Construction and Analysis of a ceRNA Network Reveals Potential Prognostic Markers in Colorectal Cancer

Li Guo[1], Guowei Yang[1], Yihao Kang[1], Sunjing Li[1], Rui Duan[2], Lulu Shen[2], Wenwen Jiang[1], Bowen Qian[1], Zibo Yin[1] and Tingming Liang[2,3]*

[1] Department of Bioinformatics, Smart Health Big Data Analysis and Location Services Engineering Lab of Jiangsu Province, School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China, [2] School of Life Science, Jiangsu Key Laboratory for Molecular and Medical Biotechnology, Nanjing Normal University, Nanjing, China, [3] Changzhou Institute of Innovation & Development, Nanjing Normal University, Nanjing, China

Colorectal cancer (CRC) is one of the leading causes of cancer-related deaths worldwide and is derived from an accumulation of genetic and epigenetic changes. This study explored potential prognostic markers in CRC *via* the construction and in-depth analysis of a competing endogenous RNA (ceRNA) network, which was generated through a three-step process. First, we screened candidate hub genes in CRC as the primary gene markers to survey their related regulatory non-coding RNAs, miRNAs. Second, the interacting miRNAs were used to search for associated lncRNAs. Thus, candidate RNAs were first constructed into ceRNA networks based on close associations with miRNAs. Further analysis at the isomiR level was also performed for each miRNA locus to understand the detailed expression patterns of the multiple variants. Finally, RNAs were performed an in-depth analysis of expression correlations, which contributed to further screening and validation of potential RNAs with close correlations to each other. Using this approach, nine hub genes, 13 related miRNAs, and 29 candidate lncRNAs were collected and used to construct the ceRNA network. Further in-depth analysis identified the MFAP5-miR-200b-3p-AC005154.6 axis as a potential prognostic marker in CRC. MFAP5 and miR-200b-3p have previously been reported to play important roles in tumorigenesis. These RNAs showed potential prognostic values, and the combination of them may have more sensitivity than using them alone. In conclusion, MFAP5, miR-200b-3p, and AC005154.6 may have potential prognostic value in CRC and may provide a prognostic reference for this patient population.

Keywords: colorectal cancer, competing endogenous RNA, ceRNA network, prognostic marker, long noncoding RNA, miRNA

HIGHLIGHTS

- Nine hub genes were screened in colorectal cancer, and then these were used to survey related microRNAs.
- Long non-coding RNAs were screened based on associations with these miRNAs. The full set of candidate RNAs were constructed into a competing endogenous RNA network.
- Further in-depth analysis verified the MFAP5-miR-200b-3p-AC005154.6 axis as a potential prognostic marker for colorectal cancer.

INTRODUCTION

Colorectal cancer is one of the most common malignancies. Although significant advancements have been made in the early diagnosis and treatment of CRC, there remains an increased risk of cancer-related death from CRC in the United States (Rahman et al., 2015; Callahan et al., 2019). CRC also has a high incidence and is the second leading cause of cancer-related deaths in European countries (Altobelli et al., 2014). The annual rate of new CRC diagnoses is increasing worldwide (Ferlay et al., 2015), but most patients can be treated at early stages (Pawa et al., 2011). Both environmental and genetic factors influence the occurrence and development of cancer. For example, the pathological progression of CRC is a multistep processes that is caused by the accumulation of genetic alterations, primarily gene mutations and epigenetic changes (Worthley et al., 2007). Chronic infections and the ensuing inflammation also contribute to pathophysiological processes, tumor initiation, and progression. Additionally, inflammation is a crucial hallmark of cancer that is caused by multiple factors (Sun and Kato, 2016; Long et al., 2017). Currently, surgery is the main therapy for localized CRC, and adjuvant chemotherapy is also used for many patients.

Some markers have been shown to have prognostic or predictive value for colon cancer patients, meaning they can contribute to making therapeutic decisions with greater precision for specific patients. For example, Ki-67 and p53 have potential prognostic value in Dukes' B and C colon cancer (Allegra et al., 2003), *KRAS* and *BRAF* have potential prognostic value in stage II and III resected colon cancer (Roth et al., 2010), the *BRAF* V600E mutation has been shown to be an independent prognostic factor for survival in stage II and stage III colon cancer patients (Farina-Sarasqueta et al., 2010), and *SNORA42* may be a prognostic biomarker in CRC (Okugawa et al., 2017). Indeed, screening potential prognostic biomarkers from multiple molecular levels is a crucial step for cancer therapy. Mounting evidence has indicated that miRNAs contribute to multiple pathophysiological processes and may be potential biomarkers for cancer diagnostics and therapy (Tong and Nemunaitis, 2008; Valihrach et al., 2019). Another type of non-coding RNA (ncRNA), lncRNAs, are implicated in diverse biological processes, especially epigenetic regulation (Lee, 2012; Morlando and Fatica, 2018). For example, miR-425-5p may be a potential prognostic biomarker for cervical cancer (Sun et al., 2017), and miRNA-based prognostic biomarkers have been found in other cancers, such as in pancreatic cancer (Guo et al., 2018). Integrating prognostic biomarkers of multiple classes, such as mRNA-miRNA-DNA methylation (Robles et al., 2015) and lncRNA-miRNA-mRNA (Zhang et al., 2018; Wang et al., 2019) can provide more references for cancer prognosis based on the potential cross-talk between the different molecular subtypes.

Additionally, recent studies have shown that many lncRNAs can compete with mRNAs for binding to miRNAs, acting as potential ceRNAs, which contribute to disease development (Cesana et al., 2011; Tan et al., 2015; Lu et al., 2016). The potential cross-talk between miRNA-lncRNA-mRNA has been widely studied (Qian et al., 2018; Chen Y. et al., 2019), and ceRNA network analysis has proven to be an effective method of screening potential prognostic biomarkers in diverse cancer types (Liu Z. et al., 2019; Song et al., 2019; Yao et al., 2019). Herein, we aimed to screen and identify potential prognostic biomarkers in CRC by constructing and analyzing a ceRNA network that was based on the integrative analysis of multiple genomics datasets (**Figure 1A**). According to potential relationships among diverse RNAs, we screened out a potential biomarker for cancer prognosis, the miR-200b-3p-MFAP5-AC005154.6 axis, which could potentially provide prognostic or predictive information for CRC patients.

MATERIALS AND METHODS

Data Sources

High-throughput sequencing data of diverse cancers were obtained from TCGA[1] using the "TCGAbiolinks" package (Colaprico et al., 2016)[2]. For CRC, there were over 500 individuals, including 480 tumor and 41 normal samples. To

---

[1] https://tcga-data.nci.nih.gov/tcga/
[2] http://doi.org/10.1093/nar/gkv1507
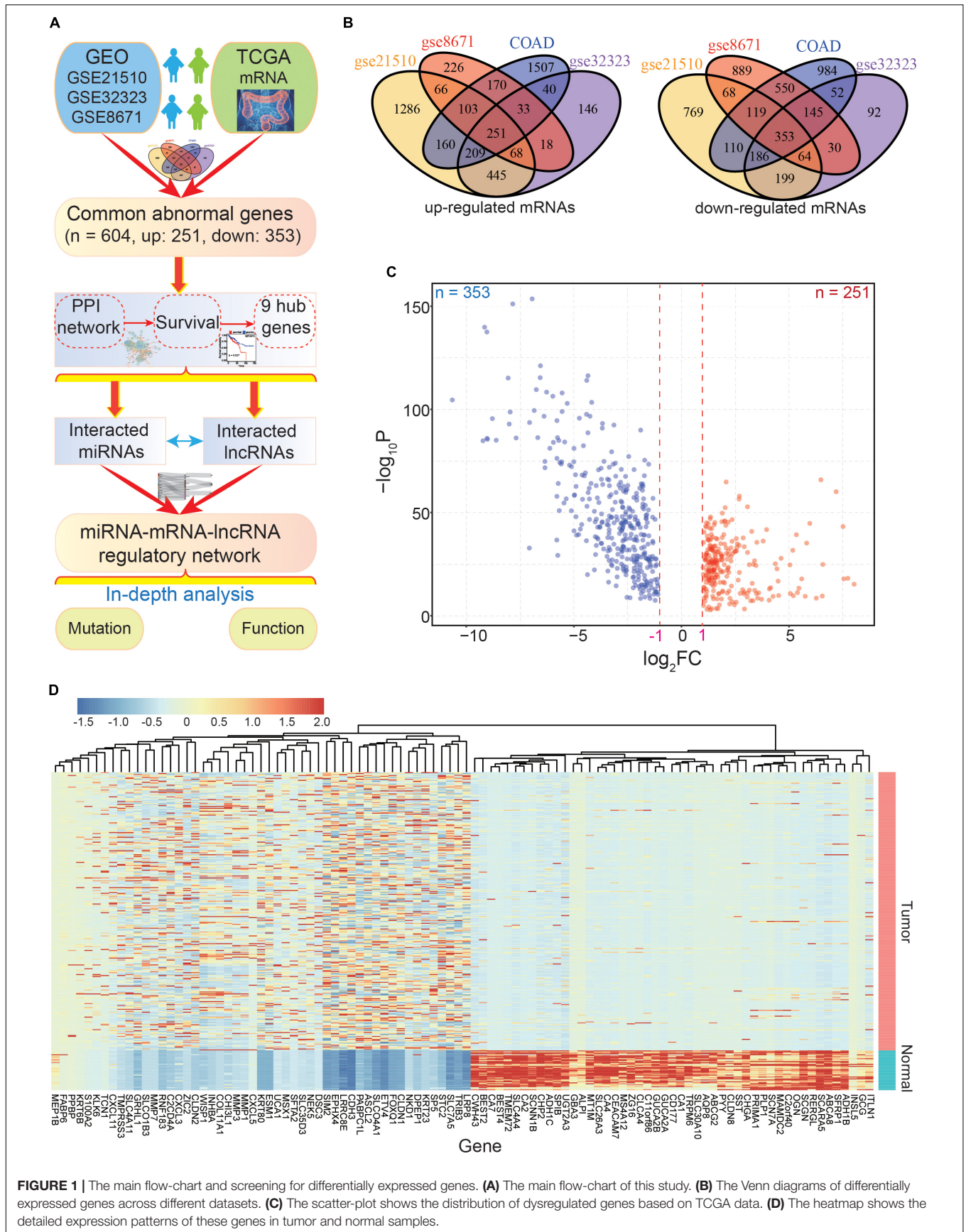
FIGURE 1 | The main flow-chart and screening for differentially expressed genes. (A) The main flow-chart of this study. (B) The Venn diagrams of differentially expressed genes across different datasets. (C) The scatter-plot shows the distribution of dysregulated genes based on TCGA data. (D) The heatmap shows the detailed expression patterns of these genes in tumor and normal samples.

