differentially expressed across the five comparison groups (Fig. 3B). These miRNAs were identified as READ-specific miRNAs. In the next step, we investigated the interactions between these intersecting miRNAs and the READ-specific lncRNAs based on miRcode (http://www.mircode.org/), and 25 key miRNAs were predicted to target 34 key lncRNAs (Table III).

The 25 key miRNAs mentioned in Table III were then used to predict key mRNAs using Targetscan (http://www.targetscan.org/), miRDB (http://www.mirdb.org/) and miRanda (http://www.microrna.org/microrna/home.do). We then compared the predicted mRNAs with the 1,530 READ-specific mRNAs and chose mRNAs that were found in both groups. Finally, 65 mRNAs were found to interact with 25 miRNAs (Table IV). Among these mRNAs, some have been verified to be transcribed from cancer-related genes.
Table IV. miRNAs that may target READ mRNAs.

| miRNAs     | mRNAs                                      |
|------------|-------------------------------------------|
| hsa-mir-107| SALL4, AXIN2, FGF2, FGFR1                 |
| hsa-mir-141| MACC1, ZEB1, EPHA7, KIAA1549, ELAVL4       |
| hsa-mir-143| COL1A1                                    |
| hsa-mir-144| FGFR2, GRIK3, NR3C1, KCNQ5                |
| hsa-mir-150| HILPDA, ZEB1                                |
| hsa-mir-152| BMP3, NPTX1                                |
| hsa-mir-155| MEIS1, CD36, GPM6B, NOVA1, PCDH9          |
| hsa-mir-17 | FAM46C, CLIP4, CFL2, FAXC, NPSAS3, FOXQ1, SLC16A9, FJX1, FAM129A, CADM2 |
| hsa-mir-182| NR3C1, NPTX1, ULBP2, CHL1, TCEAL7         |
| hsa-mir-183| ZEB1, NR3C1, AKAP12                       |
| hsa-mir-192| GRHL1, TCF7                                |
| hsa-mir-200a| EPHA7, ZEB1, MACC1, KIAA1549              |
| hsa-mir-21 | ATP2B4, PRICKLE2, CALD1, EDIL3, EPM2A, OSR1, TGFBI |
| hsa-mir-215| TCF7                                      |
| hsa-mir-217| DACH1                                     |
| hsa-mir-22 | NR3C1, RGS2                                |
| hsa-mir-223| EPB41L3                                   |
| hsa-mir-32 | ATP2B4, UGP2                               |
| hsa-mir-375| ELAVL4                                    |
| hsa-mir-424| AMOTL1, TPM2, CBX2, TMEM100, AXIN2, PSAT1, FGF2 |
| hsa-mir-425| THR8                                      |
| hsa-mir-429| ZEB1                                      |
| hsa-mir-454| SPG20, CFL2, RBM20                        |
| hsa-mir-96 | JAZ1, TRIB3, ZEB1                         |
| hsa-mir-98 | PRSS22, IGF2BP1, HAND1, SLC5A6, IGF2BP3, TRIM71, CPA4 |

miRNAs, microRNAs; READ, rectal adenocarcinoma.

Figure 6. The lncRNA/miRNA/mRNA ceRNA network. Red balls, upregulated mRNAs; blue balls, downregulated mRNAs; red squares, upregulated miRNAs; blue squares, downregulated miRNAs; red diamonds, upregulated lncRNAs; blue diamonds, downregulated lncRNAs. lncRNAs, long non-coding RNAs; miRNAs, microRNAs; ceRNA, competing endogenous RNA.
According to information provided in Tables III and IV, we constructed a miRNA/lncRNA/mRNA ceRNA network using Cytoscape 3.0. In conclusion, 25 miRNAs, 65 mRNAs and 34 lncRNAs were involved in the network (Fig. 6). We called lncRNAs involved in the ceRNA network key lncRNAs.

