

# 

**Citation:** Li F, Li Q, Wu X (2018) Construction and analysis for differentially expressed long noncoding RNAs and MicroRNAs mediated competing endogenous RNA network in colon cancer. PLoS ONE 13(2): e0192494. https://doi.org/10.1371/ journal.pone.0192494

**Editor:** Aamir Ahmad, University of South Alabama Mitchell Cancer Institute, UNITED STATES

Received: August 18, 2017

Accepted: January 24, 2018

Published: February 8, 2018

**Copyright:** © 2018 Li et al. This is an open access article distributed under the terms of the <u>Creative</u> Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data underlying this study belongs to the TCGA (The Cancer Genome Atlas), which is a collaboration between the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI). The preprocessed level 3 RNA-seq data and corresponding information of colon cancer patients can be found in The Cancer Genome Atlas (TCGA) database (http://cancergenome.nih.gov/).

**Funding:** The author(s) received no specific funding for this work.

RESEARCH ARTICLE

# Construction and analysis for differentially expressed long non-coding RNAs and MicroRNAs mediated competing endogenous RNA network in colon cancer

# Fengxi Li<sup>1</sup>, Qian Li<sup>2</sup>, Xianghua Wu<sup>1</sup>\*

1 Department of Surgery, the First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China, 2 Department of Gynaecology and Obstetrics, the First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China

\* xhwuone@163.com

# Abstract

Long non-coding RNA (IncRNA) has been confirmed to act as a key regulatory molecule in different types of cancers and play a significant role in tumors initiation and progression. LncRNA can be as acompeting endogenous RNA(ceRNA) to regulate the expression of targeted genes by sponging miRNA. In the present study, we explore the functional roles and regulatory mechanisms of IncRNAs as ceRNAs in colon cancer and their potential implications for prognosis. The IncRNAs, miRNAs and mRNAs expression profiles of 341 colon cancer tissues and 27 non-tumor colon tissues were downloaded from The Cancer Genome Atlas (TCGA) database. Differential expression of RNAs was identified using the "DESeq" bioconductor package in R. PPI network of differentially expressed genes was constructed using the STRING database. Survival analysis was estimated based on Kaplan-Meier curve analysis. We used KOBAS 3.0 to analyze the KEGG pathway of DEGs. The dysregulated IncRNA-associated ceRNA network was constructed in colon cancer based on bioinformatics generated from miRanda, PicTar, TargetScan, miRDB and miRcode. A total of 791 DEIncRNAs and 200 DEmiRNAs were identified in colon cancer compared with matched normal tissues with thresholds of log2foldChange (FC)|>3.0and adjusted P value<0.05.Twenty DEIncRNAs were identified, may be related to tumorigenesis and/or progression of colon cancer. Nine out of 20 dysregulated IncRNA were found to be significantly associated with overall survival (Pvalue<0.05). Finally, we successfully constructed colon cancer-associated ceRNA network, including 9 colon cancer-specific IncRNAs, 13 miRNAS and 70 mRNAs. In conclusion, our study will contribute to improve the understanding of ceRNA network regulatory mechanisms in colon cancer. These identified novel IncRNAs can be as candidate prognostic biomarkers or potential therapeutic targets.

# Introduction

Colon cancer is one of the most prevalent malignancies in the world and it's the third leading causes of mortality worldwide[1]. In recent years, despite colon cancer diagnosis and treatment

PLOS ONE

**Competing interests:** The authors have declared that no competing interests exist.

have developed rapidly. However, due to its high prevalence and poor prognosis, colon cancer is still an important clinical challenge worldwide. Therefore, identification of individualized treatment strategies, including potential biomarkers and therapeutic targets to combat colon cancer, is urgently needed. The present study explores how the colon cancer-specific lncRNAs act as ceRNA to regulate target genes and participate in pathogenesis and prognosis of colon cancer.

Research results show that more than 90% of the human genome sequence can be transcribed, but do not code for proteins[2]. LncRNA is a kind of non-coding RNA, with transcripts more than 200bp in length. At present, it has been identified and cloned more than fifty thousand in the human genome, but only a small part of the biological function of lncRNAs got the experimental verification and lncRNAs still remained poorly characterized. Dysregulations of lncRNAs are discovered in different types of tumors. These lncRNAs have significantpotential application prospect in the diagnosis, treatment and prognosis in malignant tumor [3]. However, lncRNA how to mediate the expression of genes has remained unclear. Considerable effort is being made to uncover that lncRNAs how to performance diverse biological functions in tumor. In 2011, Salmena et al. proposed a ceRNA hypothesis, which described a complex post-transcriptional regulatory network in which lncRNAs, mRNAs and other RNAs act as natural miRNA sponges to suppress miRNA function by sharing one or more miRNA response elements (MREs)[4]. LncRNA that harbors similar sequence to their targeted miRNAs functions as ceRNA regulates encoding protein gene level and participate in the regulation of cell biology by sponging miRNAs. Moreover, evidence is growing that lncRNA-miRNAsmRNAs regulation networkiscrucialregulatory network is crucial regulatory molecules involved in pathogenesis and progression of tumors, including breast cancer, colon cancer, liver cancer, lung cancer and kidney cancer and other malignant tumors [5-10]. Studies have shown that lncRNAXIST was over-expressed in gastric cancer, which can be as a molecular sponge to modulate EZH2 expression by sponging miR-101 and it has been suggested possibly participating in the progression of gastric cancer[11]. Moreover, lncRNA PVT1 can influence the expression of HK2 by sponging miR-497[12]. We believe lncRNAs functions as ceRNAs deserve further exploration. In addition, the analysis of colon cancer associated lncRNA-mediated ceRNA network in a whole genome is lacking, especially studies with larger sample sizes.