determine the DEGs common to different CRC samples, CRC datasets were also collected from GEO, including GSE32323 (Khamas et al., 2012) with 17 normal and 17 tumor samples, GSE21510 (Tsukamoto et al., 2011) with 44 normal and 104 tumor samples, and GSE8671 (Sabates-Bellver et al., 2007) with 32 normal and 32 tumor samples (**Supplementary Table S1**).

Screening Potential Hub Genes

Based on TCGA and GEO datasets, the common DEGs were primarily identified using the criteria: $| \log_2 FC| > 1$ and padj < 0.05. The screened genes were first queried for their biological function using The DAVID version 6.8 (Huang et al., 2009), and z scores (Walter et al., 2015) in biological processes of KEGG pathways were estimated using the following formula:

$$z \text{ score} = \frac{(up - down)}{\sqrt{count}}$$

where up and down were the numbers of up- and down-regulated genes, respectively, and count was the total number of involved genes.

To understand the potential function of these DEGs in CRC physiology, we also analyzed their contribution to the hallmarks of cancer (Subramanian et al., 2005)[3], CGC (Futreal et al., 2004)[4], core essential genes (obtained from the common genes of Hart et al. (2015); Blomen et al. (2015), and Wang et al. (2015)), and oncogenes and TSGs (Vogelstein et al., 2013). To understand the detailed expression patterns of the relevant genes, their distribution in KEGG pathways were also queried, and significantly enriched pathways were further presented using Pathview (Luo and Brouwer, 2013; Luo et al., 2017).

To survey the potential hub genes in CRC, PPI networks were constructed based on the DEGs using the STRING database[5] with default parameters (Szklarczyk et al., 2019). PPI networks were constructed with up- and down-regulated genes. For the up-regulation PPI network, the key candidate genes were first screened based on two potential modules using the MCODE plug-in in Cytoscape 3.7.1 (Shannon et al., 2003). For the down-regulation PPI network, the CytoHubba plug-in in Cytoscape was used to display the top 100 potential hub genes according to node degrees.

Potential Prognostic Values of Candidate Genes

It was important to next validate the potential prognostic values of the screened hub genes in CRC. Survival analyses were used to estimate the correlations of the candidate genes (including candidate miRNAs and lncRNAs) with prognoses. The clinical data of CRC cases, including survival status, cancer stage and grade, survival time, and molecular subtype, were obtained from TCGA using the "TCGAbiolinks" package (Colaprico et al., 2016). The log-rank test was used to estimate potential differences, and statistical significance was set at $p < 0.05$.

Furthermore, prognostic results for candidate genes were also obtained from the GEPIA[6] (Tang et al., 2017) and StarBase (Li et al., 2014) databases.

Screening Related miRNAs and lncRNAs Based on the Hub mRNAs

Candidate hub mRNAs with potential prognostic value in CRC were used to screen related miRNAs based on biological interactions. The miRNA:mRNA interactions were first collected from StarBase (Li et al., 2014), and then were further filtered based on their prognostic results. Next, the selected miRNAs were used to survey associated lncRNAs from StarBase (Li et al., 2014) and miRNet (Fan and Xia, 2018). The identified lncRNAs were further analyzed for significant correlations with cancer prognosis.

Construction and In-Depth Analysis of the ceRNA Network

The screened hub genes, interacting miRNAs, and associated lncRNAs, were constructed into a ceRNA network based on their regulatory relationships using the R package of "networkD3"[7].

According to the constructed primary ceRNA network, an in-depth analysis was performed for the involved RNAs, despite all of them being potential prognostic markers. The miRNA-mRNA interactions were further estimated based on their expression patterns and correlations, and detailed isomiR expression patterns were investigated because of the phenomenon of multiple isomiRs in the miRNA world. An increasing number of studies have shown that the small ncRNAs are not single sequences, but contain a series of multiple sequences with diverse expression patterns (Morin et al., 2008; Neilsen et al., 2012; Tan et al., 2014; Guo and Liang, 2018; Desvignes et al., 2020). Indeed, these multiple isomiRs are regulatory molecules in coding-non-coding RNA networks; therefore, it is necessary to discuss the detailed expression of isomiRs in this study. Moreover, a dysregulated expression pattern was also an important factor to identify RNAs as potential prognostic markers, and the final screened RNAs were queried for their expression patterns across diverse cancer types.

Statistical Analysis and Network Visualization

Differentially expressed gene profiles were estimated using DESeq2 (Love et al., 2014), and hypothesis testing in relevant analyses, mainly including a Wilcoxon rank-sum test, a Kruskal–Wallis test, and a paired t-test, were used to estimate the potential differences between groups. Interactions between different RNAs, PPI networks, and coding-non-coding RNA regulatory networks were visualized using Cytoscape 3.7.1 (Shannon et al., 2003). Venn distributions were generated using a publicly available tool[8]. All statistical analyses were analyzed using the R programming language (version 3.6.1).

---

[3]http://software.broadinstitute.org/gsea/msigdb/
[4]http://cancer.sanger.ac.uk/census
[5]http://string-db.org/

[6]http://gepia.cancer-pku.cn/detail.php
[7]https://CRAN.R-project.org/package=networkD3
[8]http://bioinformatics.psb.ugent.be/webtools/Venn/

RESULTS

Survey of the Common Expression Landscape in CRC

To identify the common DEGs in CRC with higher confidence levels, we performed an integrative analysis of GEO and TCGA datasets to find the commonly dysregulated genes (**Figure 1A**). Based on the four CRC datasets, a total of 604 common dysregulated mRNAs, including 251 up- and 353 down-regulated transcripts, were obtained (**Figures 1B–D**). Indeed, most of the DEGs were found in two or three datasets, and some were detected in a specific dataset. Based on the detailed expression patterns in TCGA, these DEGs were significantly divergent between normal and tumor samples (**Figure 1D**).

Functional Analysis of the DEGs

Although a series of common DEGs in CRC were collected based on multiple datasets, their potential functional implications remained unknown. Therefore, it was necessary to next investigate their potential cellular functions, as this would contribute to surveying the potential hub genes. For these common DEGs, significant GO terms and KEGG pathways were enriched (**Figure 2**). Specifically, seven significant biological processes terms were enriched, including mitotic nuclear division, cell division, and DNA replication (**Figures 2A,B**), indicating potential roles of these DEGs in tumorigenesis. Among these DEGs, 92 were identified as core essential genes, and 25 were CGC genes (**Figure 2C**), suggesting critical roles in cancer-associated pathways. The cell cycle pathway was the most common pathway among the involved genes and was also an enriched KEGG pathway *via* DAVID analysis (**Figures 2D,E**). Further analysis based on the cell cycle pathway in CRC showed crucial genes with abnormal expression patterns (**Figure 2F**), implying an important role for these abnormally expressed genes in the occurrence and development of CRC.

Screening Hub Genes Based on PPI Interaction Networks and Potential Value to Cancer Prognosis

To screen the CRC hub genes, PPI networks were constructed based on the common up- and down-regulated genes, respectively (**Figure 3**). The two modules were enriched in the network of up-regulated genes (**Figure 3A**), and these involved genes (module 1 contained 64 genes and module 2 contained 18 genes, **Supplementary Tables S2, S3**) that might be candidate hub genes worthy of performing further analysis regarding cancer prognosis. Simultaneously, the top 100 genes (≥4 interactions with other genes) with the highest scores were the potential candidate hub genes based on the network of down-regulated genes (**Supplementary Table S4** and **Figure 3B**). Thus, a total of 182 candidate hub genes were collected from PPI interaction networks, and these genes were further investigated in survival analyses to identify potential prognostic correlations with these genes.