Clinical feature analysis of key READ-specific lncRNAs. To further study the lncRNAs, the correlation between the lncRNAs involved in the ceRNA network and clinical features including sex, age, tumor stage, TNM stage and lymphatic metastasis status in TCGA database were analysed. We identified 7 lncRNAs associated with clinical features (P<0.05). The results revealed that UCA1 and HULC were age-related, CHL1-AS2, LINC00484 and HULC were sex-related, LINC00484 and ADAMTS9-AS1 were associated with TNM stage, CLDN10-AS1 was associated with tumor stage, and UCA1 was associated with lymphatic metastasis (Table V).

We also analyzed the association of the overall survival of patients with the 202 specific lncRNAs based on the clinical data of 177 samples from TCGA database. To carry out this research, we used a univariate Cox proportional hazards regression model and finally found 7 lncRNAs that were significantly associated with the overall survival of READ patients (log-rank P<0.05). As described in Fig. 7, LINC01215, LINC01602, LINC02163 and PRSS30P were positively correlated with overall survival (P<0.05). MIR497hG, PGM5P4-AS1 and PTENP1-AS were negatively correlated with overall survival (P<0.05). Based on the results of the univariate Cox regression analysis, 7 lncRNAs were associated with overall survival. We performed a multivariate Cox regression analysis with these 7 lncRNAs. LINC01602, LINC02163 and MIR497hG were found to be independent influencing factors of survival time, LINC01602 and MIR497hG were negatively related to overall survival, and
LINC02163 was positively related to overall survival. We present these results in Table VI.

$qRT$-PCR validation. To check the credibility of the bioinformatics results, we randomly selected 4 key lncRNAs (HULC, CRNDE, PVT1 and ADAMTS9-AS2) from the network. Paired t-tests were applied to analyze the differences in expression of the selected lncRNAs between READ tumor tissues and the adjacent non-tumor tissues. HULC, CRNDE and PVT1 expression levels were increased, and the expression level of ADAMTS9-AS2 was decreased in READ tumor tissues compared to adjacent non-tumor tissues. The qRT-PCR results from 30 READ patients were consistent with the bioinformatics results (Fig. 8). We also confirmed the correlation between the expression pattern and clinical features of 7 lncRNAs. Sixty patients were divided into two groups according to the expression level of the lncRNA concerned. Patients with higher and lower than the median expression of the lncRNA concerned were allocated into high and low expression groups, respectively. The results are displayed in Table VII. Thus, our bioinformatics analysis was reliable.

Discussion

CRC is one of the most common cancers around the world, and its incidence and mortality remain high. READ is a type of CRC and has a higher incidence than the other types of colon adenocarcinoma (COAD) (15). Although great progress in the treatment, prognosis and diagnosis of READ has been achieved through studies, its mortality still remains high, which may be due to the lack of efficient biomarkers and the unclear mechanisms underlying READ.

Recent studies have demonstrated that lncRNAs play vital roles during tumor progression, and some of them may be biomarkers for better prognosis and diagnosis (16-18). lncRNAs function in many ways. The ceRNA hypothesis postulates that IncRNAs may compete with mRNAs for the binding sites of miRNAs and may further affect mRNA expression through MREs (8). The hypothesis makes the relationship of miRNAs, mRNAs and lncRNAs more complicated and better explains the interaction among a variety of types of RNAs at the genetic level.

There are many studies on lncRNAs in CRC (19,20), but few of them have focused on READ. Additionally, the sample sizes of previous studies were not large enough, and almost none of the studies focused on the potential ceRNA network. In the present study, we aimed to research the IncRNAs that may have the ability to be better biomarkers and further explored
Table VII. Expression of lncRNAs related to clinical features according to the clinicopathological characteristics of patients.