By the end of 2015, TCGA collects more than 30 human tumor types and approximately 11,000 patient gene expression profiles together with relevant clinical data. Such an enormous amount of data provides a rich resource for data mining and to identify the mechanism and prognostic molecular signatures of cancer. In present research was based on TCGA, we analyzed the integrated RNA expression profiles from 341 colon cancer samples and 27 normal samples. As a result, 791 differentially expressed lncRNAs (DElncRNAs), 200 differentially expressed miRNAs (DEmiRNAs) and 1907 differentially expressed mRNAs (DEmRNAs) were identified in our study. In order to elucidate the interactions and valid potential crosstalk between RNAs, we successfully established the colon cancer associated ceRNA network based on bioinformatics generated from miRanda, PicTar, TargetScan, miRDB and miRcode, which included 9 lncRNAs, 13 miRNAs and 70mRNAs. Furthermore, relevant overall survival analyses were carried out to determine prognostic genes that possess clinical traits.

# Material and methods

# **Study population**

A total of 341 colon cancer cases were enrolled for comprehensively integrated analysis. The colon cancer level 3 RNAseq and miRNASeq of 341 colon cancer samples and corresponding clinical data (S1 File) were downloaded from TCGA database using the Data Transfer Tool

(provided by GDC Apps)(https://tcga-data.nci.nih.gov/). The RNAseq and survival data were anonymised before we accessed them. These data were included 341 colon cancer samples and 27 non-cancer samples. That sequence data was derived from Illumina HiSeqRNASeq and Illumina HiSeqmiRNASeq platforms. Our research meets publication guidelines provided by TCGA (http://cancergenome.nih.gov/publications/publicationguidelines).

# Differentially expressed analysis

Colon cancer mRNAseq and miRNASeq data derived from 368 samples, which were classified into cohort Tumor (341 colon cancer samples) and cohort Normal (27 non-cancer samples), were downloaded from TCGA (<u>S1 File</u>). In addition, we combined tumor sample and non-cancer sample data and filtered out the expressed data which closed to zero. DEmRNAs/ DEmiRNAs were identified in colon cancer samples compared with non-cancer samples using the DESeq, a Bioconductor package[13] in R software, with thresholds of |log2FC| >3.0 and adjusted P-value<0.05. We used GENCODE lncRNA annotation (V22), includes 15,877 human lncRNA genes encoding 26,414 transcripts (www.gencodegenes.org), to define and annotate the DElncRNAs.

# Construct the ceRNA network and functional annotation

In order to make clear the roles of lncRNA and miRNA with mediated ceRNA network, we built the co-expression network of lncRNAs-miRNAs-mRNAs interactions. MiRNA-targeted mRNAs were predicted by 4 programs, including miRanda, PicTar, TargetScan and miRDB. To improve accuracy, the most targeted genes were included in all the 4 datasets. To further improve the ceRNA network reliability, we retained mRNAs included in different expression of RNAs between tumor tissues and normal tissues. In addition, lncRNA-miRNA interactions were constructed based on miRcode database (http://www.mircode.org/). The ceRNA network was visualized using Cytoscape v3.5.0 software. Furthermore, KEGG pathways enrichment analysis of DEmRNAs was analyzed using KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/)[14].

# Protein-protein interaction (PPI) networks construction

Inorder to better understand protein-protein interactions between differentially expressed genes (DEGs), we constructed PPI network for DEGs using the STRING (Search Tool for the Retrieval of Interacting Genes, <u>https://string-db.org/</u>) database (minimum required interaction score>0.4)[15]. Furthermore, PPI networks were embodied using the Cytoscape v3.5.0 software. In addition, we used Cytohubb plugin to identify the hub genes in PPI network. Top 10 hub genes were identified using the ranking method of degree.

# Survival analysis

In order to identify prognostic DERNAs signature, combining the clinical data of those patients with colon cancer in TCGA, we plotted the survival curves of those samples with DElncRNAs, DEmiRNAs and DEmRNAs by "survival" package in R. The univariate survival analysis was estimated based on Kaplan Meier curve analysis. *P*-value less than 0.05 was considered as statistical significance.

# Results

#### Patient characteristics

A total of 341 patients were pathologically diagnosed as colon cancer. The median age of 341 colon cancer patients was 68 years old and forty-nine percent of them were more than 68 years

old. Gender distribution: the male was more than female (male to female ratio: 1.08). Stage II and Stage III patients with colon cancer accounted for 67.7%. All 341 colon cancer patients with detailed pathological and clinical characteristics were shown in Table 1.

#### Differentially expressed mRNAs (DEmRNAs) in colon cancer

RNAs expression profiles of colon cancer patients and corresponding clinical information were downloaded using Data Transfer Tool of TCGA database. We identified the significant DEmR-NAs and DEmiRNAs in colon cancer samples compared with the normal samples. A total of 1907 DEmRNAs were identified by DESeqpackage in R. As a result, there were 1476(77.40%) up-regulated and 431(22.60%) down-regulated DEmRNAs. A full list of DEmRNAs was shown **S1 Table**. The DEmRNAs was enriched in KEGG pathway by KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/), in order to preliminary investigation tumorigenesis of colon cancer. These DEmRNAs were mainly enriched in "Wnt signaling pathway, Transcriptional misregulation in cancer and Chemical carcinogenesis", which are closely correlated with tumor genesis (Table 2).

#### Protein-protein interaction (PPI) networks construction

Furthermore, we used STRING database to excavate the interrelationship between DEGs by constructing PPI networks. The visualization of the PPI networks was embodied using the Cytoscape software (Fig 1). Moreover, we used Cytohubb plugin of Cytoscape to identify the

Subtype	Patients n (%)
>68	167 (49.0)
≤68	174 (51.0)
Male	177 (51.9)
Female	164 (48.1)
WHITE	235 (68.9)
ASIAN	12 (3.5)
BLACK OR AFRICAN AMERICAN	66 (19.4)
UNKNOW	28 (8.2)
StageI	54 (15.8)
Stage II	131 (38.4)
Stage III	99 (29.0)
Stage IV	50 (14.7)
UNKNOW	7 (2.1)
T1	7 (2.1)
T2	56 (16.4)
T3	233 (68.3)
T4	45 (13.2)
N0	199 (58.4)
N1	84 (24.6)
NX <sup>a</sup>	58 (17.0)
M0	245 (71.6)
M1	50 (14.6)
MX <sup>a</sup>	47 (13.8)
	Subtype> $68$ $\leq 68$ MaleFemaleWHITEASIANBLACK OR AFRICAN AMERICANUNKNOWStageIStage IIStage IIIStage IVUNKNOWT1T2T3T4N0N1M1M1MX <sup>a</sup> MX <sup>a</sup>