Based on survival analyses of these primarily screened candidate hub genes in CRC, we obtained 15 hub genes (**Supplementary Table S5**, including 12 down- and three up-regulated genes) with significant correlations to prognosis (**Figures 4A,B**). Among these genes, most had multiple interactions with other genes, suggesting important roles in biological processes as hub genes. These 15 candidate genes were significantly dysregulated in CRC (**Figure 4C**), and most showed consistent expression patterns across many cancers, although some showed the opposite expression patterns in different tissues (**Figure 4D**). For example, AQP8 showed various expression patterns in diverse tissues, suggesting specific roles for AQP8 in different cancers. Indeed, AQP8 may inhibit colorectal cancer growth and metastasis by decreasing PI3K/AKT signaling and PCDH7 expression (Wu Q. et al., 2018); thus, AQP8 status in colorectal carcinoma has a potential clinical significance (Wang et al., 2012).

Next, these 15 hub genes were queried for correlations with cancer prognosis in the GEPIA database, which revealed that nine had consistent survival results. Thus, these nine genes (AQP8, CCNB1, CNN1, FABP4, KPNA2, MFAP5, PC, SCG2, and TAGLN) were selected as the final hub genes in CRC. Among them, most have been previously reported to be crucial genes for CRC tumorigenesis. For example, the HnRNPR-CCNB1/CENPF axis may contribute to gastric cancer proliferation and metastasis (Chen E. B. et al., 2019), CNN1 may be a potential prognostic marker of bladder cancer according to a bioinformatics analysis (Liu Y. et al., 2019), and the combined detection of CEA with FABP4 and FABP6 may improve the diagnostic efficacy of CRC (Zhang et al., 2019). These previous findings indicated that these nine hub genes have important roles in the occurrence and development of CRC, which suggested that they could also be potential prognostic markers.

Screening Related miRNAs Based on the Nine Hub Genes

Next, we used the nine hub genes to further survey miRNAs and lncRNAs that interacted with them to construct the ceRNA network (**Figure 5A**). miRNAs related to the hub genes were collected according to validated miRNA:mRNA interactions. These candidate miRNAs were further investigated in survival analyses to understand their potential values in cancer prognosis. Through this, we obtained 13 miRNAs with potential value as prognostic markers, and these miRNAs could interact with five hub genes (**Supplementary Table S6** and **Figure 5B**). All 13 miRNAs negatively regulated the hub genes by controlling the enrichment levels of these genes. Indeed, many studies have shown that miRNAs have crucial biological functions in CRC. For example, miR-92a and miR-144* may be potential biomarkers of non-invasive colorectal cancer (Choi et al., 2019), and miR-103a may be a new regulator of Wnt signaling as an onco-miRNA (Fasihi et al., 2018). In the miRNA-mRNA interaction network, we found that MFAP5 had the most interactions with miRNAs (it could be regulated by six miRNAs), while KPNA2 and CNN1 only interacted with one specific miRNA, miR-93-5p (**Figure 5B**). These interactions represented regulatory relationships in the coding-non-coding RNA network, and multiple interactions for specific genes suggested a complex regulatory network.
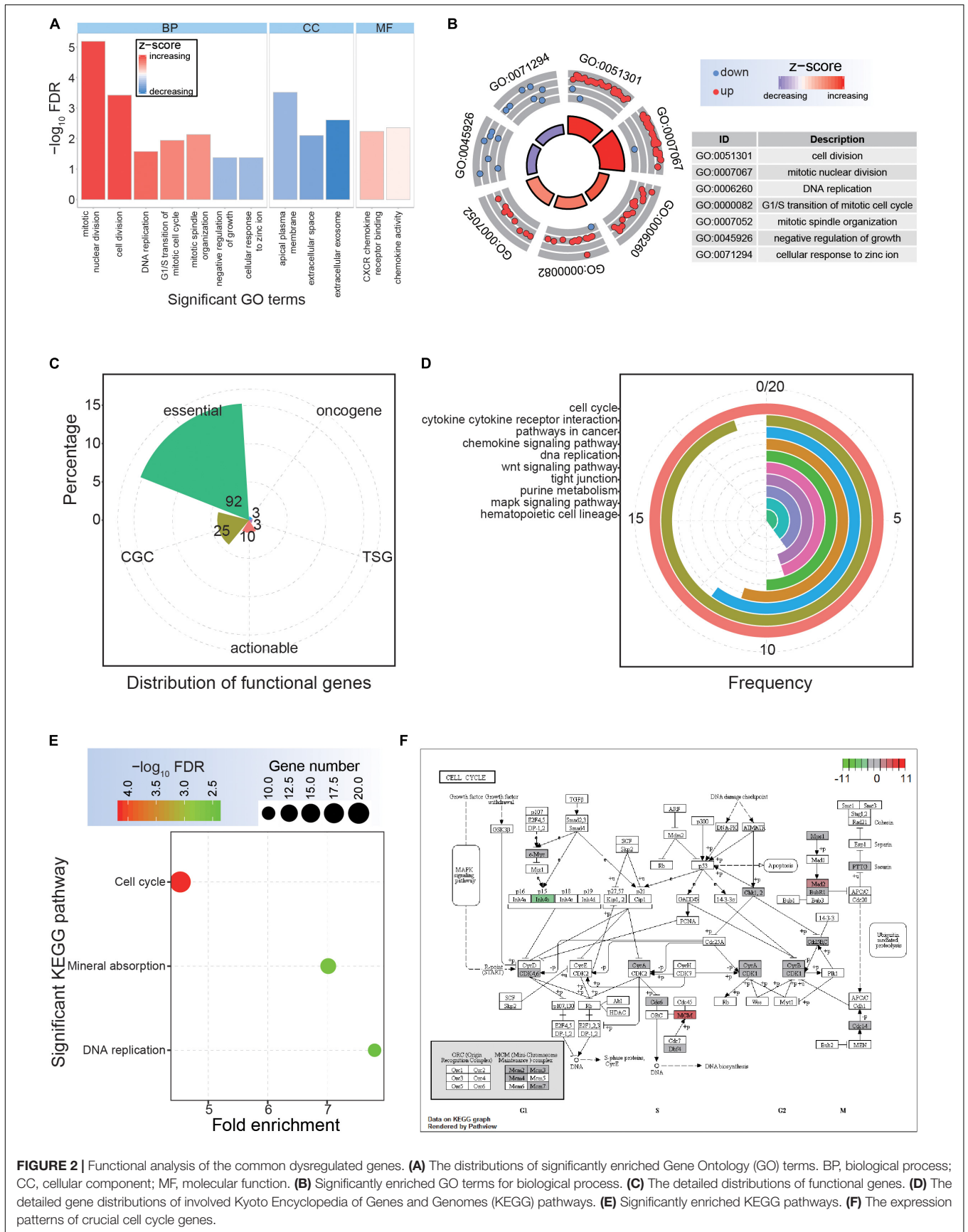
FIGURE 2 | Functional analysis of the common dysregulated genes. **(A)** The distributions of significantly enriched Gene Ontology (GO) terms. BP, biological process; CC, cellular component; MF, molecular function. **(B)** Significantly enriched GO terms for biological process. **(C)** The detailed distributions of functional genes. **(D)** The detailed gene distributions of involved Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. **(E)** Significantly enriched KEGG pathways. **(F)** The expression patterns of crucial cell cycle genes.

FIGURE 3 | The protein-protein interaction (PPI) networks constructed based on up- and down-regulated genes. **(A)** The PPI interaction network based on up-regulated genes, with two specific modules highlighted. The numbers of interacting genes are also presented. **(B)** The PPI interaction network based on down-regulated genes. The numbers of interacting genes are also presented.