| Characteristics          | UCA1 expression |          | HULC expression |          |
|--------------------------|-----------------|----------|-----------------|----------|
|                          | No.          | High group | Low group | P-value | High group | Low group | P-value |
| Age (years)              |               |           |           |         |           |           |         |
| <50                      | 18            | 5         | 13        | 0.024   | 5          | 13        | 0.024   |
| >50                      | 42            | 25        | 17        |          | 25         | 17        |          |
| Sex                      |               |           |           |         |           |           |         |
| Female                   | 32            | 17        | 15        | 0.605   | 12         | 20        | 0.038   |
| Male                     | 28            | 13        | 15        |          | 18         | 10        |          |
| Lymphatic metastasis     |               |           |           |         |           |           |         |
| No                       | 36            | 14        | 22        | 0.035   | 15         | 21        | 0.113   |
| Yes                      | 24            | 16        | 8         |          | 15         | 9         |          |
| Tumor stage              |               |           |           |         |           |           |         |
| Stage I, II              | 33            | 12        | 21        | 0.020   | 15         | 18        | 0.436   |
| Stage III, IV            | 27            | 18        | 9         |          | 15         | 12        |          |
| TNM staging system       |               |           |           |         |           |           |         |
| T1+T2                    | 32            | 17        | 15        | 0.605   | 14         | 18        | 0.301   |
| T3+T4                    | 28            | 13        | 15        |          | 16         | 12        |          |

| Characteristics          | CHL1-AS2 expression |          | LINC00484 expression |          |
|--------------------------|---------------------|----------|----------------------|----------|
|                          | No.          | High group | Low group | P-value | High group | Low group | P-value |
| Age (years)              |               |           |           |         |           |           |         |
| <50                      | 18            | 8         | 10        | 0.573   | 9          | 9         | 1.000   |
| >50                      | 42            | 22        | 20        |          | 21         | 21        |          |
| Sex                      |               |           |           |         |           |           |         |
| Female                   | 32            | 20        | 12        | 0.038   | 21         | 11        | 0.009   |
| Male                     | 28            | 10        | 18        |          | 9          | 19        |          |
| Lymphatic metastasis     |               |           |           |         |           |           |         |
| No                       | 36            | 17        | 19        | 0.598   | 19         | 17        | 0.598   |
| Yes                      | 24            | 13        | 11        |          | 11         | 13        |          |
| Tumor stage              |               |           |           |         |           |           |         |
| Stage I, II              | 33            | 16        | 17        | 0.796   | 18         | 15        | 0.436   |
| Stage III, IV            | 27            | 14        | 13        |          | 12         | 15        |          |
| TNM staging system       |               |           |           |         |           |           |         |
| T1+T2                    | 32            | 13        | 19        | 0.120   | 20         | 12        | 0.038   |
| T3+T4                    | 28            | 17        | 11        |          | 10         | 18        |          |

| Characteristics          | CLDN10-AS1 expression |          | DLX6-AS1 expression |          |
|--------------------------|-----------------------|----------|---------------------|----------|
|                          | No.          | High group | Low group | P-value | High group | Low group | P-value |
| Age (years)              |               |           |           |         |           |           |         |
| <50                      | 18            | 8         | 10        | 0.573   | 7          | 11        | 0.260   |
| >50                      | 42            | 22        | 20        |          | 23         | 19        |         |
| Sex                      |               |           |           |         |           |           |         |
| Female                   | 32            | 16        | 16        | 1.000   | 15         | 17        | 0.605   |
| Male                     | 28            | 14        | 14        |          | 15         | 13        |          |
| Lymphatic metastasis     |               |           |           |         |           |           |         |
| No                       | 36            | 21        | 15        | 0.113   | 20         | 16        | 0.292   |
| Yes                      | 24            | 9         | 15        |          | 10         | 14        |          |
the interaction among lncRNAs, miRNAs and mRNAs by constructing a ceRNA network in READ.

In this study, we identified READ-specific lncRNAs, mRNAs and miRNAs based on the intersecting differential expression between tumor tissues from all four stages and adjacent non-tumor tissues. We further analyzed the functions and pathways involving the differentially expressed genes by GO and KEGG. Then, with the use of bioinformatics tools, we constructed a ceRNA network with READ-specific mRNAs, miRNAs and lncRNAs. We further analyzed the lncRNAs involved in the ceRNA network, examining their correlations with clinical features, and we identified specific lncRNAs that were correlated with overall survival. We finally confirmed our findings in the tissues of 30 READ patients using qRT-PCR.