Table 1. Clinicopathological characteristics of 341 colon cancer patients.

a: metastasis status unknown

https://doi.org/10.1371/journal.pone.0192494.t001

Pathway ID	Description	P-value	Count	
hsa04024	cAMP signaling pathway	0.023168	18	
hsa04976	Bile secretion	2.27E-07	17	
hsa00980	Metabolism of xenobiotics by cytochrome P450	6.31E-07	17	
hsa05204	Chemical carcinogenesis	1.91E-06	17	
hsa05202	Transcriptional misregulation in cance	1.30E-13	17	
hsa00982	Drug metabolism-cytochrome P450	5.92E-06	15	
hsa04974	Protein digestion and absorption	1.22E-04	15	
hsa00830	Retinol metabolism	1.78E-05	14	
hsa04970	Salivary secretion	3.58E-04	14	
hsa04310	Wnt signaling pathway	0.022686	14	

#### Table 2. DERNAs were enriched KEGG pathways in colon cancer.

https://doi.org/10.1371/journal.pone.0192494.t002

hub genes in PPI network. We think the top 10 hub genes, ALB, F2, IL8, GCG, SST, NPY, PPBP, CASR, APOB and ACTG2, which may be closely related to colon cancer pathogenesis, using the ranking method of degree (Fig 2).

#### Differentially expressed lncRNAs (DElncRNAs) in colon cancer

A total of 791 DElncRNAs were identified in our study, with thresholds of |log2FC| > 3.0 and adjusted P-value < 0.05. A full list of DElncRNAs was shown S2 Table. To enhance the data reliability, those have not been annotated in GENCODE were removed. Finally, 12 DElncRNAs were identified in colon cancer samples compared to the normal samples (Table 3). Subsequently, to explore the relationship between DElncRNAs and the prognosis of patients with colon cancer, overall survival for 12 DElncRNAs in colon cancer patients was investigated using Kaplan-Meier curve analysis. We found that 9 of 12 DElncRNAs were considered as key DElncRNAs responsible for the prognosis of colon cancer. As a result, 9 DElncRNAs were significantly associated with overall survival, lncRNA AC079612.1 were positively correlated with overall survival, while the remaining 8 DElncRNAs were negatively associated with overall survival (log-rank P < 0.05)(Fig 3).

In addition, in order to further investigate lncRNAs took part in the progression of colon cancer. We also used "DESeq" package of R to identify the DElncRNAs (|log2FC|>3.0 and adjusted P-value<0.05). As a result, a total of 20 colon cancer specific lncRNAs were significantly different in the pathological tumor-node-metastasis (TNM) stages (Stage IV +Stage III vs. StageII+StageI, T4+T3 vs. T2+T1, N3+N2 vs. N1+N0, M1 vs. M0). We noted that 11 lncRNAs, SCARNA10, RPPH1, LINC01419, ERVH48-1, RMRP,CH507-513H4.3, MIR205HG, CH507-513H4.6, LINC00400, RP11-774D14.1 and AC006050.2, were also differentially expressed in colon cancer samples compared with the normal samples (Table 4). In addition, we found 6 out of 11 DElncRNAs, LINC01419, SCARNA10, RPPH1, LINC00400, RMRP and AC006050.2, were included in all the comparisons of pathological TNM stages. Therefore, we think these 6 DElncRNAs may be significantly associated with the tumorigenesis and/or progression of colon cancer.

# Differentially expressed miRNAs (DEmiRNAs) in colon cancer

In our study, 200 DEmiRNAs were identified in colon cancer samples compared with normal samples with thresholds of |log2FC| > 3.0 and adjusted P-value< 0.05(S3 Table). There are 146 (73.00%) up-regulated and 54 (20.21%) down-regulated DEmiRNAs. A total of 18 DEmiRNAs (15 up-regulated and 3 down-regulated) were identified and maybe play an important role in pathogenesis of colon cancer (Table 5).





PLOS ONE

As same with the DElncRNAs, the overall survival for 18 DEmiRNAs in colon cancer patients was also investigated using a Kaplan-Meier curve analysis. Seven out of 18 keys DEmiRNAs were significantly associated with overall survival (log-rank P < 0.05), and 6 DEmiRNAs, mir-139, mir-197, mir-337, mir-653, mir-4709 and mir-149 were demonstrated to be associated with high levels of DEmiRNAs and with poor prognosis. On the contrary, high levels of the remaining DEmiRNAs mir-144 were associated with prolonged patient survival time (Fig 4).



Fig 2. Top 10 hub DEGs were identified in colon cancer. Each ellipse represents gene. The solid lines between DEGs represent direct connections, and dashed lines represent indirect connections.

https://doi.org/10.1371/journal.pone.0192494.g002

### ceRNA network in colon cancer

To better understand how lncRNA mediate mRNA through combining miRNA in colon cancer, a ceRNA network graph was constructed based on above data and visualized using Cytoscape v3.5.0 (Fig 5). We found that 9 DElncRNAs could interact with the 13 DEmiRNAs retrieving miRcode database (Table 6).