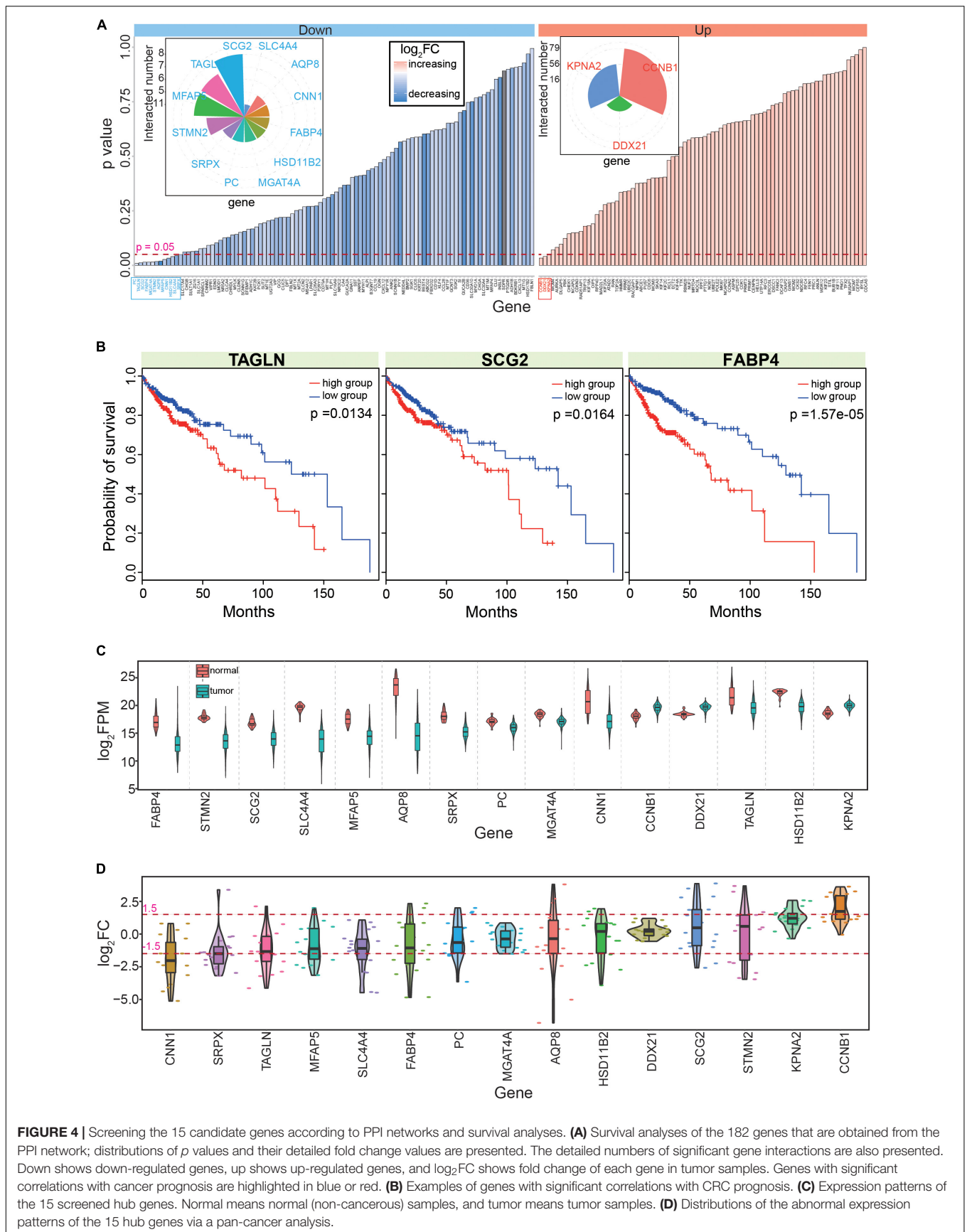**FIGURE 4 |** Screening the 15 candidate genes according to PPI networks and survival analyses. **(A)** Survival analyses of the 182 genes that are obtained from the PPI network; distributions of p values and their detailed fold change values are presented. The detailed numbers of significant gene interactions are also presented. Down shows down-regulated genes, up shows up-regulated genes, and log_2FC shows fold change of each gene in tumor samples. Genes with significant correlations with cancer prognosis are highlighted in blue or red. **(B)** Examples of genes with significant correlations with CRC prognosis. **(C)** Expression patterns of the 15 screened hub genes. Normal means normal (non-cancerous) samples, and tumor means tumor samples. **(D)** Distributions of the abnormal expression patterns of the 15 hub genes via a pan-cancer analysis.
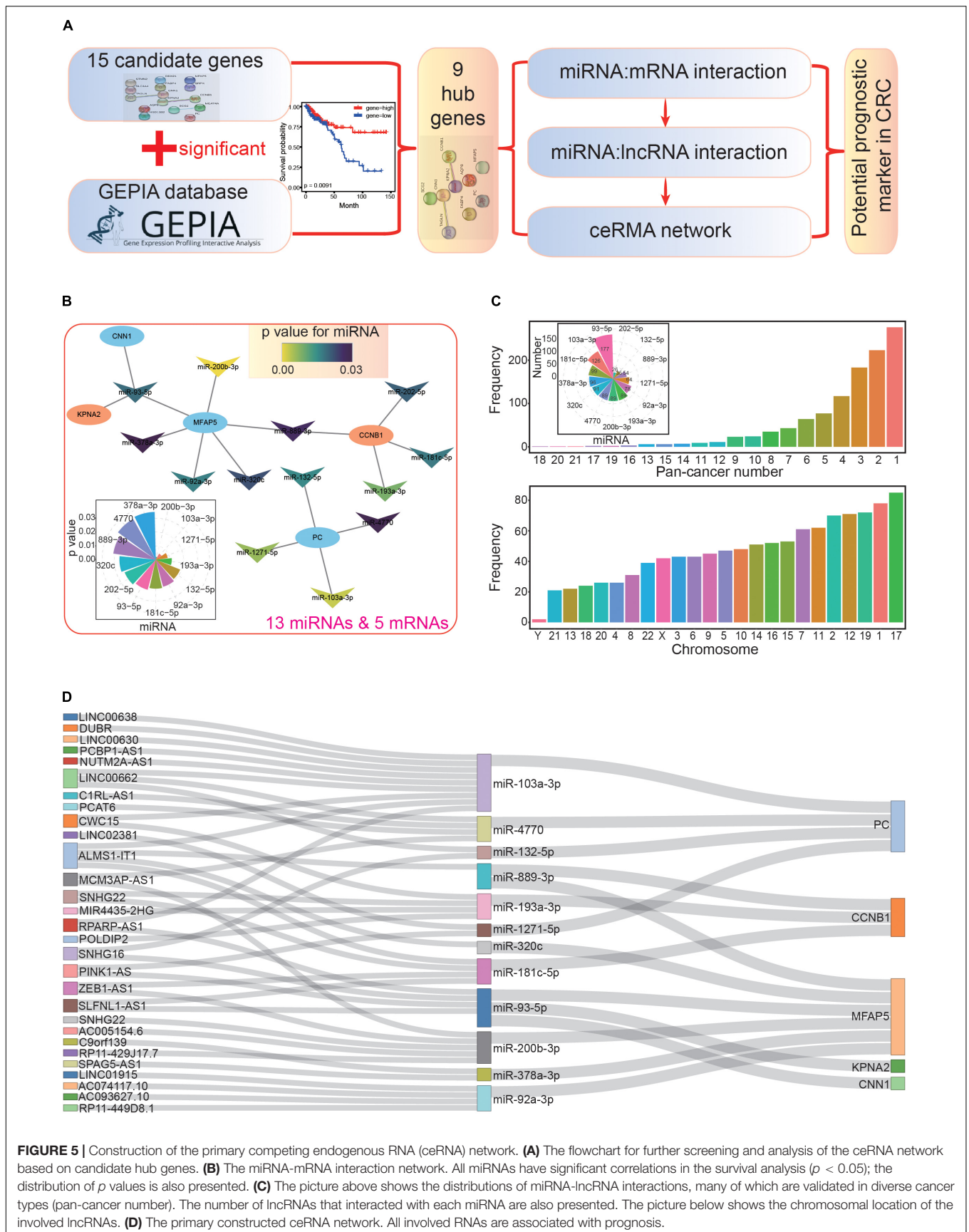
FIGURE 5 | Construction of the primary competing endogenous RNA (ceRNA) network. (A) The flowchart for further screening and analysis of the ceRNA network based on candidate hub genes. (B) The miRNA-mRNA interaction network. All miRNAs have significant correlations in the survival analysis (*p* < 0.05); the distribution of *p* values is also presented. (C) The picture above shows the distributions of miRNA-lncRNA interactions, many of which are validated in diverse cancer types (pan-cancer number). The number of lncRNAs that interacted with each miRNA are also presented. The picture below shows the chromosomal location of the involved lncRNAs. (D) The primary constructed ceRNA network. All involved RNAs are associated with prognosis.

Screening Associated lncRNAs Based on the 13 Candidate miRNAs

Based on these 13 miRNAs (herein, miRNAs were treated as intermediaries to connect mRNAs and lncRNAs), we further obtained a series of lncRNAs with possible interactions. Most of these interactions were validated in multiple cancer types (above picture in **Figure 5C**), indicating that these miRNA-lncRNA interactions existed in diverse tissues. Among the miRNAs, miR-93-5p had the most interacting lncRNAs (177), followed by miR-103a-3p (126), and miR-181c-5p (99). Furthermore, we also analyzed the distributions of these lncRNAs, and chromosome 17 contained the most, followed by chromosomes 1 and 19 (below picture in **Figure 5C**). To understand the potential prognostic values of these lncRNAs, we performed survival analyses, after which, we collected 29 lncRNAs that interacted with 12 miRNAs, all of which had potential prognostic value.

Construction of the ceRNA Network and Further Analysis at the isomiR Level

According to the screened hub genes, their interacting miRNAs, and associated lncRNAs, a ceRNA network was constructed based on their biological interactions. Although each step might involve a large number of candidate mRNAs, miRNA, or lncRNAs, only several of them were used to construct the ceRNA network after a step-by-step screening process, especially for simultaneously screening interactions and significant correlations with cancer prognosis. Ultimately, five hub genes (PC, CCNB1, MFAP5, KPNA2, and CNN1), 12 related miRNAs, and 29 associated lncRNAs were used to construct the ceRNA network (**Figure 5D**).

Although only validated miRNA:mRNA interactions were analyzed, it was necessary to determine whether there were negative correlations between them, as this would indicate that these interactions exist in CRC. According to StarBase (Li et al., 2014), there were four pairs of miRNA-mRNA interactions with significant negative correlations, including miR-93-5p:CNN1, miR-378-3p:MFAP5, miR-93-5p:MFAP5, and miR-200b-3p:MFAP5 (**Figure 6A**). These pairs contained only three miRNAs and two mRNAs, among which MFAP5 was the gene with the most (three) miRNA interactions. The expression patterns of these three miRNAs showed that miR-200b-3p was significantly up-regulated, while miR-378a-3p was significantly down-regulated (**Figure 6B**). This highlighted that examining expression patterns was a crucial step to further screen correlated RNAs based on ceRNA network.