Some lncRNAs from the READ-specific lncRNAs have been reported as playing vital roles in cancers. CCAT1 has been reported to promote gallbladder cancer development (21), and CCAT2 has been identified to play a significant role in the progression of colon cancer (22). This also supported the reliability of our analysis. Through univariate and multivariate Cox regression analyses, 3 IncRNAs (LINC01602, LINC02163 and MIR497HG) from the READ-specific IncRNAs were identified as being associated with overall survival. LINC01602 and MIR497HG were negatively related to overall survival, and LINC02163 was positively related to overall survival. These 3 IncRNAs were not involved in the ceRNA network but may still play vital roles in READ and be potential indicators of the prognosis of READ since they were related to overall survival.

Using GO and KEGG, we analysed the functions and pathways of READ-specific mRNAs. The GO results revealed that the functions enriched, involved aspects of immune function, metabolism and cellular functions. Among the results of the KEGG pathway analysis, several were confirmed to be cancer-associated. PI3K/AKT signaling plays crucial roles in reducing apoptosis, stimulating cell growth and increasing proliferation, as reported in previous studies (23). It has also been reported that many lncRNAs including PlncRNA-1 and AB073614 affect CRC through the PI3K/AKT signaling pathway (24,25). Studies have demonstrated that AMPK can also be activated by some IncRNAs. Li et al revealed that liver kinase B1 (LKB1) phosphorylates and promotes AMPK and then reduces cancer cell proliferation and metabolism (26). Available data indicate that Wnt signaling substantially impacts non-small cell lung cancer (NSCLC) tumorigenesis, prognosis, and resistance to therapy (27). Previous studies have reported that cell adhesion molecules (CAMs) play a significant role in the progression of metastasis (28). Li et al found that by activating the cGMP/PKG pathway, Wnt/β-catenin signaling can

Table VII. Continued.

| Characteristics | No. | High group | Low group | P-value |
|-----------------|-----|------------|-----------|---------|
| Age (years)     |     |            |           |         |
| <50             | 18  | 6          | 12        | 0.091   |
| >50             | 42  | 24         | 18        |         |
| Sex             |     |            |           |         |
| Female          | 32  | 17         | 15        | 0.605   |
| Male            | 28  | 13         | 15        |         |
| Lymphatic metastasis | |           |           |         |
| No              | 36  | 20         | 16        | 0.291   |
| Yes             | 24  | 10         | 14        |         |
| Tumor stage     |     |            |           |         |
| Stage I, II     | 33  | 18         | 15        | 0.436   |
| Stage III, IV   | 27  | 12         | 15        |         |
| TNM staging system |   |           |           |         |
| T1+T2           | 32  | 21         | 11        | 0.010   |
| T3+T4           | 28  | 9          | 19        |         |

IncRNAs, long non-coding RNAs.
be suppressed (29). Some other cancer-associated pathways such as those known to be involved in CRC were also revealed by our results, further demonstrating the reliability of the results. These pathways and functions may also be related to READ-specific lncRNAs due to different interactions between lncRNAs and mRNAs in READ.

With the ceRNA network in READ, we can further research the underlying mechanism of the intersections among lncRNAs, mRNAs and miRNAs in READ. Several interactions between RNAs in our network have been previously confirmed. UCA1 interacts with miR-182 to modulate glioma proliferation (30), and UCA1 regulates miR-134 to promote the invasion and EMT of bladder cancer (31). These previous studies strongly demonstrate that our analysis was reliable. In addition, the network can aid in understanding the interactions of miRNAs, lncRNAs and mRNAs in READ from various perspectives. Some lncRNAs that exist in the ceRNA network were previously reported to be READ-related lncRNAs. For example, Yang et al. discovered that HULC promotes colorectal carcinoma progression by epigenetically repressing NFK2 expression (32). Han et al. demonstrated that the lncRNA CRNDE can regulate the progression and chemoresistance of CRC via modulation of the expression levels of miR-181a-5p and Wnt/β-catenin signaling activity (33). High expression levels of PCAT-1 were involved in CRC progression and it can be a novel biomarker of poor prognosis in patients with CRC (34). PVT1 may be a new oncogene co-amplified with c-Myc in CRC tissues and extracellular vesicles and functionally correlated with the proliferation and apoptosis of CRC cells (35). Most lncRNAs that interacted with other RNAs may play significant roles in READ processes. Through further investigation of the lncRNAs involved in the network, 7 lncRNAs were identified as being associated with clinical features. Among these 7 lncRNAs, UCA1, HULC and LINC00484 appeared often, and these 3 lncRNAs may be the most clinically relevant lncRNAs, potentially acting as biomarkers. Further study still needs to be performed to understand their correlation with clinical features and their potential as efficient biomarkers.