We searched for targeted mRNAs based on 13 miRNAs using them iRanda, PicTar, TargetScan and miRDB database. The final targeted genes were selected, which most were included in all the 4 datasets. MiRNAs targeted mRNAs not included in DEmRNAs were discarded. Finally, 70 DEmRNAs were included in ceRNA network. We retrieved from the allOnco database (http://www.bushmanlab.org/links/genelists) found that most of them targeted genes involved in newly identified ceRNA network are definite tumor-related genes, such as WT1, MAPK4, MYCN, CNTN4, ADAMTSL3, KLF4, PAX3, ALK, EPHA7, ANLN, WIF1, PAX7, CTNND2, ADAM12 and FOSL1. Moreover, we noted that some of the mRNAs

IncRNA	Gene ID	Expression change	log2 FC(T/N)	FDR
AC079612.1	ENSG00000196758	Up-regulation	3.61	1.74E-10
HOTAIR	ENSG00000228630	Up-regulation	3.13	0.00072
GAS6-AS1	ENSG00000233695	Up-regulation	3.53	5.71E-18
KCNQ10T1	ENSG00000269821	Up-regulation	3.32	1.88E-09
RP11-742B18.1	ENSG00000249001	Up-regulation	4.58	2.69E-08
WT1-AS	ENSG00000183242	Up-regulation	3.46	3.64E-05
ERVH48-1	ENSG00000233056	Up-regulation	3.05	5.85E-05
MIR205HG	ENSG00000230937	Up-regulation	4.80	0.000855
LINC00400	ENSG00000229928	Up-regulation	4.75	0.001380
ELFN1-AS1	ENSG00000236081	Up-regulation	5.00	3.90E-33
TSPEAR-AS2	ENSG00000182912	Up-regulation	3.18	8.41E-10
ESRG	ENSG00000265992	Up-regulation	3.14	0.000263

Table 3. Twelve colon cancer-specific lncRNAs in ceRNA network.

Abbreviations: T: tumor; N: normal. FC: fold change

https://doi.org/10.1371/journal.pone.0192494.t003



Fig 3. Nine lncRNAs associated with overall survival in colon cancer base on Kaplan-Meier survival curves. The horizontal axis represents overall survival time (years), vertical axis represents survival function. (9DElncRNAs were presented (P < 0.05), including KCNQ1OT1, WT1-AS, HOTAIR, GAS6-AS1, AC079612.1, ELFN1-AS1, RP11-742B18.1, TSPEAR-AS2 and ESRG).

(ANLN, CFL2, FJX1, HHIP, PANX2, SCN3A, VSNL1 and ZIC2) were also associated with overall survival in patients with colon cancer (**Fig 6**).

	ONE
--	-----

Comparisons	Up-regulated	Down-regulated
Pathologic_Stage (Stage IV +Stage III vs. StageII +StageI)	LINC01419, LINC00400, RP11-774D14.1, AC006050.2	ERVH48-1, RPPH1, RMRP, SCARNA10, MIR205HG, CH507- 513H4.6
Pathologic_T (T3+T4 vs. T1+T2)	LINC01419, LINC00400, RP11-774D14.1, AC006050.2	ERVH48-1, RPPH1, RMRP, MIR205HG, SCARNA10, CH507- 513H4.6
Pathologic_N (N2+N3 vs. N0+N1)		SCARNA10, RPPH1, RMRP, LINC01419, LINC00400, AC006050.2
Pathologic_M (M1 vs. M0)	LINC01419, LINC00400 AC006050.2, RP11- 774D14.1	RMRP, RPPH1, ERVH48-1 SCARNA10, CH507-513H4.6

Table 4. 11 lncRNAs associated with the tumorigenesis and /or progression of colon cancer.

Overall survival for DEmRNAs was also investigated base on Kaplan-Meier curve analysis. To better understand the KEGG pathways involved in ceRNA network, the mRNAs were performed using KEGG pathway analysis by KOBAS 3.0, and the top 6 KEGG pathways were significantly enriched cancer-associated signal pathways, such as "Transcriptional misregulation in cancer, Proteoglycans in cancer and Wnt signaling pathway" (Table 7).

Moreover, the newly constructed ceRNA network maybe play an important role in colon cancer pathologenesis. Among ceRNA network, lncRNAs KCNQ1OT1 and WT1-AS maybe act as key lncRNAs, which interacted with 8 miRNAs (mir-143, mir-454, mir-96, mir-182, mir-17, mir-217, mir-141 and mir-223) and 5 miRNAs (mir-139, mir-429, mir-96, mir-182 and mir-424), respectively and indirectly interacted with 47and 33 miRNAs-targeted mRNAs in this network, respectively. In addition, we confirmed that the lncRNA WT1-AS was expressed in parallel with WT1 mRNA using a regression analysis (Fig 7). In addition, we found that WT1 was miR-429 and miR-424 directly targeted gene retrieved from miRanda database. Therefore, we think WT1-AS can regulate the expression levels of gene WT1 by competing combined miRNA (miR-429 and miR-424), probably.

Name	log2 FC (T/N)	P-value	FDR
hsa-mir-139	-4.39	4.26E-103	1.68E-100
hsa-mir-197	-4.04	8.25E-178	6.49E-175
hsa-mir-149	-3.34	2.22E-36	1.59E-34
hsa-mir-223	3.21	9.72E-08	7.05E-07
hsa-mir-337	3.26	7.39E-17	1.14E-15
hsa-mir-4709	3.78	0.005798	0.021126
hsa-mir-143	3.84	2.13E-08	1.63E-07
hsa-mir-17	3.87	3.78E-20	7.62E-19
hsa-mir-32	3.92	5.52E-23	1.45E-21
hsa-mir-429	4.23	9.77E-21	2.02E-19
hsa-mir-217	4.25	1.15E-08	9.14E-08
hsa-mir-182	5.31	1.75E-21	3.84E-20
hsa-mir-141	5.47	1.07E-38	8.44E-37
hsa-mir-96	5.52	7.09E-19	1.34E-17
hsa-mir-424	5.74	1.58E-24	4.79E-23
hsa-mir-144	6.17	8.40E-17	1.28E-15
hsa-mir-454	6.19	1.78E-23	4.96E-22
hsa-mir-653	7.95	2.01E-08	1.56E-07

#### Table 5. Eighteen colon cancer specific miRNAs in ceRNA network.

Abbreviations: T: tumor; N: normal.

https://doi.org/10.1371/journal.pone.0192494.t005

# Discussion

In 2012, there were 1.4 million new cases and 694,000 deaths from colorectal cancer. It is more prevalent in developed countries, where more than 65% of cases are found[16]. The



**Fig 4.** Four miRNAs associated with overall survival in colon cancer base on Kaplan-Meier survival curves. The horizontal axis represents overall survival time (years), vertical axis represents survival function. (7 DEmiRNAs were presented (P < 0.05), including miR-139, miR-144, miR-197, miR-337, miR-653, miR-4709 and miR-149).