Because of the multiple isomiRs within a single miRNA locus, we also analyzed the detailed expression patterns at the isomiR level for to each miRNA locus. Interestingly, we found there were different expression patterns for the top three dominantly expressed isomiRs of these miRNA loci (**Figure 6C**). For miR-93-5p, the dominant isomiR had a distinctly increased expression compared with the others, while the other miRNA loci showed closer expression patterns of the dominant isomiRs, especially the miR-200b-3p locus (**Figure 6C**). The non-random

expression distributions suggested that multiple isomiRs might be strictly regulated, which may contribute to maintaining miRNA:mRNA interactions that negatively regulate mRNA expression. Indeed, due to diverse sequences and enrichment levels, these isomiRs also had various expression patterns in cancer (**Figure 6D**). Some isomiRs were identified as abnormally expressed sequences, while others were normally expressed in tumor tissues (despite these diverse isomiRs were generated from a specific miRNA locus). For example, in the miR-378-3p locus, one specific isomiR was significantly dysregulated, while others did not show altered expression patterns (**Figure 6D**).

Screening Potential Prognostic Markers and Further in-Depth Analysis

We next filtered candidate RNAs according to miRNA:mRNA interactions and co-expression analysis. Thus, a further ceRNA network was constructed that contained two mRNAs (MFAP5 and CNN1), three miRNAs, and 11 lncRNAs (**Supplementary Table S7** and **Figure 6E**). In this interaction network of three RNA classes, because all of the included molecules were significantly correlated with cancer prognosis, their expression patterns were important factors to screen for potential prognostic markers. Thus, the 11 lncRNAs were further analyzed for their expression patterns in CRC. Finally, we found that AC005154.6 had a consistent expression pattern with MFAP5 (significantly down-regulated in CRC), and both had opposing expression patterns with the mediating miRNA, miR-200b-3p (**Figure 6E**).

We finally obtained MFAP5, miR-200b-3p, and AC005154.6 as potential prognostic markers in CRC. All three had potential prognostic value in CRC, and MFAP5 and AC005154.6 had consistent trends (**Figure 7A**). To further understand these three RNAs at their different molecular levels, they were queried for their detailed expression patterns in other cancer types. Based on paired samples, MFAP5 showed consistent expression patterns across different tissues, and most showed significant down-regulation (**Figure 7B**). Similar results were found in all samples (**Figure 7C**), and the relative expression patterns via pan-cancer analysis showed that MFAP5 was an important gene in tumorigenesis. Indeed, MFAP5 (microfibril-associated protein 5) may facilitate the distinction between pseudo-invasive and true-invasive lesions among colonic adenomatous polyps (Zhao et al., 2019). MFAP5 blockade can inhibit fibrosis and enhance chemosensitivity in ovarian and pancreatic cancer (Yeung et al., 2019), and can promote basal-like breast cancer progression by activating epithelial-to-mesenchymal transition (Wu et al., 2019).

As a potentially crucial miRNA that interacted with MFAP5, miR-200b-3p also showed relatively consistent expression across diverse cancer types (**Figure 7D**). We analyzed the three dominantly expressed isomiRs of the miR-200b-3p locus to track the expression patterns of these ncRNAs. The three isomiRs always showed consistent expression in specific tissues despite having heterogeneous sequences and expression, indicating potential collaborative relationships among these
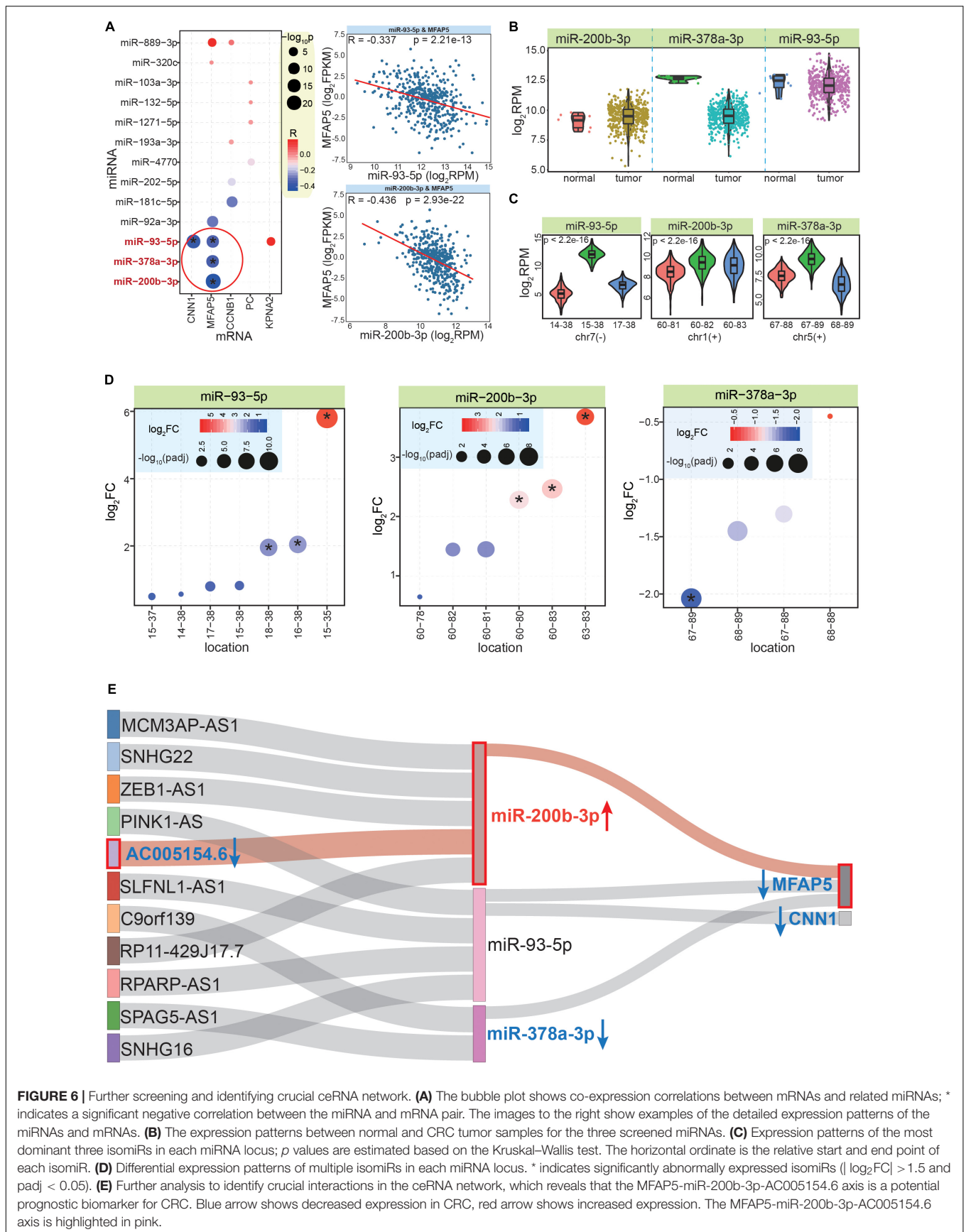
FIGURE 6 | Further screening and identifying crucial ceRNA network. **(A)** The bubble plot shows co-expression correlations between mRNAs and related miRNAs; * indicates a significant negative correlation between the miRNA and mRNA pair. The images to the right show examples of the detailed expression patterns of the miRNAs and mRNAs. **(B)** The expression patterns between normal and CRC tumor samples for the three screened miRNAs. **(C)** Expression patterns of the most dominant three isomiRs in each miRNA locus; p values are estimated based on the Kruskal–Wallis test. The horizontal ordinate is the relative start and end point of each isomiR. **(D)** Differential expression patterns of multiple isomiRs in each miRNA locus. * indicates significantly abnormally expressed isomiRs ($|\log_2 FC| > 1.5$ and padj < 0.05). **(E)** Further analysis to identify crucial interactions in the ceRNA network, which reveals that the MFAP5-miR-200b-3p-AC005154.6 axis is a potential prognostic biomarker for CRC. Blue arrow shows decreased expression in CRC, red arrow shows increased expression. The MFAP5-miR-200b-3p-AC005154.6 axis is highlighted in pink.
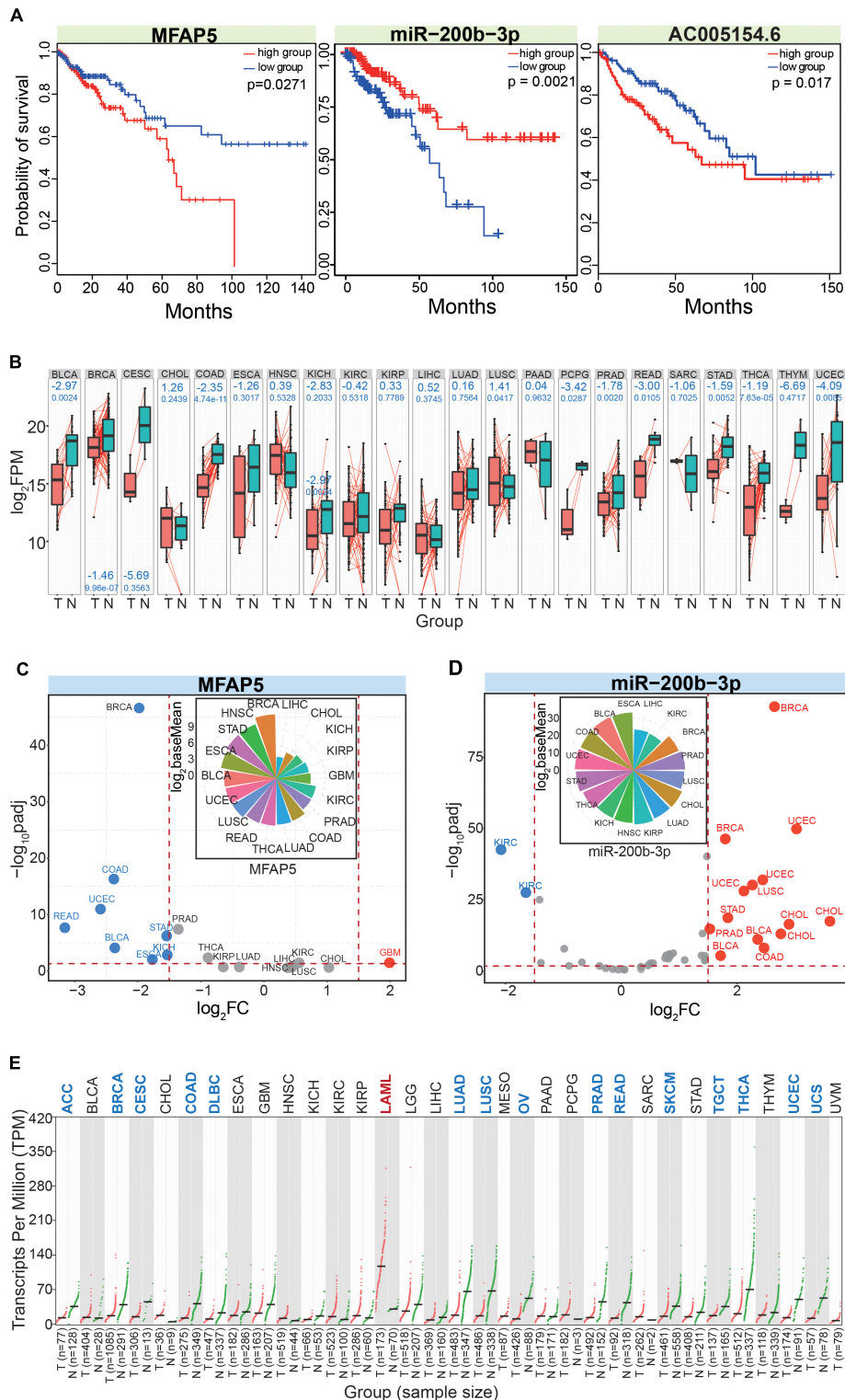
FIGURE 7 | In-depth analysis of the screened RNAs. (A) Survival curves for the involved RNAs show their potential values as prognostic markers. (B) Expression patterns of MFAP5 based on paired analysis across different cancer types. The log_2FC (above blue number) and p (below blue number) values are presented in relevant images. T, tumor samples; N, normal samples. (C) The scatter plot shows expression patterns of MFAP5 based on the DeSeq algorithm. The picture above shows distributions of baseMean values. (D) The scatter plot shows expression patterns of miR-200b-3p based on the DeSeq algorithm. The picture above shows distributions of baseMean values. Here, the three dominant isomiRs in the miR-200b-3p locus are simultaneously analyzed. (E) Expression pattern of AC005154.6 based on Gene Expression Profiling Interactive Analysis (GEPIA).

