

# RNA Language in Colorectal Cancer Using an Integrative Bioinformatics Approach

**Sepideh Kadkhoda**

Tehran University of Medical Sciences

**Farzaneh Darbeheshti**

Tehran University of Medical Sciences <https://orcid.org/0000-0002-4957-9004>

**Nima Rezaei** (✉ [rezaei\\_nima@tums.ac.ir](mailto:rezaei_nima@tums.ac.ir))

<https://orcid.org/0000-0002-3836-1827>

**Ghasem Azizi-Tabesh**

Shaheed Beheshti University of Medical Sciences

**Faezeh Zolfaghari**

Tehran University of Medical Sciences

**Sadollah Tavakolibazaz**

Mazandaran University of Medical Sciences

**Reza Taslimi**

Tehran University of Medical Sciences

**Javad Tavakkoly-Bazzaz**

Tehran University of Medical Sciences

---

## Primary research

**Keywords:** circRNA, miRNA, ceRNA, colorectal cancer, microarray, bioinformatics

**Posted Date:** March 31st, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-19929/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# RNA Language in Colorectal Cancer Using an Integrative Bioinformatics Approach

Sepideh Kadkhoda<sup>1,†</sup>, Farzaneh Darbeheshti<sup>1,2,†</sup>, Nima Rezaei<sup>3,4,5\*</sup>, Ghasem Azizi-Tabesh<sup>6,7</sup>,  
Faezeh Zolfaghari<sup>1</sup>, Sadollah Tavakolibazaz<sup>8</sup>, Reza Taslimi<sup>9</sup>, Javad Tavakkoly-Bazzaz<sup>1\*</sup>

<sup>1</sup> Department of Medical Genetics, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

<sup>2</sup> Breast Cancer Association (BrCA), Universal Scientific Education and Research Network (USERN), Tehran, Iran

<sup>3</sup> Research Center for Immunodeficiencies, Children's Medical Center, Tehran University of Medical Sciences, Tehran, Iran

<sup>4</sup> Department of Immunology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

<sup>5</sup> Network of Immunity in Infection, Malignancy and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), Tehran, Iran

<sup>6</sup> Department of Medical Genetics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>7</sup> Genomic Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>8</sup> Department of Gastrointestinal Dis., Imam Reza Hospital, Mazandaran University of Medical Sciences, Amol, Iran

<sup>9</sup> Department of Gastroenterology, Imam Khomeini Hospital, Tehran University of Medical Sciences, Tehran, Iran

\*

Sepideh Kadkhoda and Farzaneh Darbeheshti contributed equally to this study as first authors.

\*

## Correspondence

Nima Rezaei M.D, Ph.D, Research Center for Immunodeficiencies, Children's Medical Center, Tehran University of Medical Sciences, Tehran, Iran. Email: rezaei\_nima@tums.ac.ir

Javad Tavakkoly-Bazzaz M.D, Ph.D, Department of Medical Genetics, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran. Email: tavakkolybazzazj@tums.ac.ir

## Abstract

**Background:** Identification of competing endogenous RNAs (ceRNAs), especially circRNAs, have become new hotspots in cancer researches. Although, their roles and underlying mechanisms in colorectal cancer (CRC) development remain mostly unknown. The aim of this study was to integrate both coding and non-coding available microarray data in development of CRC coupled with bioinformatics analyses to understand a more inclusive pathobiologic map regarding their molecular interactions and functions.

**Methods:** The microarray data were retrieved from the Gene Expression Omnibus (GEO) database and analyzed. Several bioinformatics tools and databases including CircInteractome, CSCD, miRTarBase, TargetScan, miRmap, GEPIA, STRING, Enrichr, DAVID, and MCODE were applied for further elucidation. Principal component analysis (PCA) has separately run for four datasets. The dysregulated circRNA-miRNA-mRNA network in CRC was constructed by Cytoscape. In addition, co-expression and protein-protein interaction (PPI) networks were established based on differentially expressed (DE) protein coding genes in CRC.

**Results:** PCA disclose colorectal tumor and normal tissues could be distinguished not only by mRNAs expression profile, but also by both circRNAs and miRNAs expression profiles. We identified 14 DE mRNAs (commonly between two datasets), 85 DE miRNAs and 36 DE circRNAs in CRC tissues compared with normal tissues. Taking their potential interactions into account, a circRNA-miRNA-mRNA network was constructed. Then, according to ceRNA hypothesis, the axes with expression in the desired direction were extracted. Our results disclosed some DE circRNAs with potential oncogenic (circ\_0014879) or tumor suppressive (circ\_0001666 and circ\_0000977) effects. Finally, PPI network suggests pivotal roles for DOCK2 and PTPRC dysregulation in progression of CRC, possibly by facilitating of tumor escape from immune surveillance.

**Conclusion:** Current study proposes a novel regulatory network consisting of DE circRNAs, miRNAs and mRNAs in CRC development that in turn highlights the roles of DE circRNAs at the upstream of oncotranscriptomic cascade in CRC development, suggesting their potentiality to be utilized as both prognostic and therapeutic biomarker.

**Key words:** circRNA, miRNA, ceRNA, colorectal cancer, microarray, bioinformatics

## **Background**

Colorectal cancer (CRC), regardless of gender, is classified as the third most prevalent and second most lethal malignancy in 2018 [1]. However, it could be well treated and managed with suitable and early diagnosis. Thus, a better understanding of molecular signatures, potential biomarkers and therapeutic targets in CRC is necessary to gradually improve the diagnosis and treatment efficacy.

Nowadays, high throughput technologies have helped discover and investigate both coding and non-coding dysregulated transcripts in tumor cells. The pivotal roles of non-coding RNAs in gene expression regulation and consequently cancer initiation and progression have been determined. Recently, circular RNAs (circRNAs) as a newfound category of non-coding RNAs have been highlighted in cancer researches. circRNAs due to high stability and tissue/stage specificity show potential to be diagnostic and prognostic biomarkers [2]. In 2011, ceRNA hypothesis was proposed by Salmena et al [3]. They demonstrated that RNA transcripts communicate with each other by microRNA response elements (MREs). According to this hypothesis, the circRNA function in microRNAs sponging and consequently dysregulation of protein-coding genes expression has been established in different malignancies as long as circRNA-miRNA-mRNA axis is considered [4, 5]. However, various circRNAs have been reported as participating in the CRC pathogenesis, their molecular interactions with other transcripts are still unknown [6]. These interactions could improve current knowledge about molecular mechanisms underlying CRC and potential targeted therapies. Integrating experimental data with bioinformatics could present a powerful method to disclose molecular interactions among different transcripts in tumor cells as well as their potential functions.

The aim of this study is integration of available microarray data concerning circRNA, miRNA and mRNA in CRC with bioinformatics analyses by consensus strategy to gain a more accurate comprehension of their molecular interactions and functions. We have constructed a dysregulated circRNA-miRNA-mRNA network in colorectal tumors as well as co-expression and protein-protein interaction (PPI) networks. Our findings propose novel regulatory network underlying CRC progression and uncover molecular interactions between dysregulated transcripts in colorectal tumors.

## **Methods**

### **Differentially Expressed Genes from GEO**

To investigate and integrate differentially expressed (DE) RNAs in colorectal cancer, including both coding and non-coding transcripts, four microarray datasets have been selected from GEO database and all of them were analyzed between colorectal tumors and non-tumor tissues. Information about the four used datasets is summarized in Table 1. All raw expression data were normalized and log<sub>2</sub>-transformed (Supplementary Figure 1). Limma, a Bioconductor package for

differential analysis of microarray data was run to determine DE RNAs in each dataset with the following criterias: Two datasets, GSE41657 (platform GPL6480) and GSE128435 (GPL4133 platform), concerning mRNA expression with cutoff criteria of  $\text{adj.p.value} < 0.05$  and  $\text{log fold change} \geq |2|$  were analyzed. The subseries 128435 is part of superseries GSE128449. Finally, the common genes between the both datasets that show expression in the same direction (up- or down-expressed) have been selected for further analysis.

GSE128449 (platform GPL14767) which is part of superseries GSE128449 was analyzed to determine DE miRNAs. In this step, DE miRNAs with  $\text{adj.p.value} < 0.05$  and  $\text{log fold change} \geq |3|$  were selected.

In order to assess DE circRNA in CRC, GSE126095 (GPL19978 platform) was used and transcripts with  $\text{adj.p.value} < 0.05$  and  $\text{log fold change} \geq |3|$  were retrieved. The flowchart of data and bioinformatics analyses is visualized in Figure 1. In order to investigate similarities and dissimilarities between samples, Clustered heatmaps for top 100 DE RNAs for four datasets have been analyzed.

### **Principal component analysis (PCA)**

The PCA plot evaluates differences and similarities among samples and determine whether samples can be grouped. Hence, it could separate normal and tumor tissues based on their gene expression profiles. Scatterplots of PCA from four microarray data have been separately generated using the ggfortify package in R.

### **circRNA-miRNA-mRNA network**

First, CircInteractome online database [7] was used to bioinformatically find potential miRNAs that have a complement seed region on sequences of DE circRNAs. For each DE circRNA, the entire predicted miRNAs were obtained. Second, the overlapping miRNAs between the predicted miRNAs and the DE miRNAs were obtained. Third, common targets of DE miRNAs with DE mRNAs that were retrieved from both GSE41657 and GSE128435 were assessed. To this end, miRTarBase [8] was used for detection of experimentally validated miRNA-mRNA interactions and two databases, TargetScan [9] and miRmap [10], were run for bioinformatically predicted miRNA-mRNA interactions. It should be noted that we have selected interactions that are commonly predicted in both databases. Finally, a dysregulated circRNA-miRNA-mRNA network in colorectal cancer was constructed by Cytoscape version 3.6.1. Characteristics of 16 DE circRNAs which are involved in this network are shown in Table 2.

Then, the axes that show expression status according to ceRNA hypothesis were extracted. In this step, we have presumed that circRNAs show opposite expression direction to their downstream miRNAs and same direction to downstream mRNAs in a circRNA-miRNA-mRNA axis. Cancer-Specific CircRNA (CSCD) database [11] was used to visualize basic structural patterns of the circRNAs involved in circRNA-miRNA-gene axis.

### **Co-expression Network Construction**

The top 40 co-expressed genes in colorectal cancer for each DE mRNA were retrieved from GEPIA database [12] according to the Cancer Genome Atlas (TCGA) data. The co-expression network was visualized by Cytoscape and hub nodes were identified according to the degree and betweenness centrality. Finally, the functional analysis of hub genes in this network was carried out by DAVID database [13].

### **Protein-Protein Interaction (PPI) Network Construction**

First, PPIs were identified by String database [14] and PPI network visualized by Cytoscape. To identify hub nodes, the degree and betweenness centrality were considered. Second, MCODE, a Cytoscape app, was run to find the most significant protein module. Third, Enrichr database [15] was used to understand biological function of hub genes in PPI network.

### **Oncogenomic and Oncotranscriptomic Analysis**

cBioPortal v.3 [16] was run to identify somatic alterations of hub genes in 526 colorectal tumor samples according to TCGA, PanCancer atlas. In addition, GEPIA was used to assess expression of hub genes in CRC tissues compared with normal tissues based on TCGA data.

## **Results**

### **Differentially Expressed mRNAs, miRNAs and circRNAs**

In order to obtain differentially expressed (DE) protein coding genes in CRC, two GEO datasets (GSE 41657 and GSE 128435) were analyzed. In this step genes with  $\text{adj.p.value} < 0.05$  and  $\log \text{fold change} \geq |2|$  were selected. Subsequently, 14 common DE genes were found in both datasets that show same dysregulation direction (Supplementary Table 1). Two of the genes are up-expressed and 12 are down-expressed. Details of the analysis are shown in Figure 1.

GEO dataset concerning miRNAs expression in CRC (GSE128449) was analyzed. 85 DE miRNAs with cutoff criteria of  $\text{adj.p.value} < 0.05$  and  $\log \text{fold change} \geq |3|$  were obtained (Supplementary Table 2).

As mentioned, GSE126095 for finding DE circRNAs in CRC was checked out. Thirty six DE circRNAs with  $\text{adj.p.value} < 0.05$  and  $\log \text{fold change} \geq |3|$  were obtained (Supplementary Table 3).