https://doi.org/10.1371/journal.pone.0192494.g004



Fig 5. DElncRNAs-related ceRNA network in colon cancer. The red represent increased level of expression, while the green represent decreased level of expression. Round rectangle represent lncRNAs; hexagon represent miRNAs; ellipse represent protein-coding genes; black edges indicate lncRNA-miRNA interactions.

occurrence of colorectal cancer is caused by many factors. The expression of abnormal gene expression is one of them, and the expression of molecular biological markers is considered to be related to the occurrence of colorectal cancer[17]. In recent years, many research revealed that DElncRNAs has important biological functions by regulating gene expression at various levels in multiple tumors, including epigenetic regulation, transcription regulation and post-transcription regulation[18, 19]. In addition, rising evidence showed that DElncRNAs present opportunity for different types of cancer diagnostics, prognostics and therapeutics[20]. Compared with protein-coding genes, lncRNAs have significant advantages as diagnostic and prognostic biomarkers[21]. A CeRNA hypothesis was proposed for the mechanism of tumor-genesis, providing an important new clue and the direction of research for tumor diagnosis

lncRNA	miRNAs	
WT1-AS	miR-139, miR-429, miR-182, miR-96, miR-32, miR-17, miR-217, miR-141, miR-223, miR-424	
KCNQ10T1	miR-141, miR-143, miR-17, miR-182, miR-217, miR-223, miR-454, miR-96	
GAS6-AS1	miR-182, miR-223, miR-32, miR-96	
ERVH48-1	miR-141, miR-454, miR-182, miR-96, miR-223, miR-144	
HOTAIR	miR-143, miR-17, miR-217, miR-454	
RP11-742B18.1	miR-143, miR-429, miR-139	
AC079612.1	miR-141,miR-182, miR-96, miR-223, miR-144	
LINC00400	miR-96, miR-141, miR-182, miR-424	
MIR205HG	miR-143, miR-454	

Table 6. Nine DElncRNAs interact with the 13 DEmiRNAs retrieving miRcode database.

and treatment, providing a new guiding theory[4]. It has been confirmed that lncRNAs that harbor similar sequence to their targeted miRNAs can sponge miRNAs to regulate mRNAs expression. Several studies have shown that theDElncRNAs is closely related to the pathogenesis and prognosis of tumor, and can be used as a tumor related predictors [22, 23]. In addition, research reported that up-regulation of lncRNA H19 significantly promotes epithelial to mesenchymal transition (EMT) progression and accelerates colorectal cancer cell growth by competing binding the miR-138 and miR-200a to modulate the derepression of targeted genes Vimentin, ZEB1, and ZEB2[24]. Peng et al. found lncRNA ROR can mediate the expression of miR-145 and participated in proliferation, migration and invasion of colon cancer cell[25]. Recent studies have revealed that the lncRNA PART-1 was up-regulated in colorectal cancer, functioned as a competitive endogenous RNA to regulate the targeted gene DNMT3A by sponging miR-143, participated in colorectal cancer cell proliferation and metastasis[26].

As far as we know, few researches have been performed to predict colon cancer prognosis. And there are few reliable colon cancer-specific lncRNAs as molecular biomarkers for risk stratification and detection of colon cancer. Therefore, we comprehensively integrated 341 colon cancer samples mRNA and miRNA data derived from TCGA and determined the lncRNAsassociated ceRNA network to further investigate the regulatory mechanism of lncRNAs.

In the present study, 12 DElncRNAs were identified in colon cancer samples compared with the normal samples. We found that 9 of them were significantly associated with overall survival, could be considered as a prognostic marker for colon cancer. Moreover, we noted that the lncRNA KCNQ1OT1 and WT1-AS were interacted with most of the miRNAs included in ceRNA network. Therefore, we think KCNQ1OT1 and WT1-AS may play a more significant role in the pathogenesis and prognosis of colon cancer. KCNQ1OT1 (KCNQ1 overlapping transcript 1) is a long non-coding RNA gene found in the KCNQ1 locus [27]. KCNQ1OT1 is a nuclear, 91 kb transcript, found in close proximity to the nucleolus in certain cell types[28]. It plays a major role in the transcriptional silencing of the KCNQ1 locus by regulating histone methylation[27]. The misregulation of the KCNQ1OT1 can lead to a variety of abnormalities. Over-expression of KCNQ1OT1 has been confirmed to be associated with many types of cancers, including wilms' tumor, gliomas and colorectal cancers[29-31]. Consistent with previously reported, in our study, we identified that the lncRNA KCNQ10T1 was up-regulated in 341 colon cancers compared with 27 non-tumor samples. A latest study shows that KCNQ1OT1 can be as a ceRNA to modulate miR-370 targeted gene CCNE2 by competing combined with miR-370[32]. Moreover, Xin et al. demonstrated that KCNQ10T1 can impact expression levels of caspae-1 through directly interact with miR-214[31]. In our study, our analysis found that high expression of KCNQ10T1 may compete with 8 key DEmiRNAs (miR-141, miR-143, miR-17, miR-182, miR-217, miR-223, miR-454 and miR-96)

to mediate the expression of target genes. Moreover, patients with highly expressed KCNQ10T1 had shorter survival time than those with low expression.

Furthermore, we consider WT1-AS may also play a key role in pathogenesis and prognosis of colon cancer. lncRNA WT1-AS is the antisense transcript of Wilms tumor 1 (WT1), which



Fig 6. Eight protein-coding genes associated with overall survival in colon cancerbase on Kaplan-Meier survival curves. The horizontal axis represents overall survival time (years), vertical axis represents survival function. (8 DEmRNAs were presented (P < 0.05), including ANLN, CFL2, FJX1, HHIP, PANX2, SCN3A, VSNL1 and ZIC2).

https://doi.org/10.1371/journal.pone.0192494.g006

Pathway ID	Description	Genes
hsa04310	Wnt signaling pathway	PRKACB,FOSL1,FZD10,SFRP1,WIF1
hsa04360	Axon guidance	CFL2,EPHA5,SEMA6A,EPHA7,SEMA6D
hsa05202	Transcriptional misregulation in cancer	PAX3,PAX7,MYCN,EYA1
hsa04976	Bile secretion	PRKACB,SLC4A4,ATP1A2
hsa04971	Gastric acid secretion	PRKACB,ATP1A2,MYLK
hsa05205	Proteoglycans in cancer	ANK2,PRKACB,HPSE2,FZD10