multiple isomiRs in the coding-non-coding RNA regulatory network. miR-200 has been validated as an important regulatory molecule. Specifically, the miR-200b-3p/p38IP pair regulates monocyte/macrophage differentiation (Yu et al., 2016), miR-200b-3p in plasma may be a potential diagnostic biomarker of oral squamous cell carcinoma (Sun et al., 2018), and quercetin-induced miR-200b-3p can regulate the mode of self-renewal divisions in pancreatic cancer (Nwaeburu et al., 2017). Thus, miR-200 has an important role in the pathological processes of multiple cancers, and diverse isomiRs, especially the dominantly expressed isomiRs, largely contribute to synergistically regulating these pro-tumor processes. For the lncRNA AC005154.6, most cancer types showed down-regulated expression (**Figure 7E**), and these consistent expression patterns suggested it played a role in tumorigenesis. AC005154.6 has been reported to be highly expressed in the pituitary (Levran et al., 2018), but details of its biological functions are unknown.

DISCUSSION

Identifying new prognostic biomarkers is essential for CRC, as this also contributes to exploring the mechanisms of metastasis and surveying candidate gene targets for therapy. Herein, according to the ceRNA hypothesis, we proposed an approach to construct a ceRNA network in CRC based on multiple RNA datasets. First, the common dysregulated mRNA profiles were filtered across several sequencing datasets, and the significant DEGs were further queried for their potential functional implications. Second, potential hub genes were screened from PPI networks, and candidate hub genes were screened based on their prognostic value in CRC. Primarily screened hub genes were further validated based on their expression patterns across diverse cancer types, which revealed their expression trends and specificities in different tissues. Third, candidate hub genes were extended to interacting miRNAs because these small ncRNAs are important negative regulators in coding-non-coding RNA networks. The miRNA:mRNA interactions were validated based on their biological and expression relationships. Then, candidate miRNAs were screened by survival analyses to estimate their prognostic value. Finally, based on the screened miRNAs from the hub genes, relevant lncRNAs were collected and queried for their potential as prognostic markers. Thus, we screened hub genes, and relevant miRNAs and lncRNAs that were associated with CRC prognosis, as any of them may be a potential clinical biomarker. These obtained RNAs also have potential biological relationships and close expression correlations, so they were used to construct a ceRNA network to further screen candidate prognostic CRC biomarkers. Finally, the dysregulated expression patterns were simultaneously analyzed to predict possible regulatory relationships among the different identified RNAs. Furthermore, the expression patterns of miRNAs were dissected at the isomiR level. Thus, a mRNA-miRNA-lncRNA axis could be identified as a potential prognostic marker, and the involved RNAs were analyzed in-depth across diverse cancer types to identify trends in their expression.

In this study, miRNAs were used to link mRNAs and lncRNAs, which is an important step when constructing a ceRNA network. Given tumor heterogeneity and the multitude of variables that influence clinical progress, the combination of multiple RNAs provides a more comprehensive prognostic analysis. Indeed, many previous studies have found great prospects for the regulation of CRCs by lncRNAs, miRNAs and mRNAs (Marisa et al., 2013; Pizzini et al., 2013; Zhou et al., 2016). Aberrant lncRNA expression is associated with tumorigenesis, tumor progression, and metastasis (Shi et al., 2013). Although only a few lncRNAs have been investigated in CRC, existing results have demonstrated that lncRNAs may be ideal prognostic biomarkers for this malignancy (Smolle et al., 2014). In the integrative analysis of the different RNA classes, we also analyzed multiple isomiRs within the different miRNA loci. Diverse isomiRs expand miRNA:mRNA interactions, and the widespread interactions between small ncRNAs and mRNAs contribute to flexible regulation and regulatory effectiveness. Multiple isomiRs provide a synergistic regulatory pattern among diverse small ncRNAs, although these isomiRs usually show expression and sequence heterogeneities. Similar to homologous and/or clustered classical miRNAs, the phenomenon of isomiRs enriches and complicates studying small regulatory RNAs.

Based on first screening hub genes and then identifying relevant miRNAs followed by lncRNAs, all of the candidate RNAs are potential prognostic markers in CRC at the individual molecular level. Based on the potential biological and expression relationships, especially for expression patterns of isomiRs within the miRNA loci, we further screened the MFAP5-miR-200b-3p-AC005154.6 axis, which may be a potential prognostic biomarker for CRC. MFAP5 promotes tumor progression and bone metastasis in several cancers (Leung et al., 2014; Wu Z. et al., 2018), and miR-200b-3p negatively regulates MFAP5 expression; finally, AC005154.6 interacts with miRNA-200b-3p. Thus, the interactions among these RNAs may further contribute to relevant biological pathways and CRC pathophysiology. Although previous studies have shown that these RNAs have important roles in multiple biological processes, especially for MFAP5 and miR-200b-3p in tumorigenesis, more studies are needed that are focused on the mechanistic interactions between MFAP5, miR-200b-3p and AC005154.6, especially regarding their functional relationships in the occurrence and progression of CRC.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the manuscript.