The clustered heatmaps for all analyzed datasets are shown in Figure 3 and Supplementary Figures 2-4. In addition the volcano plots of DE RNAs in four datasets have been drawn (Figure 2A).

The gene expression PCA plot provides insights into the association between samples. As it is shown in Figure 2B, colorectal tumor and normal tissues could be distinguished not only by mRNAs expression profile, but also by both circRNAs and miRNAs expression profiles.

### **Potential DE circRNA-DE miRNA and DE miRNA-mRNA interactions**

According to circInteractome database, 54 DE circRNA-DE miRNA interactions are bioinformatically predicted.

Three databases were run to investigate DE gene-DE miRNA interactions. Six experimentally validated interactions by mirTarbase and 50 bioinformatically predicted interactions by TargetScan and miRmap were consistently obtained (Supplementary Table 4).

### **CircRNA- miRNA- mRNA network**

The dysregulated circRNA- miRNA- mRNA network in colorectal cancer including 16 circRNAs, 14 miRNAs and 11 mRNAs was constructed (Figure 4A). The extracted axes according to ceRNA hypothesis have been visualized in Figure 4B including five circRNAs, four miRNAs and eight mRNAs. The involved circRNAs in these axes are available in Cancer-Specific CircRNA (CSCD) database (Figure 5).

### **Co-expression network construction**

As mentioned, 11 dysregulated genes related to DE miRNAs were detected. Their top 40 co-expressed genes in CRC were used to construct co-expression network (Figure 6). In this network, GAS7, C10ORF54, RELL1, and TSPAN1 are considered as hub genes because of having higher degree and betweenness centrality. In addition, functional analysis of hub genes indicates their roles in cancer related processes such as cell proliferation and migration (Table 3).

### **Protein-Protein interaction (PPI) network**

The PPI network of 11 DE genes and their co-expressed genes has been established after removing unconnected nodes based on String output (Figure 7A). This network shows PTPRC, SPI1, C3AR1, LCP2 and DOCK2 genes as hubnodes. We run MCODE to detect a significant module (Figure 7B). It should be noted that all hubnodes are involved in the detected module.

Functional annotation according to Enrichr web server revealed the roles of hub genes in immune system (Figure 7C).

### **Oncotranscriptomic and Oncogenomic Analysis**

Expression analysis for five hub nodes of PPI network using GEPIA revealed that while all of them are down-expressed in tumors compared with normal tissues, only DOCK2 and PTPRC dysregulation are statistically significant (Figure 8A).

Investigation of DOCK2 and PTPRC somatic mutations in colorectal cancer reveals a fairly high frequency (10% and 8% respectively) of genetic alterations in tumors. As it is shown in Figure 8B, the main proportion of these alterations is composed of missense and truncating mutations.

## Conclusion

Investigation of non-coding RNAs as well as identification of competing endogenous RNAs (ceRNAs) have become new hotspots in cancer research. Among them, circRNAs are more prominent due to their specific characterizations such as high stability, time- and tissue-specificity. circRNAs can play an important role in gene expression regulation at both transcriptional or posttranscriptional levels [17]. Various studies have demonstrated that dysregulated circRNAs are involved in cancer initiation and progression. For example, up-regulation of hsa\_circ\_0000069 and its role in cell proliferation, migration, and invasion were observed in CRC [18]. Recent evidences have disclosed that circRNAs could sponge miRNA and suppress their functions. Although, there are rare studies concerning circRNA molecular interactions in CRC.

miRNAs are a class of small noncoding RNAs which have pivotal roles in regulating gene expression at post-transcriptional level through suppressing target mRNA. It is well known that dysregulated miRNAs are involved in several oncopathways and result in cancer development. In this study, 36 DE circRNAs and 85 DE miRNA with log fold change  $\geq |3|$  were identified in CRC tissues compared with non-tumor tissues as well as 14 DE mRNAs, which consistently show dysregulation in both GSE41657 and GSE128435 GEO datasets. These DE transcripts were integrated using several databases and bioinformatics tools in order to construct a dysregulated circRNA-miRNA-mRNA network in CRC. In addition, mRNA, miRNA and circRNA datasets have been separately analyzed to assess their ability in separating colorectal tumor and normal samples. The results revealed that each of them could alone distinguish colorectal tumor and normal samples.

According to ceRNA hypothesis, we have presumed that circRNAs, as miRNA sponges, show opposite expression direction to their downstream miRNAs and same direction to downstream mRNAs in a circRNA-miRNA-mRNA axis. Since, 11 axes with this expression status were extracted from circRNA-miRNA-mRNA network, including five circRNAs, four miRNAs, and eight mRNAs. Previous cancer researches concerning these transcripts are summarized in Table 4. However, most of these circRNAs and miRNAs have not been previously reported in CRC. It should be noted that six of seven detected circRNAs are available in Cancer-Specific CircRNA Database (CSCD) and the location of their microRNA response elements (MRE) are shown in Fig. 4.

We have established potential dysregulated circRNA-miRNA-mRNA axes that could play pivotal roles in CRC oncotranscriptomic. Our analyses reveal that circRNAs could show both oncogenic (oncoCirc) and tumor suppressive functions in CRC. For example, circ\_0014879, an up-expressed exonic circRNA, could be an oncoCirc through circ\_0014879/miR-885-5p/ARG2 axis. miR-885-5p is bioinformatically predicted to be sponged by circ\_0014879. Interestingly, miR-885-5p is a tumor suppressor miRNA and down-expressed in different cancer (Table 4) as well as in our results. ARG2 is an enzyme that plays a part in immunosuppressive tumor microenvironment and tumorigenesis. ARG2 is predicted to be miR-885-5p target. This protein shows overexpression and association with poor prognosis in different cancers. In 2019, Youjun Wu et al. demonstrated

that expression of ARG2 is significantly higher in CRC samples than normal tissues [19]. In addition, they revealed downregulation of ARG2 results in inhibiting colorectal cancer cells growth. According to our investigation, ARG2 is significantly up-expressed in both GSE41657 and GSE128435 datasets with log fold change 2.15 and 2.05, respectively. Since, it is suggested that overexpression of circ\_0014879 as an oncoCirc could result in upregulation of ARG2 oncogene through inhibition of miR-885-5p function in CRC.

On the other hand, circ\_0001666, a down-expressed exonic circRNA, could be a tumor suppressive circRNA in CRC through sponging miR-1276 in circ\_0001666/miR-1276/RASSF2 axis. It is revealed that RASSF2 inhibits tumor cell growth and suggested that it is a tumor-suppressor gene in CRC [20]. Moreover, our results disclose potential tumor suppressive roles of circ\_0000977 in CRC through sponging miR-1208. Interestingly, this miRNA might post-transcriptionally inhibit two tumor suppressor genes in CRC including IL10RA and AKR1B10. In 2018, Zadka et al. reported that IL10RA is down-expressed in CRC tissues and shows negative correlation with the clinical stage and proliferation [21]. It is shown that AKR1B10 expression is significantly decreased in CRC samples and its down-expression correlates with decreased survival and poor prognosis of patients [22]. In addition, AKR1B10 knock down resulted in inhibition of apoptosis in colorectal cancer cells. It should be noted that expression of RASSF2, IL10RA and AKR1B10 was down-regulated in both GSE41657 and GSE128435 datasets with log fold change  $< -2$ . Collectively, it can be proposed that under-expression of both circ\_0001666 and circ\_0000977, as tumor suppressive circRNAs, contributes to down-regulation of some tumor suppressive genes in CRC including RASSF2, IL10RA and AKR1B10. These finding should be more validated by functional studies in future.

In the next step, we constructed a dysregulated co-expression network in CRC, including 11 DE genes and their co-expressed genes. This network shows GAS7, C10ORF54, TSPAN1, RELL1 genes are hub nodes and could play important roles in progression of CRC tumors as far as expression profile is concerned. Functional annotation based on DAVID database revealed their roles in cell proliferation, cell-cell adhesion, cell migration and cell differentiation (Table 3). Consistently, previous studies have demonstrated their dysregulation in CRC. Remarkable down-regulation of GAS7 due to promoter hypermethylation was reported in CRC [23, 24]. It is suggested that the downregulation of GAS7 results in high metastatic ability of tumor cells. RELL1 is a family member of tumor necrosis factor receptor (TNFR) that induces apoptosis [25, 26]. Here we proposed the important role of RELL1 in CRC oncotranscriptome for the first time; according to our results, RELL1 shows significant expression correlation with two down-regulated DE genes including RAB3B and ACER3 in CRC related TCGA data. C10ORF54 (VISTA) is a protein-coding gene, which codes immunoregulatory receptor. This protein inhibits the T-cell response and acts during adaptive immune responses [27]. Since, its expression can be a predictive biomarker as long as cancer immunotherapies are concerned [28]. Interestingly, our analyses based on TCGA data (Supplementary Figure 5) shows that GAS7 and C10ORF54 are significantly downregulated in CRC tissues. However, in 2018, Shan Xie et al. have reported that C10ORF54 expression was significantly high in CRC tissues compared with normal tissues [29]. It seems that

this controversy should be noticed in CRC patients who are supposed to receive immunotherapies. Another hub gene in co-expression network is TSPAN1, which functions in cell mitosis and differentiation. TCGA data show significant up-expression of TSPAN1 in CRC tumors than normal tissues (Supplementary Figure 5). It is reported that overexpression of TSPAN1 is correlated to the poor prognostic factors of CRC patients [30]. In 2010, Chen et al. disclosed that suppression of TSPAN1 results in reducing proliferation and invasion of colon cancer cells [31]. To assess the protein interactions between DE mRNAs and their co-expressed genes, we established a PPI network by STRING search tool. This network revealed two hub nodes DOCK2 and PTPRC which significantly downregulate in CRC tumors according to TCGA data (Fig. 7A). Oncogenomic analysis shows that a notable proportion of their downregulation is probably attributed to truncating mutations (Fig. 7B). Although it seems that the transcriptional and post-translational regulatory mechanisms are involved in down regulation of DOCK2 and PTPRC in CRC. Hematopoietic cells express DOCK2, which modulates activation and migration of immune cells. DOCK2 was shown to be involved in inflammatory diseases such as malignancies [32]. Yu et al. have demonstrated that DOCK2 mutation has a high frequency (7.7%) in colon tumors and low expression of DOCK2 is associated with poor prognosis of patients [33]. Another investigation has documented that DOCK2 is a prominent hypermethylated gene in CRC tissues [34]. PTPRC protein is necessary for T-cell activation and its mutation is associated with severe combined immunodeficiency. Down-regulation of PTPRC is reported in several cancer types [35, 36]. Here we show that DOCK2 and PTPRC are not only hub nodes of dysregulated PPI in CRC based on degree and betweenness centrality but also involved in identified module with MCODE algorithm (Fig. 6B). Functional enrichment analysis of module members indicates their roles in immune system (Fig. 6C). Taken together, it is suggested that down regulation of DOCK2 and PTPRC in CRC contributes to the escape of tumor from immune surveillance; accordingly, their dysregulation could be involved in resistance to immunotherapy.

In conclusion, in this study, combination of the both coding and non-coding microarray data as well as bioinformatics tools were applied to construct the dysregulated circRNA-miRNA-mRNA network and detect ceRNA axes in CRC. We identified potential oncoCirc (circ\_0014879) and tumor suppressive circRNAs (circ\_0001666 and circ\_0000977) in CRC that could become both prognostic and therapeutic biomarkers. In addition, PPI network implies pivotal roles of immune system related proteins such as DOCK2 and PTPRC in CRC progression, which could be important in immunotherapies.

## **Abbreviations**

CRC: colorectal cancer; circRNA: circular RNA; miRNA: microRNA; ceRNA: competing endogenous RNA; MRE: miRNA response elements; TCGA: the Cancer Genome Atlas; GEO: Gene Expression Omnibus; DE: differentially expressed; CSCD: Cancer-Specific CircRNA Database; PPI: Protein-Protein Interaction; PCA: Principal component analysis.

## **Acknowledgments**

We would like to acknowledge the GEO, GEPIA, STRING, DIANA, Enrichr, circInteractome, DAVID, cBioportal, CSCD, TargetScan, miRmap and miRTarBase databases as well as Cytoscape software for free use.