Table 7. Top 6 KEEG pathways enriched by the coding genes involved in ceRNA network.

located upstream of WT1. WT1-AS is found on chromosome 11 and is expressed in kidney [33]. K. Moorwood et al. also found WT1-AS was expressed in kidney, positively regulated WT1 protein levels, and participated in the developing kidney[34]. Studies have shown that WT1-AS was expressed in parallel with WT1 mRNA. Moreover, several studies demonstrated WT1-AS may play an important role in the progression of ovarian cancer, acute myeloid leukemia and gastric cancer[35, 36]. In present research, WT1-AS was up-regulated in 341 colon cancer samples compared to 27 non-cancer samples. Survival analysis confirmed that high expression of WT1-AS was associated with poor prognosis in patient with colon cancer. Interesting, in our study, we performed a regression analysis between the expression levels of WT1-AS and WT1 targeted gene, which were both included in the newly constructed ceRNA network. The results revealed a very strong positive correlation between WT1-AS and WT1 expression levels. Furthermore, we noted that WT1 was miR-429 and miR-424 directly targeted gene retrieved from miRanda database. Based on the above results, we think that lncRNA WT1-AS may compete with 2 key DEmiRNAs (miR-429 and miR-424) to mediate the expression of WT1 target gene. However, these findings need further research to identify whether the lncRNA WT1-AS can be as a ceRNA to regulate the expression of gene WT1 by sponging miR-429 and miR-424.

Although lncRNA has received wide attentions in recent years, miRNAs also warrant increased attention. There is no doubt that tumorigenic-related pathways of research based on regulation of miRNAs are indispensable. Dysregulated expression miRNAs is reported to play various roles in carcinogenesis. In the present study, we identified tumor initiation-related miRNAs in colon cancer. Moreover, miR-139 and miR-144 not only involved in ceRNA network, but also closely associated with survival of colon cancer patient. MiR-139 aberration is found in wide variety types of human tumor, including gastric cancer, breast cancer, bladder cancer, hepatocellular carcinoma, ovarian cancer, colorectal cancer and so on. It has been confirmed that miR-139 was significantly down-regulated in colorectal cancer and play a crucial role in colorectal cancer metastasis and progression[37]. Consistent with previously reported, we confirmed miR-139 was down-regulated in 341 colon cancers compared with non-cancer samples. However, our survival analysis results indicated that the prognosis of colon cancer patients with high expression of miR-139 is poor. In addition, previous reports have revealed that the differential expression of miR-144 was associated with bladder cancer cell proliferation by regulating Wnt signaling pathway [38]. Moreover, Lwaya et al. [39] found miR-144 is related to the prognosis of colorectal cancer patients. In our study, survival analysis demonstrated that low expression of mir-144 leads to poor prognosis of colon cancer patients. However, those findings needed more research to identify whether those miRNAs have a specific role in tumorigenesis and prognosis of colon cancer.

The miRNAs-targeted genes involved in ceRNA network were performed KEGG pathways enrichment analysis results showed that targeted genes were mainly enriched Wnt signaling pathway, Transcriptional misregulation in cancer and Proteoglycans in cancer<sup>°</sup> cancer-



Pearson's correlation=0.728

Fig 7. Regression analysis between the expression levels of WT1-AS and WT1 in ceRNA network. The horizontal axis represents WT1 expression, and vertical axis represents WT1-AS expression.

associated pathways. These genes, SLC4A4, FOSL1, EYA1, FZD10, SFRP1, ATP1A2 and WIF1, not only enriched in KEGG pathway but also included in the ceRNA network. We found that these genes play a key regulatory role in the occurrence and development of tumors. The study found that FZD10 silencing can inhibit aberrant activation of Wnt3-FZD10-s. a specific role in tumorigenesis and prognosis of colon[40]. In addition, it has been confirmed that WIF1 (Wnt inhibitory factor-1), a Wnt pathway inhibitor, can inhibit human invasive urinary bladder cancer cells growth [41]. Interestingly, in our study, we also found that FZD10 and WIF1 were enriched in Wnt signaling pathway. Moreover, T Fukui et al. studies demonstrated that the SFRP1 gene is frequently downregulated and suppresses tumor growth activity of lung cancer cells<sup>[42]</sup>. They suggested that SFRP1 is a candidate tumor suppressor gene for lung cancer. Parks SK et al. found SLC4A4 plays a role in growth and migration of colon and breast cancer cells[43]. In addition, we noted that the targeted genes ANLN, CFL2, FJX1, HHIP, PANX2, SCN3A, VSNL1 and ZIC2, were also associated with overall survival of colon cancer patients (P value<0.05). In our study, we revealed how specific lncRNAs interact with miRNAs and coding genes through the successful construction of lncRNA-miRNAs-mRNA ceRNA network of colon cancer.

# Conclusion

In conclusion, 20 cancer-specific lncRNAs were identified from hundreds of candidate lncRNAs in large scale colon cancer samples. The research indicates that dysregulation of ceRNA network can lead to tumorigenesis[44]. We found that some lncRNA were remarkably associated with overall survival in patient with colon cancer. Importantly, we have successfully

constructed a lncRNA-associated ceRNA network, which brings a new approach to lncRNA research in colon cancer and provides novel lncRNAs as candidate prognosis biomarkers or potential therapeutic targets.

# Supporting information

**S1 Table.** A full list of DEmRNAs in colon cancer. (CSV)

S2 Table. A full list of DElncRNAs in colon cancer. (CSV)

**S3 Table. A full list of DEmiRNAs in colon cancer.** (CSV)

**S1** File. The minimal anonymized data set of colon cancer and non-cancer. (ZIP)

# **Author Contributions**

Conceptualization: Fengxi Li, Xianghua Wu.

Data curation: Fengxi Li, Qian Li, Xianghua Wu.

Formal analysis: Fengxi Li, Qian Li.

Investigation: Fengxi Li, Qian Li.