To verify our bioinformatic analysis, we randomly selected 4 lncRNAs from the network to assess their expression levels in 30 paired READ tumor tissues and adjacent non-tumor tissues. We also confirmed the correlation between the expression pattern and the clinical features of 7 lncRNAs using qRT-PCR. The results revealed that our analysis was credible.

Our analysis has great meaning in the study of a ceRNA network in READ and some results were confirmed by qRT-PCR. However, there were still several limitations in our study. Firstly, the sample number of normal tissues was not very large although previous studies with a small sample size also exist (36). Secondly the follow-up information of our patients was not enough to study the overall survival, most follow-up information after surgery was ~2-3 years. In the future, we will validate our analysis results as soon as we obtain enough follow-up information. Thirdly, further research is also warranted on the functions of key lncRNAs in vivo and in vitro.

In conclusion, our study revealed READ-specific lncRNAs using bioinformatics analysis and studied their associations with clinical features based on data from TCGA database. To the best of our knowledge, studies about lncRNA profiles with such a large sample size are rare. Some key lncRNAs may become efficient biomarkers for READ diagnosis and prognosis. In addition, we constructed a ceRNA network with READ-specific lncRNAs, miRNAs and mRNAs, and it revealed the relationship among these three types of RNAs and aided in elucidating the mechanisms underlying READ on the genetic level.

Acknowledgements

Not applicable.

Funding

The study was funded by Jiangsu Key Medical Discipline (General Surgery) (no. ZDXKA2016005).

Availability of data and materials

The datasets used during the present study are available from the corresponding author upon reasonable request.

Authors' contributions

ZZ conceived, designed and further performed the study. SeW, DJ, WQ, QW, JL, JG, WP, TH, BJ, YZ, ShW and YS were also involved in the conception of the study and gave their advice in the process of the research. SW and DJ helped ZZ did the data analysis. ZZ wrote the paper. ZZ and DJ reviewed and edited the manuscript. All authors have read and approved the manuscript and agree to be accountable for all aspects of the research in ensuring that the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethics approval and consent to participate