#### **Author's contributions**

S. Kadkhoda, F. Darbeheshti and G. Azizi-Tabesh conceived and designed the model and analyzed the data. S. Kadkhoda and J. Tavakkoly Bazzaz contributed to the interpretation of the results. F. Darbeheshti wrote the manuscript with support from N. Rezaei, F. Zolfaghari, R. Taslimi and S. Tavakolibazaz. All authors provided critical feedback and helped shape the research, analysis and manuscript.

#### **Funding**

No funding was used in this study.

#### **Availability of data**

The data that support these findings are available in the supplementary information files.

#### **Ethics approval and consent to participate**

Not applicable.

#### **Competing interests**

The authors declare that they have no competing interests

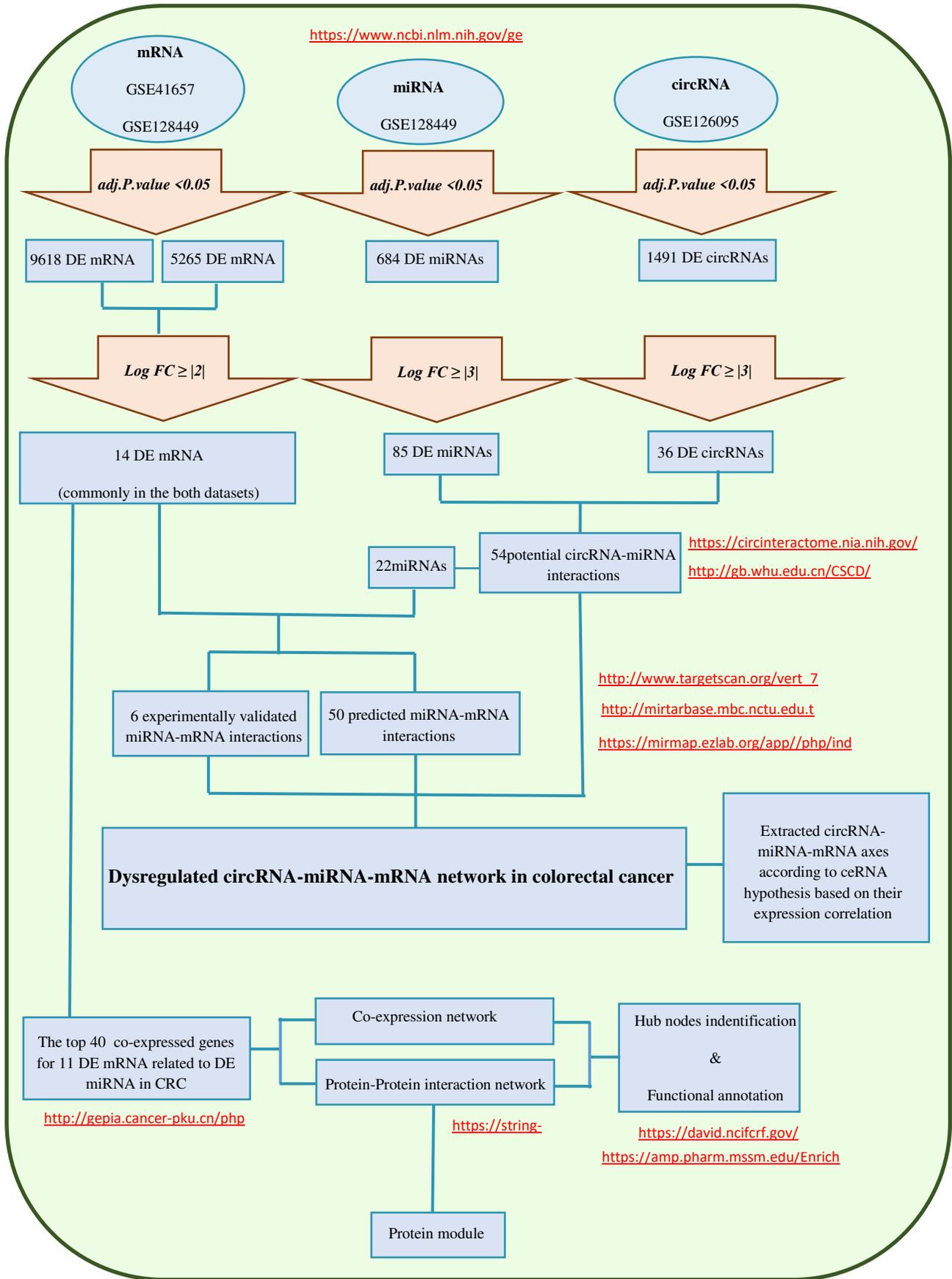
## References

1. Bray, F., J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*. 2018; 68(6): 394-424.
2. Vo, J. N., M. Cieslik, Y. Zhang, S. Shukla, L. Xiao, Y. Zhang, Y.-M. Wu, S. M. Dhanasekaran, C. G. Engelke and X. Cao. The landscape of circular RNA in cancer. *Cell*. 2019; 176(4): 869-881. e813.
3. Salmena, L., L. Poliseno, Y. Tay, L. Kats and P. P. Pandolfi. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011;146(3): 353-358.
4. Guan, Y.-j., J.-y. Ma and W. Song. Identification of circRNA-miRNA-mRNA regulatory network in gastric cancer by analysis of microarray data. *Cancer Cell International*. 2019; 19(1): 183.
5. Sun, X., X. Ge, Z. Xu and D. Chen. Identification of circRNA-miRNA-mRNA regulatory network in hepatocellular carcinoma by integrated analysis. *Journal of Gastroenterology and Hepatology*. 2019
6. Xu, H., C. Wang, H. Song, Y. Xu and G. Ji. RNA-Seq profiling of circular RNAs in human colorectal Cancer liver metastasis and the potential biomarkers. *Molecular cancer*. 2019; 18(1): 8.
7. Dudekula, D. B., A. C. Panda, I. Grammatikakis, S. De, K. Abdelmohsen and M. Gorospe. CircInteractome: a web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA biology*. 2016; 13(1): 34-42.
8. Chou, C.-H., N.-W. Chang, S. Shrestha, S.-D. Hsu, Y.-L. Lin, W.-H. Lee, C.-D. Yang, H.-C. Hong, T.-Y. Wei and S.-J. Tu. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic acids research*. 2015; 44(D1): D239-D247.
9. Lewis, B. P., I.-h. Shih, M. W. Jones-Rhoades, D. P. Bartel and C. B. Burge. Prediction of mammalian microRNA targets. *Cell*. 2003; 115(7): 787-798.
10. Vejnar, C. E., M. Blum and E. M. Zdobnov. miRmap web: comprehensive microRNA target prediction online. *Nucleic acids research*. 2013; 41(W1): W165-W168.
11. Xia, S., J. Feng, K. Chen, Y. Ma, J. Gong, F. Cai, Y. Jin, Y. Gao, L. Xia and H. Chang. CSCD: a database for cancer-specific circular RNAs. *Nucleic acids research*. 2017; 46(D1): D925-D929.
12. Tang, Z., C. Li, B. Kang, G. Gao, C. Li and Z. Zhang. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research*. 2017; 45(W1): W98-W102.
13. Huang, D. W., B. T. Sherman, Q. Tan, J. Kir, D. Liu, D. Bryant, Y. Guo, R. Stephens, M. W. Baseler and H. C. Lane. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic acids research*. 2007; 35(suppl\_2): W169-W175.
14. Szklarczyk, D., J. H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N. T. Doncheva, A. Roth and P. Bork. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic acids research*: gkw937. 2016.
15. Kuleshov, M. V., M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik and A. Lachmann. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research*. 2016; 44(W1): W90-W97.

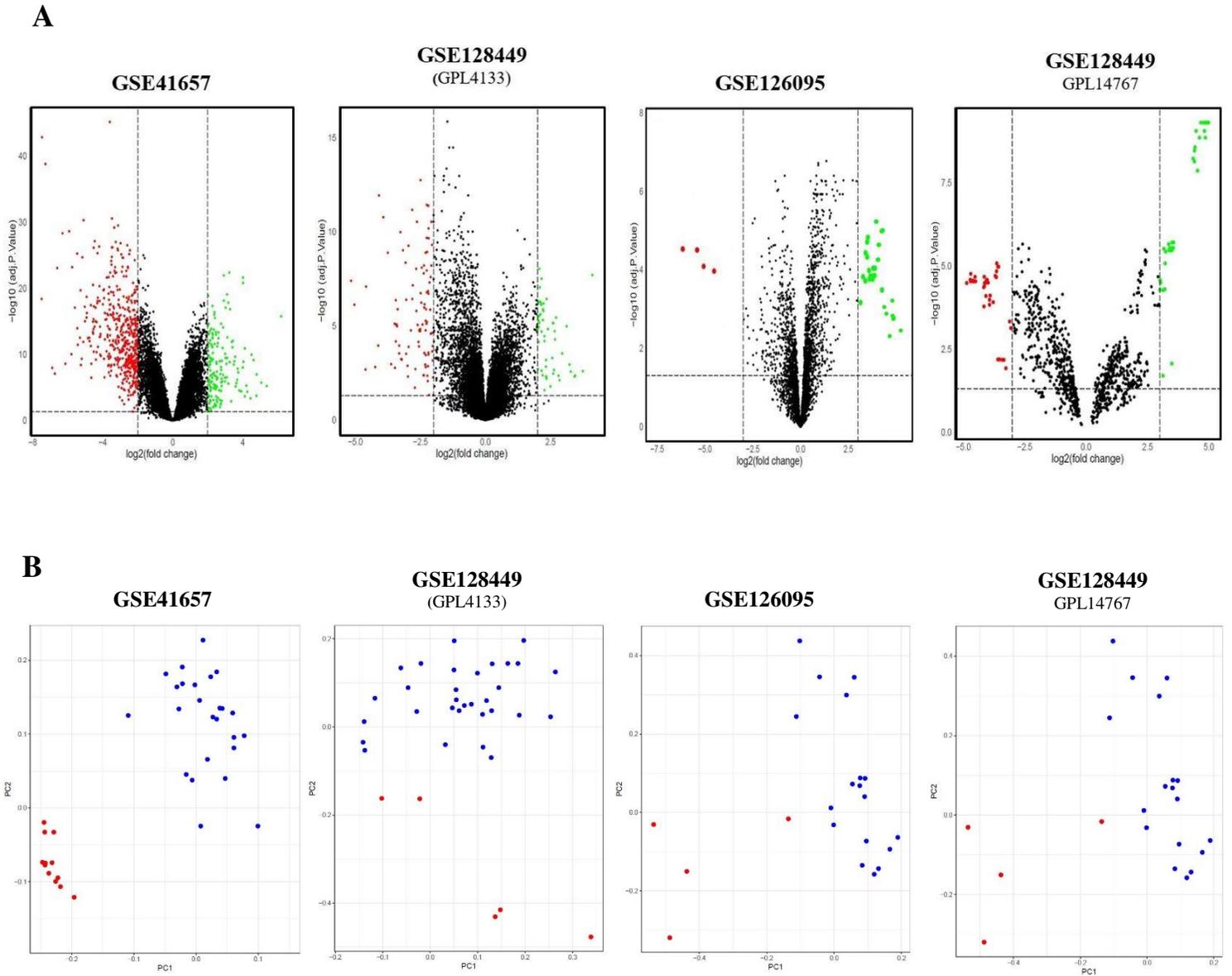
16. Gao, J., J. Lindsay, S. Watt, I. Bahceci, P. Lukasse, A. Abeshouse, H.-W. Chen, I. de Bruijn, B. Gross and D. Li. The cBioPortal for cancer genomics and its application in precision oncology, AACR. 2016.
17. Chen, S. and Y. Zhao. Circular RNAs: Characteristics, function, and role in human cancer. *Histology and histopathology*. 2018; 33(9): 887-893.
18. Guo, J.-n., J. Li, C.-l. Zhu, W.-t. Feng, J.-x. Shao, L. Wan, M.-d. Huang and J.-d. He. Comprehensive profile of differentially expressed circular RNAs reveals that hsa\_circ\_0000069 is upregulated and promotes cell proliferation, migration, and invasion in colorectal cancer. *OncoTargets and therapy*. 2016; 9: 7451.
19. Wu, Y., C. He, S. Hu, Z. Hu, Y. Li, X. Xing and X. Du. Downregulation of ARG2 inhibits growth of colorectal cancer cells and increases expression of the CD3 $\zeta$  chain in co-cultured T-cells. *international journal of clinical and experimental medicine*. 2019; 12(6): 6946-6957.
20. Akino, K., M. Toyota, H. Suzuki, H. Mita, Y. Sasaki, M. Ohe-Toyota, J.-P. J. Issa, Y. Hinoda, K. Imai and T. Tokino. The Ras effector RASSF2 is a novel tumor-suppressor gene in human colorectal cancer. *Gastroenterology*. 2005; 129(1): 156-169.
21. Zadka, L., M. J. Kulus, K. Kurnol, A. Piotrowska, N. Glatzel-Plucińska, T. Jurek, M. Czuba, A. Nowak, M. Chabowski and D. Janczak. The expression of IL10RA in colorectal cancer and its correlation with the proliferation index and the clinical stage of the disease. *Cytokine*. 2018; 110: 116-125.
22. Ohashi, T., M. Idogawa, Y. Sasaki, H. Suzuki and T. Tokino. AKR1B10, a transcriptional target of p53, is downregulated in colorectal cancers associated with poor prognosis. *Molecular Cancer Research*. 2013; 11(12): 1554-1563.
23. D'Arrigo, A., C. Belluco, A. Ambrosi, M. Digito, G. Esposito, A. Bertola, M. Fabris, V. Nofrate, E. Mammano and A. Leon. Metastatic transcriptional pattern revealed by gene expression profiling in primary colorectal carcinoma. *International journal of cancer*. 2005; 115(2): 256-262.
24. Ashktorab, H., M. Darempouran, A. Goel, S. Varma, R. Leavitt, X. Sun and H. Brim. DNA methylome profiling identifies novel methylated genes in African American patients with colorectal neoplasia. *Epigenetics*. 2014; 9(4): 503-512.
25. Cusick, J. K., A. Mustian, K. Goldberg and M. E. Reyland. RELT induces cellular death in HEK 293 epithelial cells. *Cellular immunology*. 2010; 261(1): 1-8.
26. Moua, P., M. Checketts, L.-G. Xu, H.-B. Shu, M. E. Reyland and J. K. Cusick. RELT family members activate p38 and induce apoptosis by a mechanism distinct from TNFR1. *Biochemical and biophysical research communications*. 2017; 491(1): 25-32.
27. Lines, J. L., L. F. Sempere, T. Broughton, L. Wang and R. Noelle. VISTA is a novel broad-spectrum negative checkpoint regulator for cancer immunotherapy. *Cancer immunology research*. 2014; 2(6): 510-517.
28. Lines, J. L., E. Pantazi, J. Mak, L. F. Sempere, L. Wang, S. O'Connell, S. Ceeraz, A. A. Suriawinata, S. Yan and M. S. Ernstoff. VISTA is an immune checkpoint molecule for human T cells. *Cancer research*. 2014; 74(7): 1924-1932.
29. Xie, S., J. Huang, Q. Qiao, W. Zang, S. Hong, H. Tan, C. Dong, Z. Yang and L. Ni. Expression of the inhibitory B7 family molecule VISTA in human colorectal carcinoma tumors. *Cancer Immunology, Immunotherapy*. 2018; 67(11): 1685-1694.
30. Chen, L., Y.-Y. Zhu, X.-J. Zhang, G.-L. Wang, X.-Y. Li, S. He, J.-B. Zhang and J.-W. Zhu. TSPAN1 protein expression: a significant prognostic indicator for patients with colorectal adenocarcinoma. *World Journal of Gastroenterology: WJG*. 2009; 15(18): 2270.