Methodology: Fengxi Li, Xianghua Wu.

Project administration: Fengxi Li, Xianghua Wu.

Resources: Fengxi Li.

Software: Fengxi Li, Qian Li.

Supervision: Qian Li.

Validation: Fengxi Li, Qian Li, Xianghua Wu.

Visualization: Fengxi Li, Qian Li.

Writing - original draft: Fengxi Li.

Writing - review & editing: Xianghua Wu.

#### References

- Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet. 2014; 383(9927):1490–502. https://doi.org/10. 1016/S0140-6736(13)61649-9 PMID: 24225001
- Xiao K, Yu Z, Li X, Li X, Tang K, Tu C, et al. Genome-wide Analysis of Epstein-Barr Virus (EBV) Integration and Strain in C666-1 and Raji Cells. Journal of Cancer. 2016; 7(2):214–24. https://doi.org/10.7150/ jca.13150 PMID: 26819646
- 3. Yarmishyn AA, Kurochkin IV. Long noncoding RNAs: a potential novel class of cancer biomarkers. Frontiers in genetics. 2015; 6:145. Epub 2015/05/09. https://doi.org/10.3389/fgene.2015.00145 PMID: 25954300; PubMed Central PMCID: PMCPmc4407501.
- 4. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta stone of a hidden RNA language? Cell. 2011; 146(3):353–8. https://doi.org/10.1016/j.cell.2011.07.014 PMID: 21802130
- Wang J, Liu X, Wu H, Ni P, Gu Z, Qiao Y, et al. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. Nucleic Acids Research. 2010; 38 (16):5366–83. https://doi.org/10.1093/nar/gkq285 PMID: 20423907

- 6. Liu X, Sun M, Nie F, Ge Y, Zhang E, Yin D, et al. Lnc RNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. Molecular Cancer. 2014; 13(1):92.
- Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature. 2010; 465(7301):1033. https://doi. org/10.1038/nature09144 PMID: 20577206
- Yu G, Yao W, Gumireddy K, Li A, Wang J, Xiao W, et al. Pseudogene PTENP1 functions as a competing endogenous RNA to suppress clear-cell renal cell carcinoma progression. Molecular Cancer Therapeutics. 2014; 13(12):3086–97. https://doi.org/10.1158/1535-7163.MCT-14-0245 PMID: 25249556
- Hou P, Zhao Y, Li Z, Yao R, Ma M, Gao Y, et al. LincRNA-ROR induces epithelial-to-mesenchymal transition and contributes to breast cancer tumorigenesis and metastasis. Cell Death & Disease. 2014; 5(5): e1287.
- Song X, Cao G, Jing L, Lin S, Wang X, Zhang J, et al. Analysing the relationship between IncRNA and protein-coding gene and the role of IncRNA as ceRNA in pulmonary fibrosis. Journal of Cellular & Molecular Medicine. 2014; 18(6):991–1003.
- 11. Chen D, Ju H, Lu Y, Chen L, Zeng Z, Zhang D, et al. Long non-coding RNA XIST regulates gastric cancer progression by acting as a molecular sponge of miR-101 to modulate EZH2 expression. Journal of Experimental & Clinical Cancer Research Cr. 2016; 35(1):142.
- Song J, Wu X, Liu F, Li M, Sun Y, Wang Y, et al. Long non-coding RNA PVT1 promotes glycolysis and tumor progression by regulating miR-497/HK2 axis in osteosarcoma. Biochemical & Biophysical Research Communications. 2017.
- Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biology. 2010; 11(10):R106. https://doi.org/10.1186/gb-2010-11-10-r106 PMID: 20979621
- Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res. 2011; 39(Web Server issue):W316–22. Epub 2011/07/26. https://doi.org/10.1093/nar/gkr483 PMID: 21715386; PubMed Central PMCID: PMCPmc3125809.
- Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. 2017; 45(D1): D362–d8. https://doi.org/10.1093/nar/gkw937 PMID: 27924014.
- 16. Stewart B, Wild C. World cancer report 2014: International Agency for Research on Cancer; 2014.
- Peeters M, Price T. Biologic therapies in the metastatic colorectal cancer treatment continuum–Applying current evidence to clinical practice. Cancer Treatment Reviews. 2012; 38(5):397–406. https://doi. org/10.1016/j.ctrv.2011.08.002 PMID: 21899955
- Muers M. RNA: Genome-wide views of long non-coding RNAs. Nature Reviews Genetics. 2011; 12 (11):742.
- Caley DP, Pink RC, Trujillano D, Carter DR. Long noncoding RNAs, chromatin, and development. Thescientificworldjournal. 2010; 10(1):90.
- Zhang F, Zhang L, Zhang C. Long noncoding RNAs and tumorigenesis: genetic associations, molecular mechanisms, and therapeutic strategies. Tumor Biology. 2016; 37(1):163–75. <u>https://doi.org/10.1007/s13277-015-4445-4</u> PMID: 26586396
- Hauptman N, Glavač D. Long Non-Coding RNA in Cancer. International Journal of Molecular Sciences. 2013; 14(3):4655–69. https://doi.org/10.3390/ijms14034655 PMID: 23443164
- Augoff K, Mccue B, Plow EF, Sossey-Alaoui K. miR-31 and its host gene IncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. Molecular Cancer. 2012; 11(1):5.
- Li H, Yu B, Li J, Su L, Yan M, Zhu Z, et al. Overexpression of IncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. Oncotarget. 2014; 5(8):2318–29. https://doi.org/10.18632/oncotarget. 1913 PMID: 24810858
- Liang WC, Fu WM, Cheuk-Wa W, Wang Y, Wang WM, Hu GX, et al. The IncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer. Oncotarget. 2015; 6(26):22513–25. https://doi.org/10.18632/oncotarget.4154 PMID: 26068968
- Zhou P, Sun L, Liu D, Liu C, Sun L. Long Non-Coding RNA lincRNA-ROR Promotes the Progression of Colon Cancer and Holds Prognostic Value by Associating with miR-145. Pathology & Oncology Research. 2016; 22(4):1–8.
- **26.** Hu Y, Ma Z, He Y, Liu W, Su Y, Tang Z. PART-1 functions as a competitive endogenous RNA for promoting tumor progression by sponging miR-143 in colorectal cancer. Biochemical & Biophysical Research Communications. 2017.
- 27. Kanduri C. Kcnq1ot1: a chromatin regulatory RNA. Seminars in Cell & Developmental Biology. 2011; 22 (4):343.