All experimental protocols were approved by the Nanjing Medical University (Nanjing, China).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J and Jemal A: Global cancer statistics, 2012. CA Cancer J Clin 65: 87-108, 2015.
2. Lee YJ, Kang YR, Lee SY, Jih Y, Han DC and Kwon BM: Ginkgetin induces G2-phase arrest in HCT116 colon cancer cells through the modulation of b-Myb and miRNA34a expression. Int J Oncol 51: 1331-1342, 2017.
3. Wu S, Wu F and Jiang Z: Identification of hub genes, key miRNAs and potential molecular mechanisms of colorectal cancer. Oncol Rep 38: 2043-2050, 2017.
4. Yamamura S, Imai-Sumida M, Tanaka Y and Dahiya R: Interaction and cross-talk between non-coding RNAs. Cell Mol Life Sci 75: 467-484, 2018.
5. Sana J, Faltejskova P, Svoboda M and Slaby O: Novel classes of non-coding RNAs and cancer. J Transl Med 10: 103, 2012.
6. Takahashi K, Yan I, Haga H and Patel T: Long noncoding RNA in liver diseases. Hepatology 60: 744-753, 2014.
7. Svoboda M, Slyslova J, Schneiderova M, Makovicky P, Bielik L, Levy M, Lipska L, Hemmelova B, Kala Z, Protivanova M, et al.: HOTAIR long non-coding RNA is a negative prognostic factor not only in primary tumors, but also in the blood of colorectal cancer patients. Carcinogenesis 35: 1510-1515, 2014.
8. Salmena L, Poliseno L, Tay Y, Kats L and Pandolfi PP: A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? Cell 146: 353-358, 2011.
9. Takahashi K, Yan I, Haga H and Patel T: Portrait of the PI3K/AKT pathway in colorectal cancer. Biochim Biophys Acta 1855: 104-121, 2015.
10. Song W, Mei JZ and Zhang M: LncRNA PcnrRNA-1 promotes colorectal cancer cell progression by regulating PI3K/Akt signaling pathway. Oncol Rep: Aug 23, 2017 (Epub ahead of print).
11. Yang L, Qiu M, Xu L, Wang J, Zheng Y, Li M, Xu L and Yin R: Dysregulation of non-coding RNAs and cancer. J Transl Med 10: 103, 2012.
12. Cao Y, Lin M, Bu Y, Ling h, He Y, Huang C, Chen Y, Song B and Cao D: p53-inducible long non-coding RNA PICART1 mediates cancer cell proliferation and migration. Int J Oncol 50: 1671-1682, 2017.
13. Chen DL, Chen LZ, Lu YX, Zhang DS, Zeng ZL, Pan ZZ, Huang P, Wang FH, Li YH, Ju HQ, et al.: lncRNA XIST expedites metastasis and modulates epithelial-mesenchymal transition in colorectal cancer. Cell Death Dis 8: e3011, 2017.
14. Ding J, Li J, Wang H, Tian Y, Xie M, He X, Ji H, Ma Z, Hui B, Wang K, et al.: Long noncoding RNA CRNDE promotes colorectal cancer cell proliferation via epigenetically silencing DUSP5/CDKN1A expression. Cell Death Dis 8: e2997, 2017.
15. Turati F, Bravi F, Di Maso M, et al.: Adherence to the world Cancer Research Fund/American Institute for Cancer Research recommendations and colorectal cancer risk. Eur J Cancer 85: 86-94, 2017.
16. Zhang L, Liang X and Li Y: Long non-coding RNA MEG3 inhibits cell growth of gliomas by targeting miR-93 and inactivating PI3K/AKT pathway. Oncol Rep 38: 2408-2416, 2017.
17. Li Y, Huang S, Li Y, Zhang W, He K, Zhao M, Lin H, Li D, Zhang H, Zheng Z, et al.: Decreased expression of LncRNA SLC25A25-AS1 promotes proliferation, chemoresistance, and EMT in colorectal cancer cells. Tumor Biol 37: 14205-14215, 2016.
18. Qiu JJ, Zhang XD, Tang YT, Zhang Y and Hua KQ: ElncRNA1, a long non-coding RNA that is transcriptionally induced by oestrogen, promotes epithelial ovarian cancer cell proliferation. Int J Oncol 51: 507-514, 2017.
19. Ozawa T, Matsuyama T, Toyama Y, Takahashi N, Ishikawa T, Uetake H, Yamada Y, Kusunoki M, Calin G and Goel A: CCAT1 and CCAT2 long noncoding RNAs, located within the 8q24.21 ‘gene desert’, serve as important prognostic biomarkers in colorectal cancer. Ann Oncol 28: 1882-1888, 2017.