AUTHOR CONTRIBUTIONS

LG and TL designed this study. LG, GY, YK, SL, RD, LS, WJ, BQ, and ZY participated in data analysis. LG and TL wrote the manuscript. All authors read and accepted the final version. TL was responsible for coordinating and supervising the entire project.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00418/full#supplementary-material

REFERENCES

Allegra, C. J., Paik, S., Colangelo, L. H., Parr, A. L., Kirsch, I., Kim, G., et al. (2003). Prognostic value of thymidylate synthase, Ki-67, and p53 in patients with Dukes' B and C colon cancer: a National Cancer Institute-national surgical adjuvant breast and bowel project collaborative study. *J. Clin. Oncol.* 21, 241–250. doi: 10.1200/JCO.2003.05.044

Altobelli, E., Lattanzi, A., Paduano, R., Varassi, G., and di Orio, F. (2014). Colorectal cancer prevention in Europe: burden of disease and status of screening programs. *Prev. Med.* 62, 132–141. doi: 10.1016/j.ypmed.2014.02.010

Blomen, V. A., Majek, P., Jae, L. T., Bigenzahn, J. W., Nieuwenhuis, J., Staring, J., et al. (2015). Gene essentiality and synthetic lethality in haploid human cells. *Science* 350, 1092–1096. doi: 10.1126/science.aac7557

Callahan, K. E., Ponce, C. P., Cross, C. L., Sy, F. S., and Pinheiro, P. S. (2019). Low colorectal cancer survival in the Mountain West state of Nevada: a population-based analysis. *PLoS One* 14:e0221337. doi: 10.1371/journal.pone.0221337PONE-D-18-35819

Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., et al. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147, 358–369. doi: 10.1016/j.cell.2011.09.028

Chen, E. B., Qin, X., Peng, K., Li, Q., Tang, C., Wei, Y. C., et al. (2019). HnRNPR-CCNB1/CENPF axis contributes to gastric cancer proliferation and metastasis. *Aging* 11, 7473–7491. doi: 10.18632/aging.102254102254

Chen, Y., Mao, Z. D., Shi, Y. J., Qian, Y., Liu, Z. G., Yin, X. W., et al. (2019). Comprehensive analysis of miRNA-mRNA-lncRNA networks in severe asthma. *Epigenomics* 11, 115–131. doi: 10.2217/epi-2018-2132

Choi, H. H., Cho, Y. S., Choi, J. H., Kim, H. K., Kim, S. S., and Chae, H. S. (2019). Stool-Based miR-92a and miR-144* as noninvasive biomarkers for colorectal cancer screening. *Oncology* 97, 173–179. doi: 10.1159/000500639000500639

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44:e71. doi: 10.1093/nar/gkv1507

Desvignes, T., Loher, P., Eilbeck, K., Ma, J., Urgese, G., Fromm, B., et al. (2020). Unification of miRNA and isomiR research: the mirGFF3 format and the mirtop API. *Bioinformatics* 36, 698–703. doi: 10.1093/bioinformatics/btz675

Fan, Y., and Xia, J. (2018). miRNet-functional analysis and visual exploration of miRNA-target interactions in a network context. *Methods Mol. Biol.* 1819, 215–233. doi: 10.1007/978-1-4939-8618-7_10

Farina-Sarasqueta, A., van Lijnschoten, G., Moerland, E., Creemers, G. J., Lemmens, V. E., Rutten, H. J., et al. (2010). The BRAF V600E mutation is an independent prognostic factor for survival in stage II and stage III colon cancer patients. *Ann. Oncol.* 21, 2396–2402. doi: 10.1093/annonc/mdq258mdq258

Fasihi, A., Soltani, B. M., Atashi, A., and Nasiri, S. (2018). Introduction of hsa-miR-103a and hsa-miR-1827 and hsa-miR-137 as new regulators of Wnt signaling pathway and their relation to colorectal carcinoma. *J. Cell. Biochem.* 119, 5104–5117. doi: 10.1002/jcb.26357

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136, E359–E386. doi: 10.1002/ijc.29210

Futreal, P. A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., et al. (2004). A census of human cancer genes. *Nat. Rev. Cancer* 4, 177–183. doi: 10.1038/nrc1299

Guo, L., and Liang, T. (2018). microRNAs and their variants in an RNA world: implications for complex interactions and diverse roles in an RNA regulatory network. *Brief. Bioinform.* 19, 245–253. doi: 10.1093/bib/bbw124

Guo, S., Fesler, A., Wang, H., and Ju, J. (2018). microRNA based prognostic biomarkers in pancreatic Cancer. *Biomark. Res.* 6:18. doi: 10.1186/s40364-018-0131-1131

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K. R., MacLeod, G., et al. (2015). High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* 163, 1515–1526. doi: 10.1016/j.cell.2015.11.015

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Khamas, A., Ishikawa, T., Shimokawa, K., Mogushi, K., Iida1, S., Ishiguro, M., et al. (2012). Screening for epigenetically masked genes in colorectal cancer Using 5-Aza-2′-deoxycytidine, microarray and gene expression profile. *Cancer Genom. Proteom.* 9, 67–75.

Lee, J. T. (2012). Epigenetic regulation by long noncoding RNAs. *Science* 338, 1435–1439. doi: 10.1126/science.1231776338/6113/1435

Leung, C. S., Yeung, T. L., Yip, K. P., Pradeep, S., Balasubramanian, L., Liu, J., et al. (2014). Calcium-dependent FAK/CREB/TNNC1 signalling mediates the effect of stromal MFAP5 on ovarian cancer metastatic potential. *Nat. Commun.* 5:5092. doi: 10.1038/ncomms6092ncomms6092

Levran, O., Correa da Rosa, J., Randesi, M., Rotrosen, J., Adelson, M., and Kreek, M. J. (2018). A non-coding CRHR2 SNP rs255105, a cis-eQTL for a downstream lincRNA AC005154.6, is associated with heroin addiction. *PLoS One* 13:e0199951. doi: 10.1371/journal.pone.0199951PONE-D-18-15461

Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–D97. doi: 10.1093/nar/gkt1248gkt1248

Liu, Y., Wu, X., Wang, G., Hu, S., Zhang, Y., and Zhao, S. (2019). CALD1, CNN1, and TAGLN identified as potential prognostic molecular markers of bladder cancer by bioinformatics analysis. *Medicine* 98:e13847. doi: 10.1097/MD.00000000000001384700005792-201901110-201901117

Liu, Z., Mi, M., Li, X., Zheng, X., Wu, G., and Zhang, L. (2019). lncRNA OSTN-AS1 may represent a novel immune-related prognostic marker for triple-negative breast cancer based on integrated analysis of a ceRNA network. *Front. Genet.* 10:850. doi: 10.3389/fgene.2019.00850

Long, A. G., Lundsmith, E. T., and Hamilton, K. E. (2017). Inflammation and colorectal cancer. *Curr. Colorectal. Cancer Rep.* 13, 341–351. doi: 10.1007/s11888-017-0373-376

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/S13059-014-0550-558

Lu, M. H., Tang, B., Zeng, S., Hu, C. J., Xie, R., Wu, Y. Y., et al. (2016). Long noncoding RNA BC032469, a novel competing endogenous RNA, upregulates hTERT expression by sponging miR-1207-5p and promotes proliferation in gastric cancer. *Oncogene* 35, 3524–3534. doi: 10.1038/onc.2015.413onc2015413

Luo, W., and Brouwer, C. (2013). Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* 29, 1830–1831. doi: 10.1093/bioinformatics/btt285btt285

Luo, W., Pant, G., Bhavnasi, Y. K., Blanchard, S. G. Jr., and Brouwer, C. (2017). Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Res.* 45, W501–W508. doi: 10.1093/nar/gkx3723804420

Marisa, L., de Reynies, A., Duval, A., Selves, J., Gaub, M. P., Vescovo, L., et al. (2013). Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* 10:e1001453. doi: 10.1371/journal.pmed.1001453PMEDICINE-D-12-03295

Morin, R. D., O'Connor, M. D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A. L., et al. (2008). Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* 18, 610–621. doi: 10.1101/gr.7179508

Morlando, M., and Fatica, A. (2018). Alteration of epigenetic regulation by long noncoding RNAs in cancer. *Int. J. Mol. Sci.* 19:E570. doi: 10.3390/ijms19020570

Neilsen, C. T., Goodall, G. J., and Bracken, C. P. (2012). IsomiRs - the overlooked repertoire in the dynamic microRNAome. *Trends Genet.* 28, 544–549. doi: 10.1016/j.tig.2012.07.005

Nwaeburu, C. C., Abukiwan, A., Zhao, Z., and Herr, I. (2017). Quercetin-induced miR-200b-3p regulates the mode of self-renewing divisions in pancreatic cancer. *Mol. Cancer* 16:23. doi: 10.1186/s12943-017-0589-588

Okugawa, Y., Toiyama, Y., Toden, S., Mitoma, H., Nagasaka, T., Tanaka, K., et al. (2017). Clinical significance of SNORA42 as an oncogene and a prognostic biomarker in colorectal cancer. *Gut* 66, 107–117. doi: 10.1136/gutjnl-2015-309359

Pawa, N., Arulampalam, T., and Norton, J. D. (2011). Screening for colorectal cancer: established and emerging modalities. *Nat. Rev. Gastroenterol. Hepatol.* 8, 711–722. doi: 10.1038/nrgastro.2011.205

Pizzini, S., Bisognin, A., Mandruzzato, S., Biasiolo, M., Facciolli, A., Perilli, L., et al. (2013). Impact of microRNAs on regulatory networks and pathways in human colorectal carcinogenesis and development of metastasis. *BMC Genomics* 14:589. doi: 10.1186/1471-2164-14-589