- Fedoriw AM, Mauro CJ, Mu W, Della Y, Terry M. Differentiation-Driven Nucleolar Association of the Mouse ImprintedKcnq1Locus. G3 Genesgenetics. 2012; 2(12):1521–8.
- Yoshizawa S, Fujiwara K, Sugito K, Uekusa S, Kawashima H, Hoshi R, et al. Pyrrole-imidazole polyamide-mediated silencing of KCNQ1OT1 expression induces cell death in Wilms' tumor cells. International journal of oncology. 2015; 47(1):115. https://doi.org/10.3892/ijo.2015.3018 PMID: 25998555
- **30.** Sunamura N, Ohira T, Kataoka M, Inaoka D, Tanabe H, Nakayama Y, et al. Regulation of functional KCNQ1OT1 IncRNA by β-catenin. Scientific Reports. 2016; 6(2):20690.
- Xin J, Hao J, Yan S, Guo Y, Hong Z. Long Non-Coding RNA KCNQ1OT1 Promotes Cataractogenesis via miR-214 and Activation of the Caspase-1 Pathway. Cellular Physiology & Biochemistry International Journal of Experimental Cellular Physiology Biochemistry & Pharmacology. 2017; 42(1):295.
- Wei G, Zheng J, Liu X, Liu Y, Guo J, Gao Y, et al. Knockdown of Long Non-Coding RNA KCNQ1OT1 Restrained Glioma Cells' Malignancy by Activating miR-370/CCNE2 Axis. Frontiers in Cellular Neuroscience. 2017; 11.
- Huang A, Campbell CE, Bonetta L, Mcandrews-Hill MS, Chilton-Macneil S, Coppes MJ, et al. Tissue, Developmental, and Tumor-Specific Expression of Divergent Transcripts in Wilms Tumor. Science. 1990; 250(4983):991–4. PMID: 2173145
- Moorwood K, Charles AK, Salpekar A, Wallace JI, Brown KW, Malik K. Antisense WT1 transcription parallels sense mRNA and protein expression in fetal kidney and can elevate protein levels in vitro. Journal of Pathology. 1998; 185(4):352–9. https://doi.org/10.1002/(SICI)1096-9896(199808) 185:4<352::AID-PATH119>3.0.CO;2-# PMID: 9828833
- Dallosso AR, Hancock AL, Malik S, Salpekar A, King-Underwood L, Pritchard-Jones K, et al. Alternately spliced WT1 antisense transcripts interact with WT1 sense RNA and show epigenetic and splicing defects in cancer. RNA. 2007; 13(12):2287–99. https://doi.org/10.1261/rna.562907 PMID: 17940140
- 36. Du T, Zhang B, Zhang S, Jiang X, Zheng P, Li J, et al. Decreased expression of long non-coding RNA WT1-AS promotes cell proliferation and invasion in gastric cancer. Biochimica Et Biophysica Acta. 2016; 1862(1):12–9. https://doi.org/10.1016/j.bbadis.2015.10.001 PMID: 26449525
- Song M, Yin Y, Zhang J, Zhang B, Bian Z, Quan C, et al. MiR-139-5p inhibits migration and invasion of colorectal cancer by downregulating AMFR and NOTCH1. Protein & cell. 2014; 5(11):851–61. Epub 2014/08/26. https://doi.org/10.1007/s13238-014-0093-5 PMID: 25149074; PubMed Central PMCID: PMCPmc4225484.
- Guo Y, Ying L, Tian Y, Yang P, Zhu Y, Wang Z, et al. miR-144 downregulation increases bladder cancer cell proliferation by targeting EZH2 and regulating Wnt signaling. The FEBS journal. 2013; 280 (18):4531–8. Epub 2013/07/03. https://doi.org/10.1111/febs.12417 PMID: 23815091.
- Iwaya T, Yokobori T, Nishida N, Kogo R, Sudo T, Tanaka F, et al. Downregulation of miR-144 is associated with colorectal cancer progression via activation of mTOR signaling pathway. Carcinogenesis. 2012; 33(12):2391–7. Epub 2012/09/18. https://doi.org/10.1093/carcin/bgs288 PMID: 22983984.
- Gong C, Qu S, Lv XB, Liu B, Tan W, Nie Y, et al. BRMS1L suppresses breast cancer metastasis by inducing epigenetic silence of FZD10. Nature Communications. 2014; 5(7):5406.
- Tang Y, Simoneau AR, Liao WX, Yi G, Hope C, Liu F, et al. WIF1, a Wnt pathway inhibitor, regulates SKP2 and c-myc expression leading to G1 arrest and growth inhibition of human invasive urinary bladder cancer cells. Molecular Cancer Therapeutics. 2009; 8(2):458. https://doi.org/10.1158/1535-7163. MCT-08-0885 PMID: 19174556
- **42.** Fukui T, Kondo M, Ito G, Maeda O, Sato N, Yoshioka H, et al. Transcriptional silencing of secreted frizzled related protein 1 (SFRP1) by promoter hypermethylation in non-small-cell lung cancer. Oncogene. 2005; 24(41):6323–7. https://doi.org/10.1038/sj.onc.1208777 PMID: 16007200
- **43.** Parks SK, Pouyssegur J. The Na(+)/HCO3(-) Co-Transporter SLC4A4 Plays a Role in Growth and Migration of Colon and Breast Cancer Cells. Journal of Cellular Physiology. 2015; 230(8):1954–63. https://doi.org/10.1002/jcp.24930 PMID: 25612232
- Dai Q, Jixia LI, Zhou K, Liang T. Competing endogenous RNA: A novel posttranscriptional regulatory dimension associated with the progression of cancer. Oncology Letters. 2015; 10(5):2683. https://doi. org/10.3892/ol.2015.3698 PMID: 26722227