20. Yang L, Xu L, Wang Q, Wang M and An G: Dysregulation of long non-coding RNA profiles in human colorectal cancer and its association with overall survival. Oncol Lett 12: 4068-4074, 2016.
21. Ma MZ, Chu BF, Zhang Y, Weng MZ, Qin YY, Gong W and Quan ZW: Long non-coding RNA CCAT1 promotes gallbladder cancer development via negative modulation of miRNA-218-5p. Cell Death Dis 6: e1583, 2015.
22. Ling H, Spizzo R, Atsasi Y, Nicoloso M, Shimizu M, Redis RS, Nishida N, Gafà R, Song J, Guo Z, et al.: CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res 23: 1446-1461, 2013.
23. Danielsen SA, Eide PW, Nesbakken A, Guren T, Leithe E and Lothe RA: Expression of the LKB1/AMPK pathway in colorectal cancer. Biochim Biophys Acta 1855: 104-121, 2015.
24. Song W, Mei JZ and Zhang M: LncRNA PcnrRNA-1 promotes colorectal cancer cell progression by regulating PI3K/Akt signaling pathway. Oncol Rep: Aug 23, 2017 (Epub ahead of print).
25. Wang Y, Huang H, Xue J, Liao L, Yin F and Zhou X: LncRNA AB073614 regulates proliferation and metastasis of colorectal cancer cells via the PI3K/AKT signaling pathway. Biomed Pharmacother 93: 1230-1237, 2017.
26. Li N, Zhang Li, Liu H, He S, and Cao D: p53-inducible long non-coding RNA PICART1 mediates cancer cell proliferation and cell cycle progression. Oncol Rep 35: 318-324, 2016.
27. Cao Y, Lin M, Bu Y, Ling H, He Y, Huang C, Chen Y, Song B and Cao D: p53-inducible long non-coding RNA PICART1 mediates cancer cell proliferation and migration. Int J Oncol 50: 1671-1682, 2017.
28. Chen DL, Chen LZ, Lu YX, Zhang DS, Zeng ZL, Pan ZZ, Huang P, Wang FH, Li YH, Ju HQ, et al.: lncRNA XIST expedites metastasis and modulates epithelial-mesenchymal transition in colorectal cancer. Cell Death Dis 8: e3011, 2017.
29. Ding J, Li J, Wang H, Tian Y, Xie M, He X, Ji H, Ma Z, Hui B, Wang K, et al.: Long noncoding RNA CRNDE promotes colorectal cancer cell proliferation via epigenetically silencing DUSP5/CDKN1A expression. Cell Death Dis 8: e2997, 2017.
30. Turati F, Bravi F, Di Maso M, et al.: Adherence to the world Cancer Research Fund/American Institute for Cancer Research recommendations and colorectal cancer risk. Eur J Cancer 85: 86-94, 2017.
31. Zhang L, Liang X and Li Y: Long non-coding RNA MEG3 inhibits cell growth of gliomas by targeting miR-93 and inactivating PI3K/AKT pathway. Oncol Rep 38: 2408-2416, 2017.
32. Li Y, Huang S, Li Y, Zhang W, He K, Zhao M, Lin H, Li D, Zhang H, Zheng Z, et al.: Decreased expression of LncRNA SLC25A25-AS1 promotes proliferation, chemoresistance, and EMT in colorectal cancer cells. Tumor Biol 37: 14205-14215, 2016.
33. Qiu JJ, Zhang XD, Tang YT, Zhang Y and Hua KQ: ElncRNA1, a long non-coding RNA that is transcriptionally induced by oestrogen, promotes epithelial ovarian cancer cell proliferation. Int J Oncol 51: 507-514, 2017.
34. Ozawa T, Matsuyama T, Toyama Y, Takahashi N, Ishikawa T, Uetake H, Yamada Y, Kusunoki M, Calin G and Goel A: CCAT1 and CCAT2 long noncoding RNAs, located within the 8q24.21 ‘gene desert’, serve as important prognostic biomarkers in colorectal cancer. Ann Oncol 28: 1882-1888, 2017.
35. Yang L, Xu L, Wang Q, Wang M and An G: Dysregulation of long non-coding RNA profiles in human colorectal cancer and its association with overall survival. Oncol Lett 12: 4068-4074, 2016.
36. Ma MZ, Chu BF, Zhang Y, Weng MZ, Qin YY, Gong W and Quan ZW: Long non-coding RNA CCAT1 promotes gallbladder cancer development via negative modulation of miRNA-218-5p. Cell Death Dis 6: e1583, 2015.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.