31. Chen, L., D. Yuan, R. Zhao, H. Li and J. Zhu. Suppression of TSPAN1 by RNA interference inhibits proliferation and invasion of colon cancer cells in vitro. *Tumori Journal*. 2010; 96(5): 744-750.
32. Chen, Y., F. Meng, B. Wang, L. He, Y. Liu and Z. Liu. Dock2 in the development of inflammation and cancer. *European journal of immunology*. 2018; 48(6): 915-922.
33. Yu, J., W. K. Wu, X. Li, J. He, X.-X. Li, S. S. Ng, C. Yu, Z. Gao, J. Yang and M. Li. Novel recurrently mutated genes and a prognostic mutation signature in colorectal cancer. *Gut*. 2015; 64(4): 636-645.
34. Liu, J., H. Li, L. Sun, Z. Wang, C. Xing and Y. Yuan. Aberrantly methylated-differentially expressed genes and pathways in colorectal cancer. *Cancer Cell International*. 2017; 17(1): 75.
35. Bankaitis-Davis, D. M., L. Siconolfi, K. Storm and K. Wassmann. Gene Expression Profiling for Identification, Monitoring and Treatment of Prostate Cancer, Google Patents. 2010
36. Kafetzopoulou, L. E., D. J. Boocock, G. K. R. Dhondalay, D. G. Powe and G. R. Ball. Biomarker identification in breast cancer: beta-adrenergic receptor signaling and pathways to therapeutic response. *Computational and structural biotechnology journal*. 2013; 6(7): e201303003.
37. Guo, S., A. Fesler, H. Wang and J. Ju. microRNA based prognostic biomarkers in pancreatic Cancer. *Biomarker research*. 2018; 6(1): 18.
38. Hou, S., J. Tan, B. Yang, L. He and Y. Zhu. Effect of alkylglycerone phosphate synthase on the expression profile of circRNAs in the human thyroid cancer cell line FRO. *Oncology letters*. 2018; 15(5): 7889-7899.
39. Su, H., F. Lin, X. Deng, L. Shen, Y. Fang, Z. Fei, L. Zhao, X. Zhang, H. Pan and D. Xie. Profiling and bioinformatics analyses reveal differential circular RNA expression in radioresistant esophageal cancer cells. *Journal of translational medicine*. 2016; 14(1): 225.
40. Qiu, L., T. Wang, Q. Ge, H. Xu, Y. Wu, Q. Tang and K. Chen. Circular RNA Signature in Hepatocellular Carcinoma. *Journal of Cancer*. 2019; 10(15): 3361.
41. Ou, M., H. Lin, W. Gong, F. Liu, P. Chen, Y. Zhang, J. Chen, W. Xue, W. Sui and Y. Dai. Comprehensive analysis of circRNA expression patterns in small hepatocellular carcinoma by integrating circRNA and gene expression data. *Int J Clin Exp Med*. 2017; 10(2): 2858-2865.
42. Li, Q.-h., Y. Liu, S. Chen, Z.-h. Zong, Y.-p. Du, X.-j. Sheng and Y. Zhao. circ-CSPP1 promotes proliferation, invasion and migration of ovarian cancer cells by acting as a miR-1236-3p sponge. *Biomedicine & Pharmacotherapy*. 2019; 114: 108832.
43. Reid, J. F., V. Sokolova, E. Zoni, A. Lampis, S. Pizzamiglio, C. Bertan, S. Zanutto, F. Perrone, T. Camerini and G. Gallino. miRNA profiling in colorectal cancer highlights miR-1 involvement in MET-dependent proliferation. *Molecular Cancer Research*. 2012; 10(4): 504-515.
44. Afanasyeva, E. A., P. Mestdagh, C. Kumps, J. Vandesompele, V. Ehemann, J. Theissen, M. Fischer, M. Zapatka, B. Brors and L. Savelyeva . MicroRNA miR-885-5p targets CDK2 and MCM5, activates p53 and inhibits proliferation and survival. *Cell death and differentiation*. 2011; 18(6): 974.
45. Zhang, Z., J. Yin, J. Yang, W. Shen, C. Zhang, W. Mou, J. Luo, H. Yan, P. Sun and Y. Luo. miR-885-5p suppresses hepatocellular carcinoma metastasis and inhibits Wnt/ $\beta$ -catenin signaling pathway. *Oncotarget*. 2016; 7(46): 75038.
46. Chen, L., Y. Li, Y. Fu, J. Peng, M.-H. Mo, M. Stamatakos, C. B. Teal, R. F. Brem, A. Stojadinovic and M. Grinkemeyer. Role of deregulated microRNAs in breast cancer progression using FFPE tissue. *PloS one*. 2013; 8(1): e54213.

47. Jin, S., Y. Dai, C. Li, X. Fang, H. Han and D. Wang. MicroRNA-544 inhibits glioma proliferation, invasion and migration but induces cell apoptosis by targeting PARK7. *American journal of translational research*. 2016; 8(4): 1826.
48. Xiong, D.-d., Y.-w. Dang, P. Lin, D.-y. Wen, R.-q. He, D.-z. Luo, Z.-b. Feng and G. Chen. A circRNA–miRNA–mRNA network identification for exploring underlying pathogenesis and therapy strategy of hepatocellular carcinoma. *Journal of translational medicine*. 2018; 16(1): 220.
49. Hou, B.-h., Z.-x. Jian, P. Cui, S.-j. Li, R.-q. Tian and J.-r. Ou. miR-216a may inhibit pancreatic tumor growth by targeting JAK2. *FEBS letters*. 2015; 589(17): 2224-2232.
50. Anauate, A. C., M. F. Leal, F. Wisnieski, L. C. Santos, C. O. Gigeck, E. S. Chen, D. Q. Calcagno, P. P. Assumpção, S. Demachki and C. H. Arasaki. Analysis of 8q24. 21 miRNA cluster expression and copy number variation in gastric cancer. *Future medicinal chemistry*. 2019; 11(09): 947-958.
51. Huppi, K., N. Volfovsky, T. Runfola, T. L. Jones, M. Mackiewicz, S. E. Martin, J. F. Mushinski, R. Stephens and N. J. Caplen. The identification of microRNAs in a genomically unstable region of human chromosome 8q24. *Molecular Cancer Research*. 2008; 6(2): 212-221.



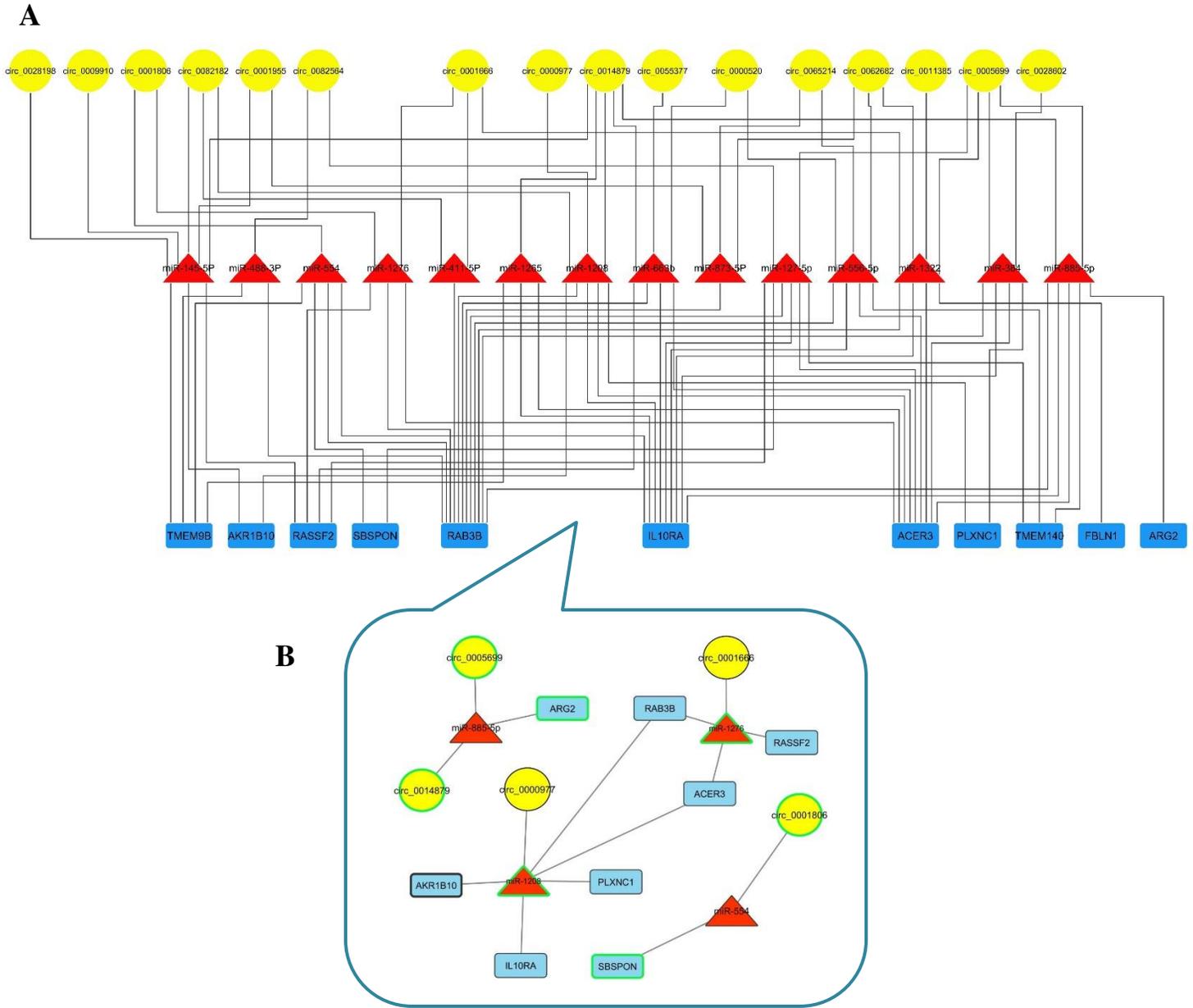
**Figure 1.** Flowchart of data and bioinformatics analyses