Qian, Y., Mao, Z. D., Shi, Y. J., Liu, Z. G., Cao, Q., and Zhang, Q. (2018). Comprehensive analysis of miRNA-mRNA-lncRNA networks in non-smoking and smoking patients with chronic obstructive pulmonary disease. *Cell Physiol. Biochem.* 50, 1140–1153. doi: 10.1159/000494541

Rahman, R., Schmaltz, C., Jackson, C. S., Simoes, E. J., Jackson-Thompson, J., and Ibdah, J. A. (2015). Increased risk for colorectal cancer under age 50 in racial and ethnic minorities living in the United States. *Cancer Med.* 4, 1863–1870. doi: 10.1002/cam4.560

Robles, A. I., Arai, E., Mathe, E. A., Okayama, H., Schetter, A. J., Brown, D., et al. (2015). An integrated prognostic classifier for stage I lung adenocarcinoma based on mRNA, microRNA, and DNA methylation biomarkers. *J. Thorac. Oncol.* 10, 1037–1048. doi: 10.1097/JTO.0000000000000560S1556-0864(15)33491-33492

Roth, A. D., Tejpar, S., Delorenzi, M., Yan, P., Fiocca, R., Klingbiel, D., et al. (2010). Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. *J. Clin. Oncol.* 28, 466–474. doi: 10.1200/JCO.2009.23.3452

Sabates-Bellver, J., Van der Flier, L. G., de Palo, M., Cattaneo, E., Maake, C., Rehrauer, H., et al. (2007). Transcriptome profile of human colorectal adenomas. *Mol. Cancer Res.* 5, 1263–1275. doi: 10.1158/1541-7786.MCR-07-0267

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Shi, X., Sun, M., Liu, H., Yao, Y., and Song, Y. (2013). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* 339, 159–166. doi: 10.1016/j.canlet.2013.06.013S0304-3835(13)00463-461

Smolle, M., Uranitsch, S., Gerger, A., Pichler, M., and Haybaeck, J. (2014). Current status of long non-coding RNAs in human cancer with specific focus on colorectal cancer. *Int. J. Mol. Sci.* 15, 13993–14013. doi: 10.3390/ijms150813993

Song, X., Zhang, C., Liu, Z., Liu, Q., He, K., and Yu, Z. (2019). Characterization of ceRNA network to reveal potential prognostic biomarkers in triple-negative breast cancer. *PeerJ* 7:e7522. doi: 10.7717/peerj.7522

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102

Sun, G., Cao, Y., Wang, P., Song, H., Bie, T., Li, M., et al. (2018). miR-200b-3p in plasma is a potential diagnostic biomarker in oral squamous cell carcinoma. *Biomarkers* 23, 137–141. doi: 10.1080/1354750X.2017.1289241

Sun, J., and Kato, I. (2016). Gut microbiota, inflammation and colorectal cancer. *Genes Dis.* 3, 130–143. doi: 10.1016/j.gendis.2016.03.004

Sun, L., Jiang, R., Li, J., Wang, B., Ma, C., Lv, Y., et al. (2017). MicoRNA-425-5p is a potential prognostic biomarker for cervical cancer. *Ann. Clin. Biochem.* 54, 127–133. doi: 10.1177/0004563216649377

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky11315198476

Tan, G. C., Chan, E., Molnar, A., Sarkar, R., Alexieva, D., Isa, I. M., et al. (2014). 5 ' isomiR variation is of functional and evolutionary importance. *Nucleic Acids Res.* 42, 9424–9435. doi: 10.1093/nar/gku656

Tan, J. Y., Sirey, T., Honti, F., Graham, B., Piovesan, A., Merkenschlager, M., et al. (2015). Extensive microRNA-mediated crosstalk between lncRNAs and mRNAs in mouse embryonic stem cells. *Genome Res.* 25, 655–666. doi: 10.1101/gr.181974.114

Tang, Z., Li, C., Kang, B., Gao, G., and Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45, W98–W102. doi: 10.1093/nar/gkx2473605636

Tong, A. W., and Nemunaitis, J. (2008). Modulation of miRNA activity in human cancer: a new paradigm for cancer gene therapy? *Cancer Gene Ther.* 15, 341–355. doi: 10.1038/cgt.2008.8cgt20088

Tsukamoto, S., Ishikawa, T., Iida, S., Ishiguro, M., Mogushi, K., Mizushima, H., et al. (2011). Clinical significance of osteoprotegerin expression in human colorectal cancer. *Clin. Cancer Res.* 17, 2444–2450. doi: 10.1158/1078-0432.CCR-10-2884

Valihrach, L., Androvic, P., and Kubista, M. (2019). Circulating miRNA analysis for cancer diagnostics and therapy. *Mol. Aspects Med.* [Epub ahead of print]. doi: 10.1016/j.mam.2019.10.002

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. Jr., and Kinzler, K. W. (2013). Cancer genome landscapes. *Science* 339, 1546–1558. doi: 10.1126/science.1235122

Walter, W., Fátima, S., and Mercedes R. (2015). GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics* 31, 2912–2914. doi: 10.1093/bioinformatics/btv300

Wang, T., Birsoy, K., Hughes, N. W., Krupczak, K. M., Post, Y., Wei, J. J., et al. (2015). Identification and characterization of essential genes in the human genome. *Science* 350, 1096–1101. doi: 10.1126/science.aac7041

Wang, W., Li, Q., Yang, T., Bai, G., Li, D., and Sun, H. (2012). Expression of AQP5 and AQP8 in human colorectal carcinoma and their clinical significance. *World J. Surg. Oncol.* 10:242. doi: 10.1186/1477-7819-10-242

Wang, Y., Huang, T., and Sun, X. (2019). Identification of a potential prognostic lncRNA-miRNA-mRNA signature in endometrial cancer based on the competing endogenous RNA network. *J. Cell. Biochem.* 120, 18845–18853. doi: 10.1002/jcb.29200

Worthley, D. L., Whitehall, V. L., Spring, K. J., and Leggett, B. A. (2007). Colorectal carcinogenesis: road maps to cancer. *World J. Gastroenterol.* 13, 3784–3791. doi: 10.3748/wjg.v13.i28.3784

Wu, Q., Yang, Z. F., Wang, K. J., Feng, X. Y., Lv, Z. J., Li, Y., et al. (2018). AQP8 inhibits colorectal cancer growth and metastasis by down-regulating PI3K/AKT signaling and PCDH7 expression. *Am. J. Cancer Res.* 8, 266–279.

Wu, Y., Wu, P., Zhang, Q., Chen, W., Liu, X., and Zheng, W. (2019). MFAP5 promotes basal-like breast cancer progression by activating the EMT program. *Cell. Biosci.* 9:24. doi: 10.1186/s13578-019-0284-0284

Wu, Z., Wang, T., Fang, M., Huang, W., Sun, Z., Xiao, J., et al. (2018). MFAP5 promotes tumor progression and bone metastasis by regulating ERK/MMP signaling pathways in breast cancer. *Biochem. Biophys. Res. Commun.* 498, 495–501. doi: 10.1016/j.bbrc.2018.03.007

Yao, Y., Zhang, T., Qi, L., Zhou, C., Wei, J., Feng, F., et al. (2019). Integrated analysis of co-expression and ceRNA network identifies five lncRNAs as prognostic markers for breast cancer. *J. Cell Mol. Med.* 23, 8410–8419. doi: 10.1111/jcmm.14721

Yeung, T. L., Leung, C. S., Yip, K. P., Sheng, J., Vien, L., Bover, L. C., et al. (2019). Anticancer immunotherapy by MFAP5 blockade inhibits fibrosis and enhances chemosensitivity in ovarian and pancreatic cancer. *Clin. Cancer Res.* 25, 6417–6428. doi: 10.1158/1078-0432.CCR-19-0187

Yu, X., Wang, Q. L., Li, Y. F., Wang, X. D., Xu, A., and Li, Y. (2016). A novel miR-200b-3p/p38IP pair regulates monocyte/macrophage differentiation. *Cell Discov.* 2:15043. doi: 10.1038/celldisc.2015.43

Zhang, G., Li, S., Lu, J., Ge, Y., Wang, Q., Ma, G., et al. (2018). LncRNA MT1JP functions as a ceRNA in regulating FBXW7 through competitively binding

to miR-92a-3p in gastric cancer. *Mol. Cancer* 17:87. doi: 10.1186/s12943-018-0829-826

Zhang, Y., Zhao, X., Deng, L., Li, X., Wang, G., Li, Y., et al. (2019). High expression of FABP4 and FABP6 in patients with colorectal cancer. *World J. Surg. Oncol.* 17:171. doi: 10.1186/s12957-019-1714-1715

Zhao, L., Westerhoff, M., Hornick, J. L., Krausz, T., Antic, T., Xiao, S. Y., et al. (2019). Loss of microfibril-associated protein 5 (MFAP5) expression in colon cancer stroma. *Virchows Arch.* 476, 383–390. doi: 10.1007/s00428-019-02649-y

Zhou, M., Diao, Z., Yue, X., Chen, Y., Zhao, H., Cheng, L., et al. (2016). Construction and analysis of dysregulated lncRNA-associated ceRNA network identified novel lncRNA biomarkers for early diagnosis of human pancreatic cancer. *Oncotarget* 7, 56383–56394. doi: 10.18632/oncotarget.10891