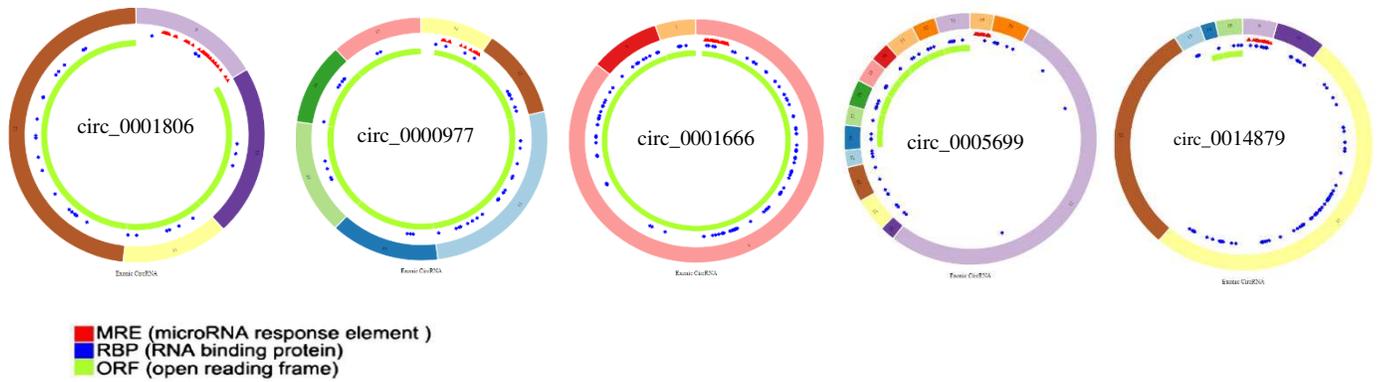
**Figure 2. A:** The Volcano plot of significant differentially expressed mRNAs (GSE41657 and GSE128449,GPL4133) with  $\geq |2|$  log fold change, circRNAs (GSE126095) and miRNAs (GSE128449,GPL14767) with  $\geq |3|$  log fold change in colorectal tumors and normal colorectal tissues. Red and green dots show down and up regulated RNAs, respectively. **B:** The scatterplot of PCA from the gene expression profiles in tumor and normal colorectal tissues, including two mRNA microarray datasets (GSE41657 and GSE128449,GPL4133), one circRNA microarray dataset (GSE126095), and one miRNA microarray dataset (GSE128449,GPL14767). The blue points represent colorectal tumor cells, whereas the red ones show normal cells. Distribution of information with respect to differential expression between tumor and normal tissues.



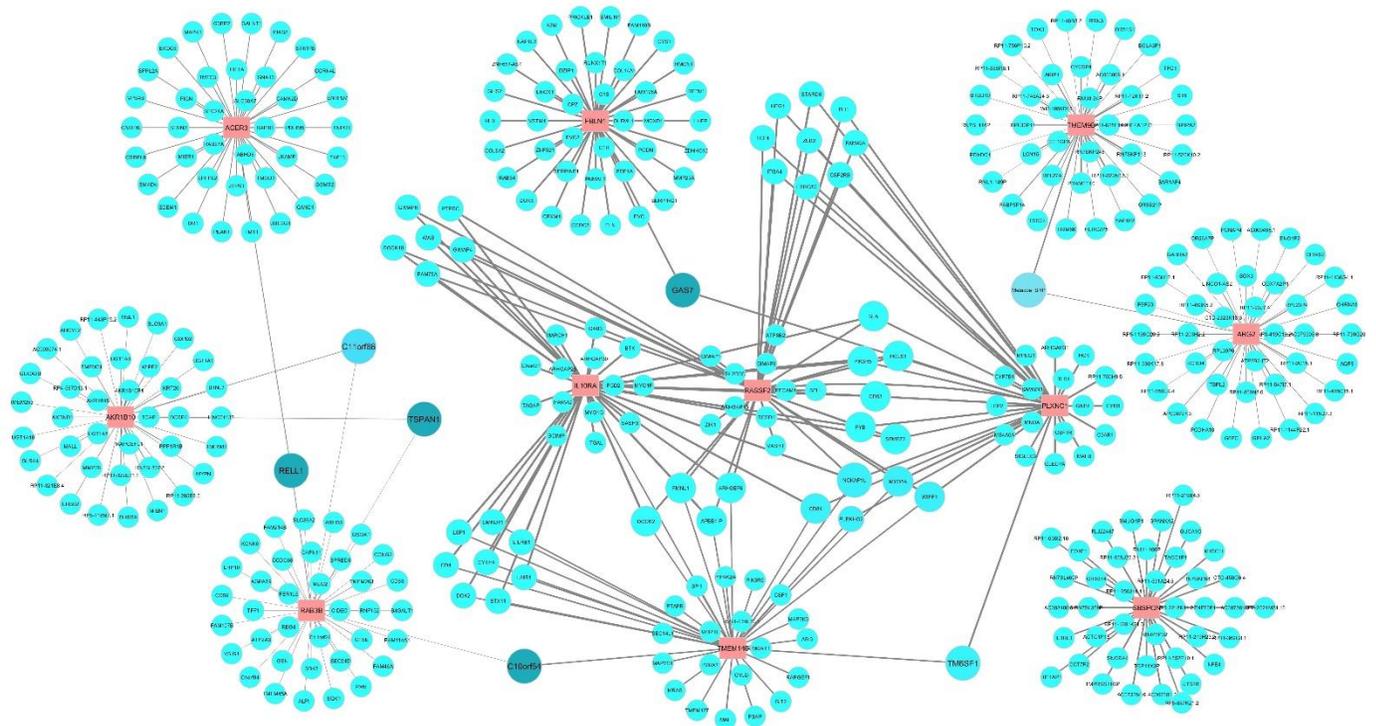
**Figure 3.** Clustered heatmap for the top 100 differentially expressed circRNAs according to GSE126095. Rows represent circRNAs and columns represent tissue types.



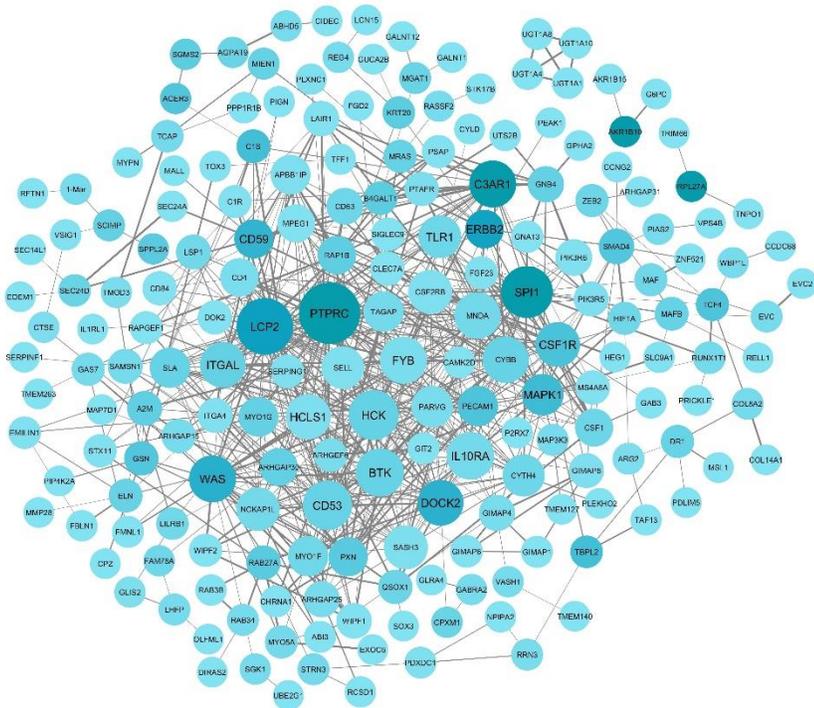
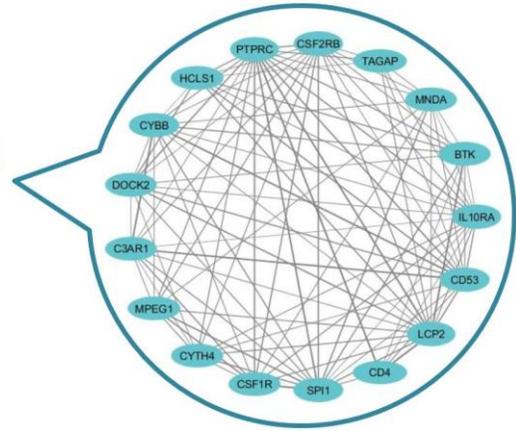
**Figure 4. A:** Dysregulated circRNA-miRNA-mRNA network in CRC. The yellow ellipses present circRNAs, the red triangles present miRNAs and the blue rectangles present mRNAs. The network was constructed by Cytoscape. **B:** Extracted axes from circRNA-miRNA-mRNA network that have an expression correlation according to ceRNA hypothesis. The nodes with black borderline are down-expressed and the nodes with green borderline are up-expressed in analyzed GEO datasets. The yellow ellipses present circRNAs, the red triangles present miRNAs and the blue rectangles present mRNAs.



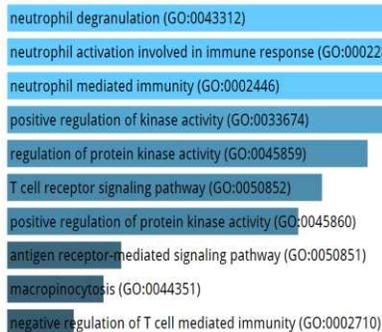
**Figure 5.** Structural patterns of the six circRNAs involved in circRNA-miRNA-gene axis (ceRNA hypothesis). Figures were retrieved from Cancer-Specific CircRNA (CSCD).



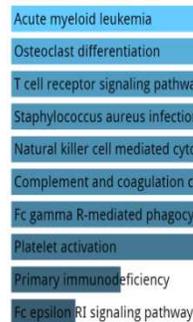
**Figure 6.** Co-expression network of 11 differentially expressed (DE) genes in colorectal cancer. The top 40 co-expressed gene for each DE gene were achieved from GEPIA webserver based on TCGA data. Data are visualized by Cytoscape. The nodes in red present 11 DE genes. The blue nodes present co-expressed genes, their size changes according to the node degree and the node color changes gradually from light blue to dark blue according to the betweenness centrality. The width of the edge indicates Pearson Correlation Coefficient (PCC).

**A****B****C**

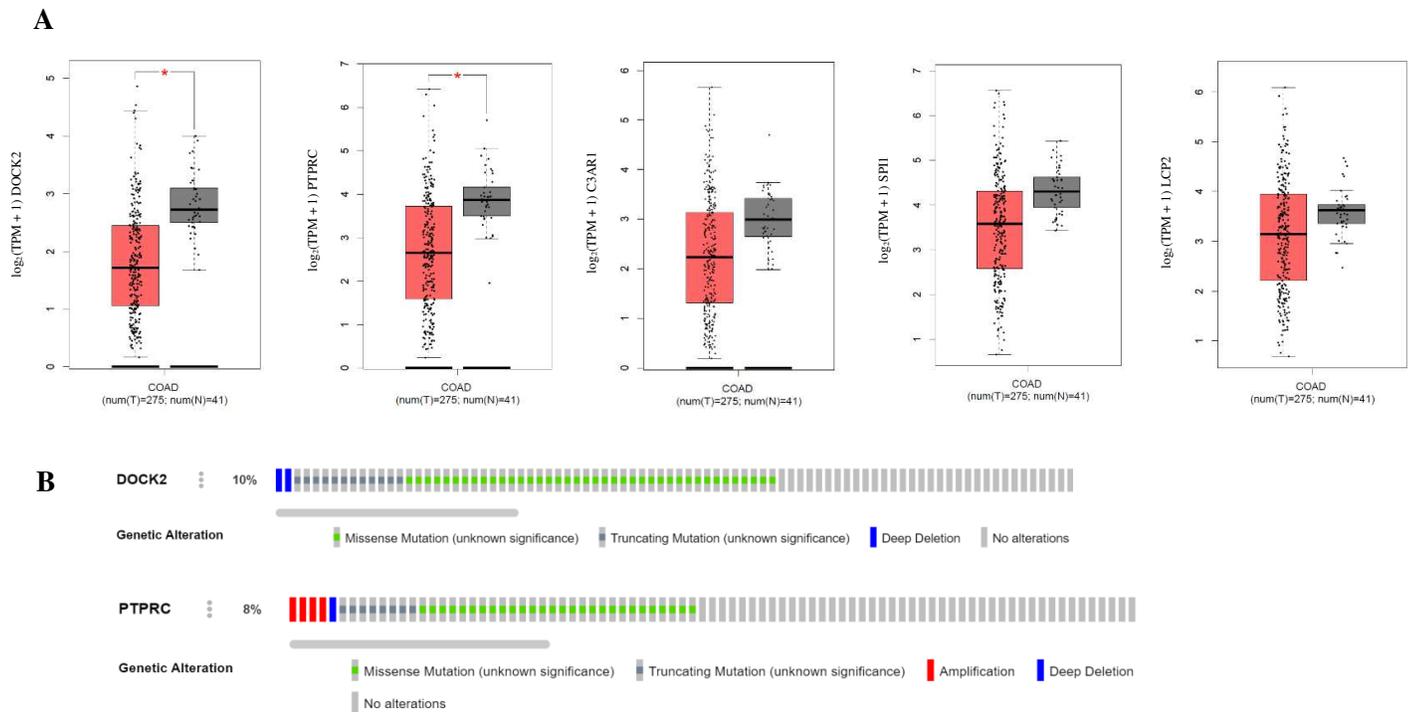
GO biological process



KEGG pathway



**Figure 7. A:** Protein-Protein Interaction (PPI) network. Data were retrieved from String based on 11 differentially expressed (DE) genes and their co-expressed genes in CRC and visualized by Cytoscape. The color of nodes were determined based on betweenness centrality and their size based on degree; darker color and bigger nodes have more betweenness centrality and degree. Also for determining of the edges width, combined score was considered. **B:** The main module is extracted from PPI by MCODE cytoscape app. **C:** Functional annotation for the five hub nodes including SPI1, DOCK2, PTPRC, C3AR1, and LCP2. The graphs were achieved from Enrichr and sorted based on combined score. The longer bar and brighter color represent term that is more



**Figure 8. A:** Expression analysis of five hub genes of PPI network between tumors and normal samples. The boxplots were retrieved from GEPIA based on TCGA data including expression of DOCK2, PTPRC, C3AR1, SPI1, and LCP2 from left to right, respectively. Tumors and normal samples present in red and gray, respectively. Red star indicates significant difference between tumors and normal tissues (P.value  $\leq 0.01$ ). **B:** Genetic alterations of DOCK2 and PTPRC in 526 colorectal tumor samples (the two hubgenes with significant difference expression between tumors and normal tissues). Figures were extracted from cBioPortal based on TCGA, PanCancer atlas.

**Table 1.** Characteristics of four colorectal cancer-related GEO microarray datasets. T: tumor; N: normal

<b>GEO</b>	<b>Platform</b>	<b>Year</b>	<b>Country</b>	<b>Sample size</b>
GSE41657	GPL6480	2015	China	25 T/12 N
GSE128449	GPL4133	2019	Spain	31 T/5 N
GSE128449	GPL14767	2019	Spain	18 T/4 N
GSE126095	GPL19978	2019	China	10 T/10 N

**Table 2.** Characteristics of differentially expressed circRNAs in GSE126095 with log fold change  $\geq 3$ . The red and black rows show down-expressed and up-expressed circRNAs, respectively.

circRNA	Alias	Log FC	Category	Chromosome	Host gene
hsa_circRNA_102619	hsa_circ_0000977	-5.2	exonic	chr2	NOL10
hsa_circRNA_104270	hsa_circ_0001666	-4.2	exonic	chr6	FAM120B
hsa_circRNA_104475	hsa_circ_0082182	3.6	exonic	chr7	FAM71F2
hsa_circRNA_100367	hsa_circ_0014879	3.3	exonic	chr1	DCAF8
hsa_circRNA_104640	hsa_circ_0001806	3.6	exonic	chr8	CSPP1
hsa_circRNA_103188	hsa_circ_0062682	3.4	exonic	chr22	TPST2
hsa_circRNA_103348	hsa_circ_0065214	3.2	exonic	chr3	SCAP
hsa_circRNA_101555	hsa_circ_0001955	4.1	exonic	chr15	CSNK1G1
hsa_circRNA_001846	hsa_circ_0000520	3.3	intragenic	chr14	RPPH1
hsa_circRNA_104499	hsa_circ_0082564	3.3	exonic	chr7	CREB3L2
hsa_circRNA_101744	hsa_circ_0005699	3.2	exonic	chr16	C16orf62
hsa_circRNA_100146	hsa_circ_0011385	3.9	exonic	chr1	EIF3I
hsa_circRNA_101145	hsa_circ_0028198	3.4	exonic	chr12	ANAPC7
hsa_circRNA_100053	hsa_circ_0009910	3	exonic	chr1	MFN2
hsa_circRNA_101164	hsa_circ_0028602	3.1	exonic	chr12	RNFT2
hsa_circRNA_102771	hsa_circ_0055377	3.1	exonic	chr2	CTNNA2

**Table 3.** Functional annotation of hub genes in co-expression network by DAVID database.

Hub genes	GO terms
GAS7	Cell cycle arrest, Regulation of transcription, DNA-templated, Nervous system development, Cell differentiation
C10ORF54	Single organismal cell-cell adhesion
TSPAN1	Cell proliferation, Cell migration, Cell surface receptor signaling pathway, Positive regulation of endocytosis, Protein stabilization
RELL1	Plasma membrane, Microtubule cytoskeleton, Integral component of membrane

**Table 4.** Summary of previous researches about dysregulation of circRNAs and miRNAs involved in circRNAs/miRNAs/mRNAs axes (Fig. 3) in different cancer types. The red and black transcripts respectively present down- and up-expression in CRC based on our analyses.

<b>Transcript</b>	<b>Article</b>	<b>Dysregulation</b>	<b>Cancer Type</b>
<b>circRNAs</b>			
hsa_circ_0000977	(Guo, Fesler et al. 2018)(37)	up	pancreatic Cancer
hsa_circ_0001666	(Hou, Tan et al. 2018)(38)	down	thyroid cancer
hsa_circ_0014879	(Su, Lin et al. 2016)(39)	up	esophageal cancer
	(Qiu, Wang et al. 2019)(40)	up	Hepatocellular Carcinoma
hsa_circ_0005699	(Ou, Lin et al. 2017)(41)	up	Hepatocellular Carcinoma
	(Guan, Ma et al. 2019)(4)	down	gastric cancer
hsa_circ_0001806	(Li, Liu et al. 2019)(42)	up	ovarian cancer
	(Qiu, Wang et al. 2019)(40)	up	Hepatocellular Carcinoma
<b>miRNAs</b>			
mir-885-5p	(Reid, Sokolova et al. 2012)(43)	down	colorectal cancer
	(Afanasyeva, Mestdagh et al. 2011)(44)	down	Neuroblastoma
	(Zhang, Yin et al. 2016)(45)	down	Hepatocellular Carcinoma
mir-554	(Chen, Li et al. 2013)(46)	down	breast cancer
	(Jin, Dai et al. 2016)(47)	down	glioma
mir-1276	(Xiong, Dang et al. 2018)(48)	up	Hepatocellular Carcinoma
	(Hou, Jian et al. 2015)(49)	up	pancreatic Cancer
Mir-1208	(Anauate, Leal et al. 2019)(50)	up	gastric cancer
	(Huppi, Volfovsky et al. 2008)(51)	down	breast cancer
	(Huppi, Volfovsky et al. 2008)(51)	down	colorectal cancer

# Figures

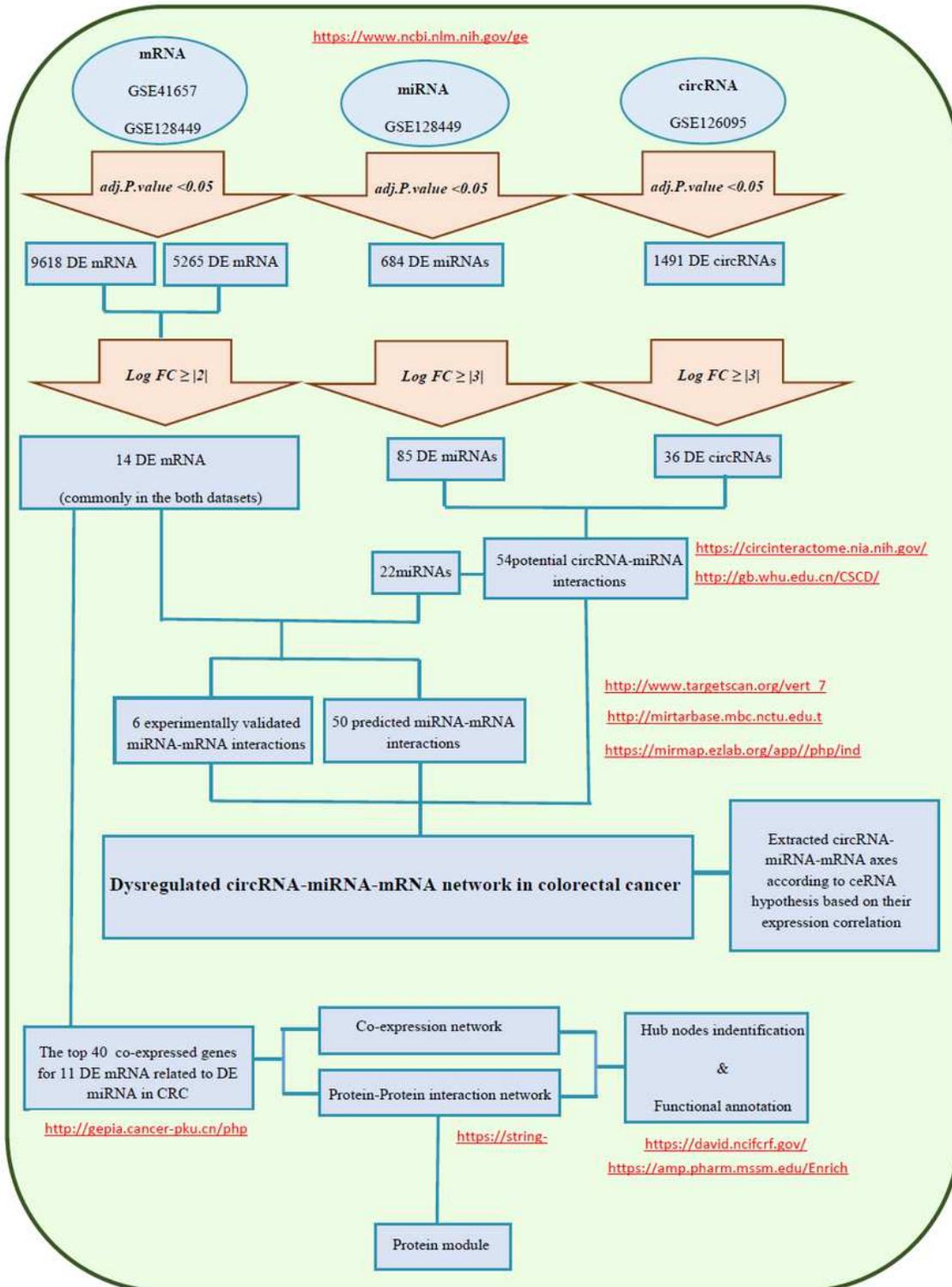
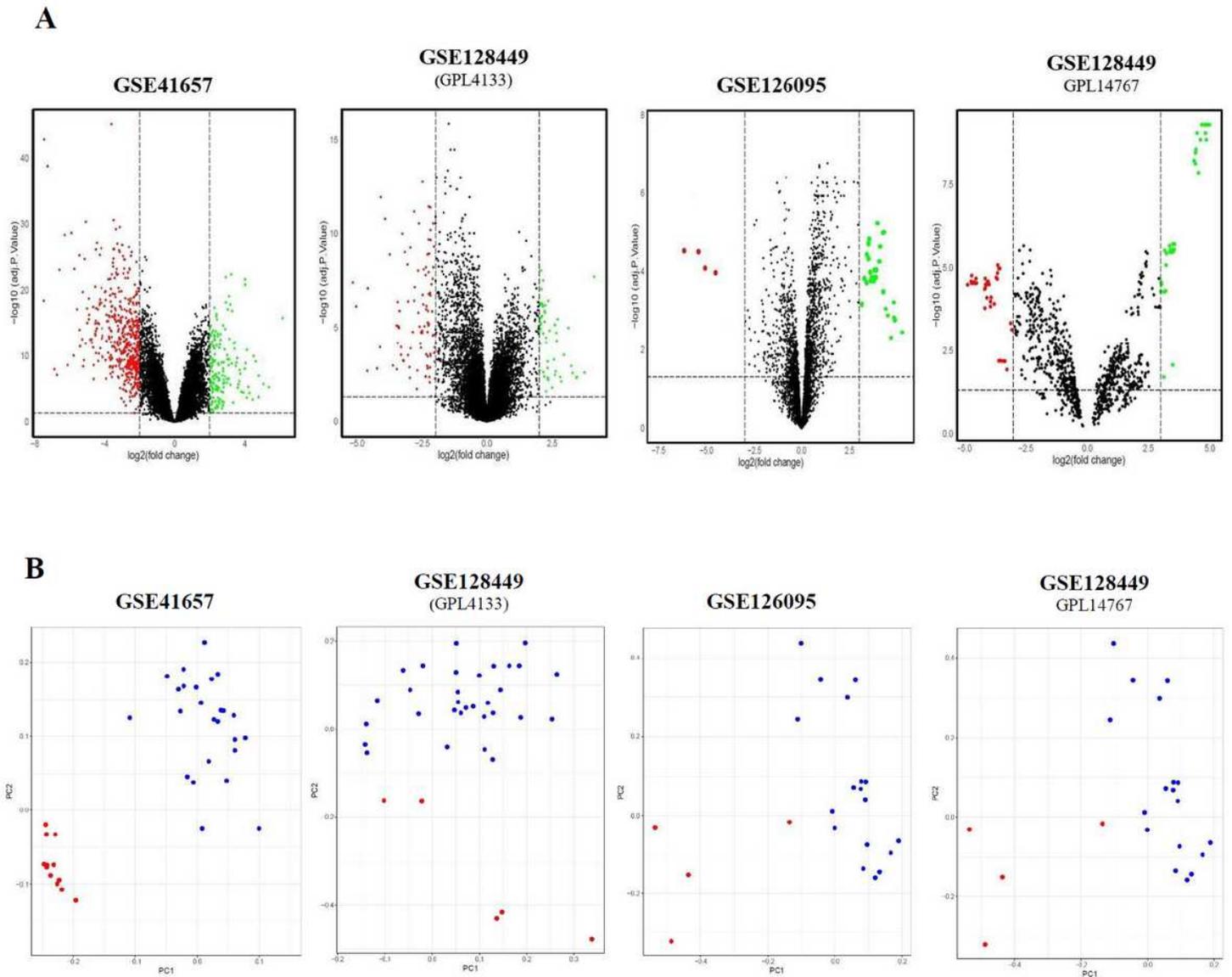


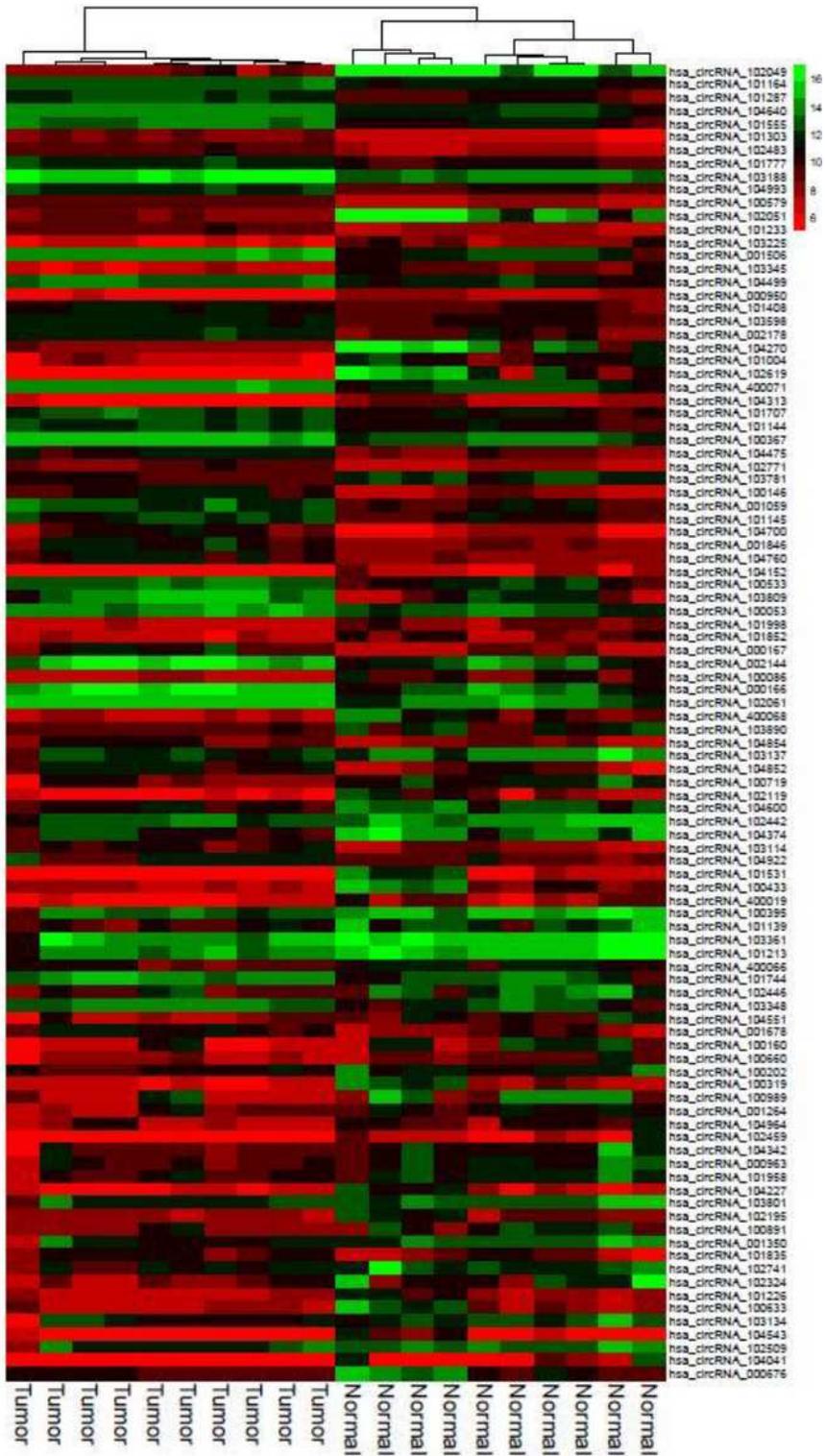
Figure 1

Flowchart of data and bioinformatics analyses



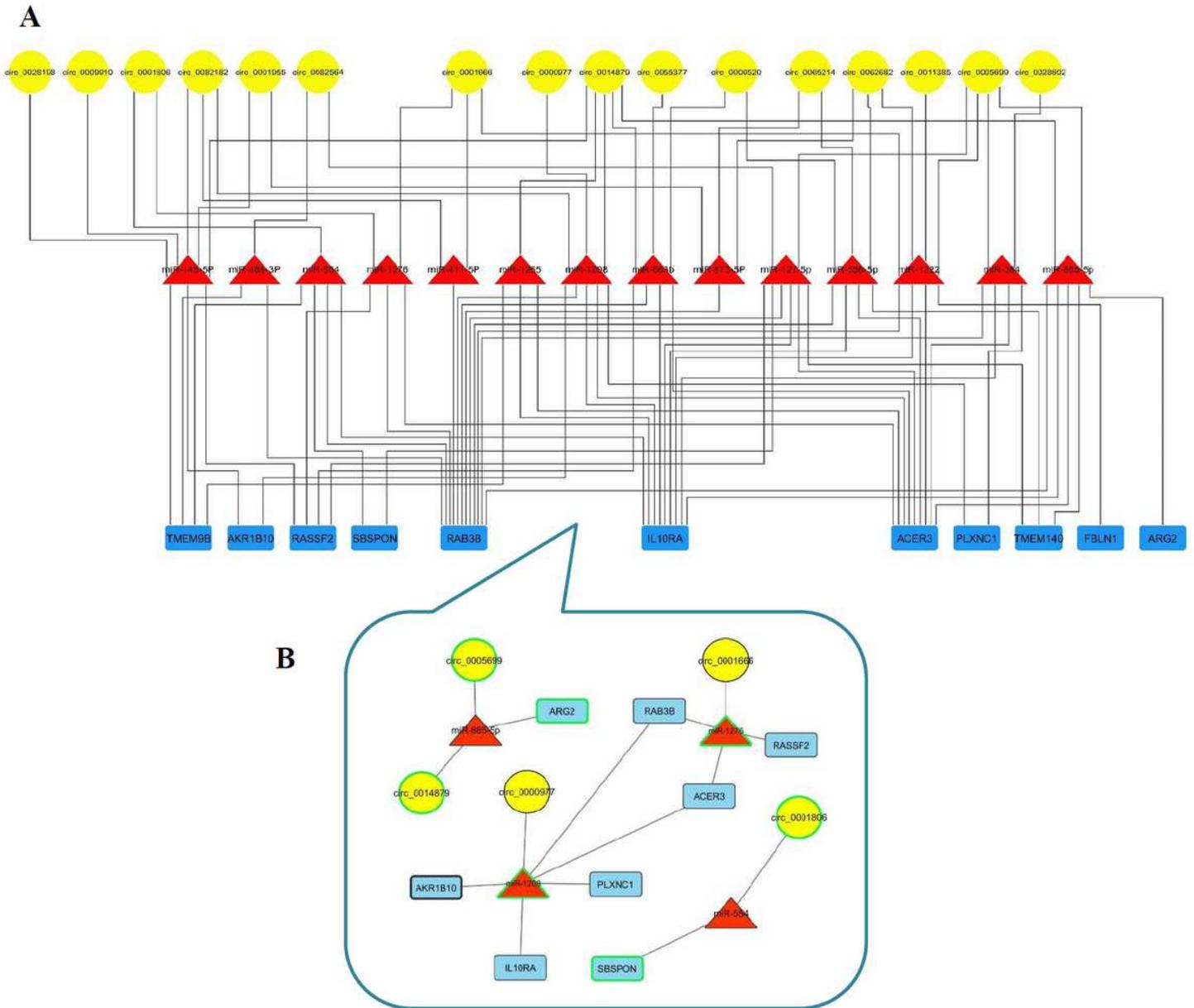
**Figure 2**

A: The volcano plot of significant differentially expressed mRNAs (GSE41657 and GSE128449,GPL4133) with  $\geq |2|$  log fold change, circRNAs (GSE126095) and miRNAs (GSE128449,GPL14767) with  $\geq |3|$  log fold change in colorectal tumors and normal colorectal tissues. Red and green dots show down and up regulated RNAs, respectively. B: The scatterplot of PCA from the gene expression profiles in tumor and normal colorectal tissues, including two mRNA microarray datasets (GSE41657 and GSE128449,GPL4133), one circRNA microarray dataset (GSE126095), and one miRNA microarray dataset (GSE128449,GPL14767). The blue points represent colorectal tumor cells, whereas the red ones show normal cells. Distribution of information with respect to differential expression between tumor and normal tissues.



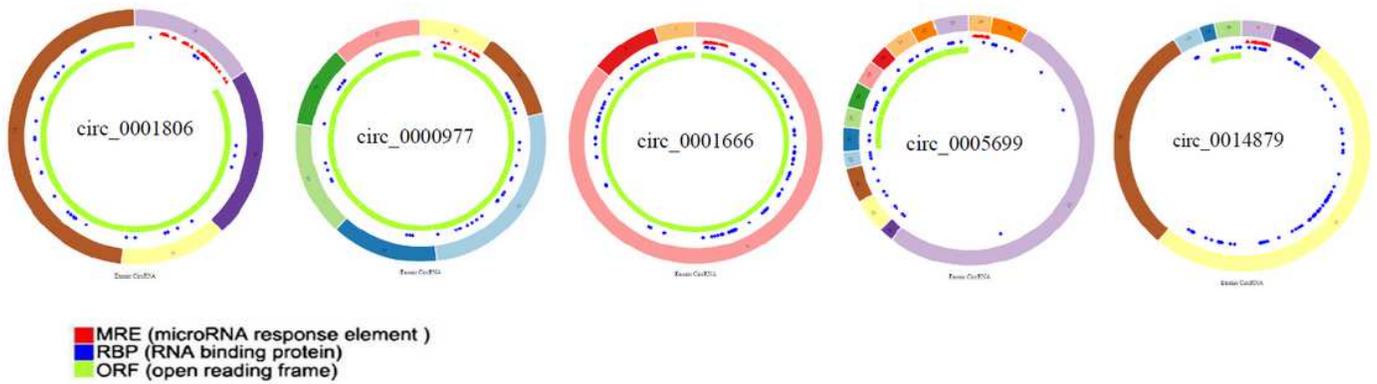
**Figure 3**

Clustered heatmap for the top 100 differentially expressed circRNAs according to GSE126095. Rows represent circRNAs and columns represent tissue types.



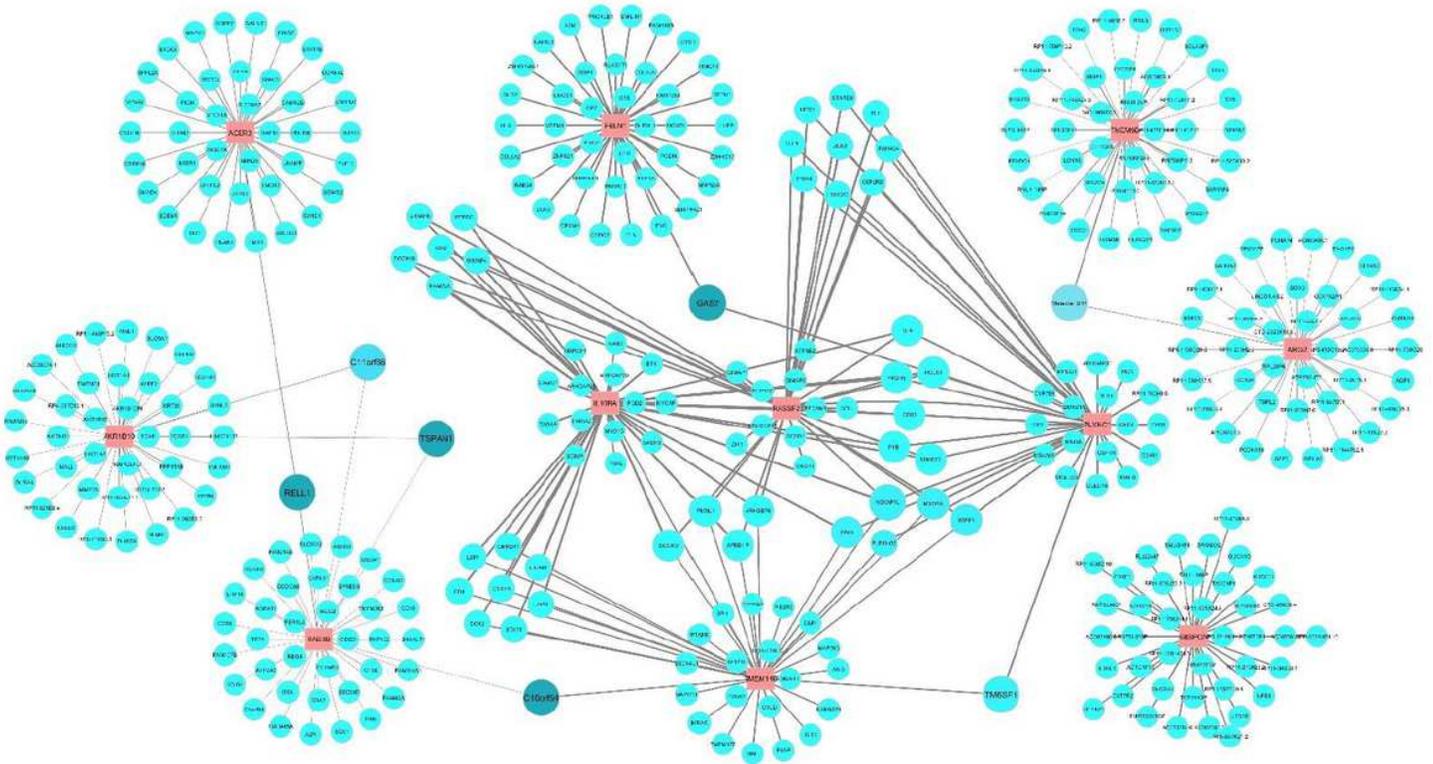
**Figure 4**

A: Dysregulated circRNA-miRNA-mRNA network in CRC. The yellow ellipses present circRNAs, the red triangles present miRNAs and the blue rectangles present mRNAs. The network was constructed by Cytoscape. B: Extracted axes from circRNA-miRNA-mRNA network that have an expression correlation according to ceRNA hypothesis. The nodes with black borderline are down-expressed and the nodes with green borderline are up-expressed in analyzed GEO datasets. The yellow ellipses present circRNAs, the red triangles present miRNAs and the blue rectangles present mRNAs.



**Figure 5**

Structural patterns of the six circRNAs involved in circRNA miRNA gene axis (ceRNA hypothesis). Figures were retrieved from Cancer Specific CircRNA (CSCD).

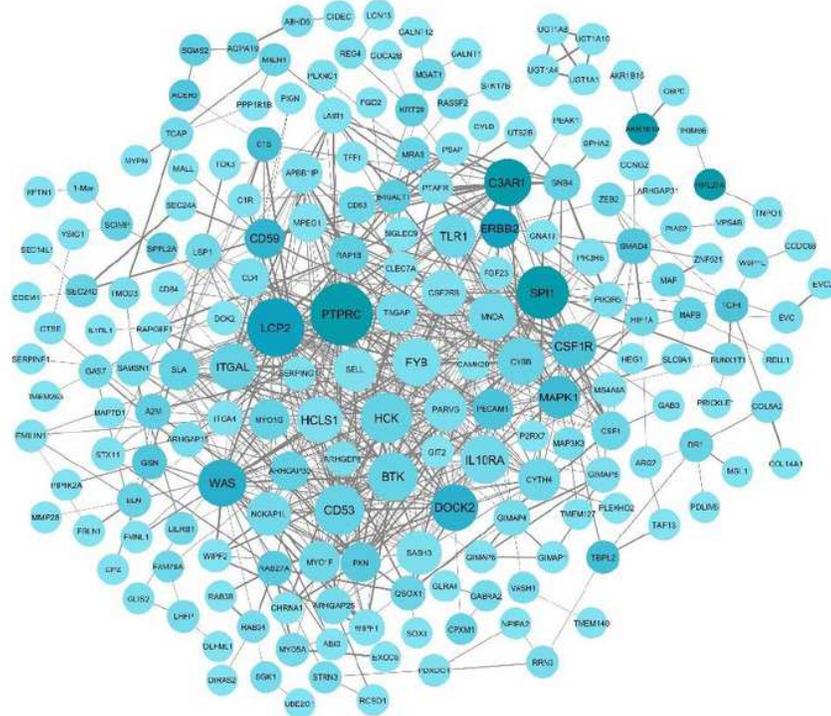


**Figure 6**

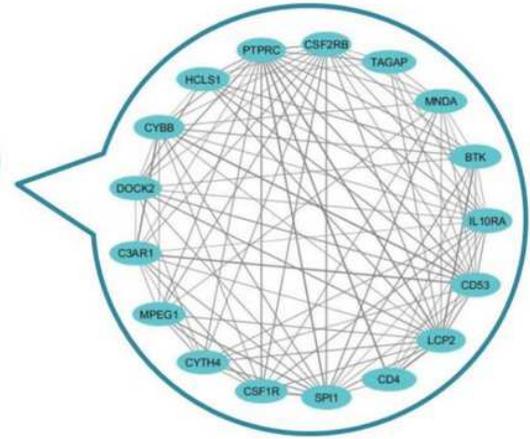
Co-expression network of 11 differentially expressed (DE) genes in colorectal cancer. The top 40 co-expressed gene for each DE gene were achieved from GEPIA webserver based on TCGA data. Data are visualized by Cytoscape. The nodes in red present 11 DE genes. The blue nodes present co-expressed genes, their size changes according to the node degree and the node color changes gradually from light

blue to dark blue according to the betweenness centrality. The width of the edge indicates Pearson Correlation Coefficient (PCC).

A



B



C

GO biological process

KEGG pathway

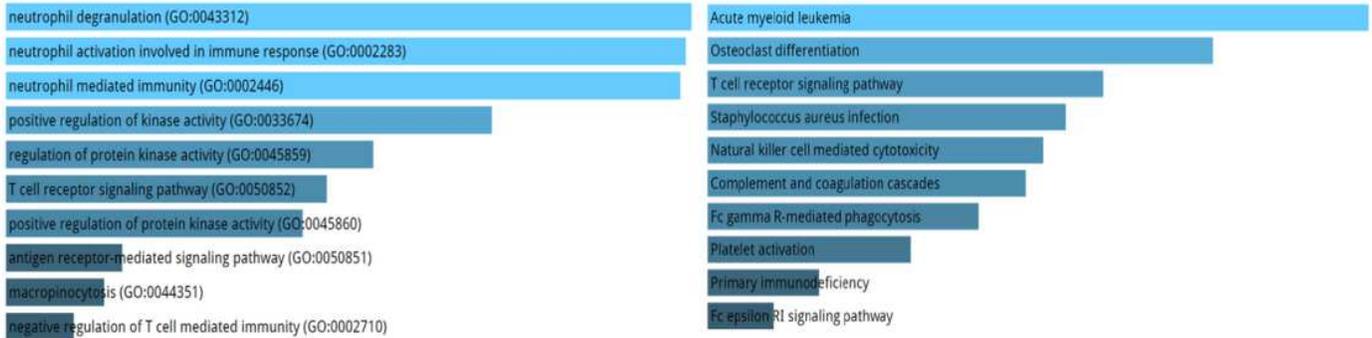
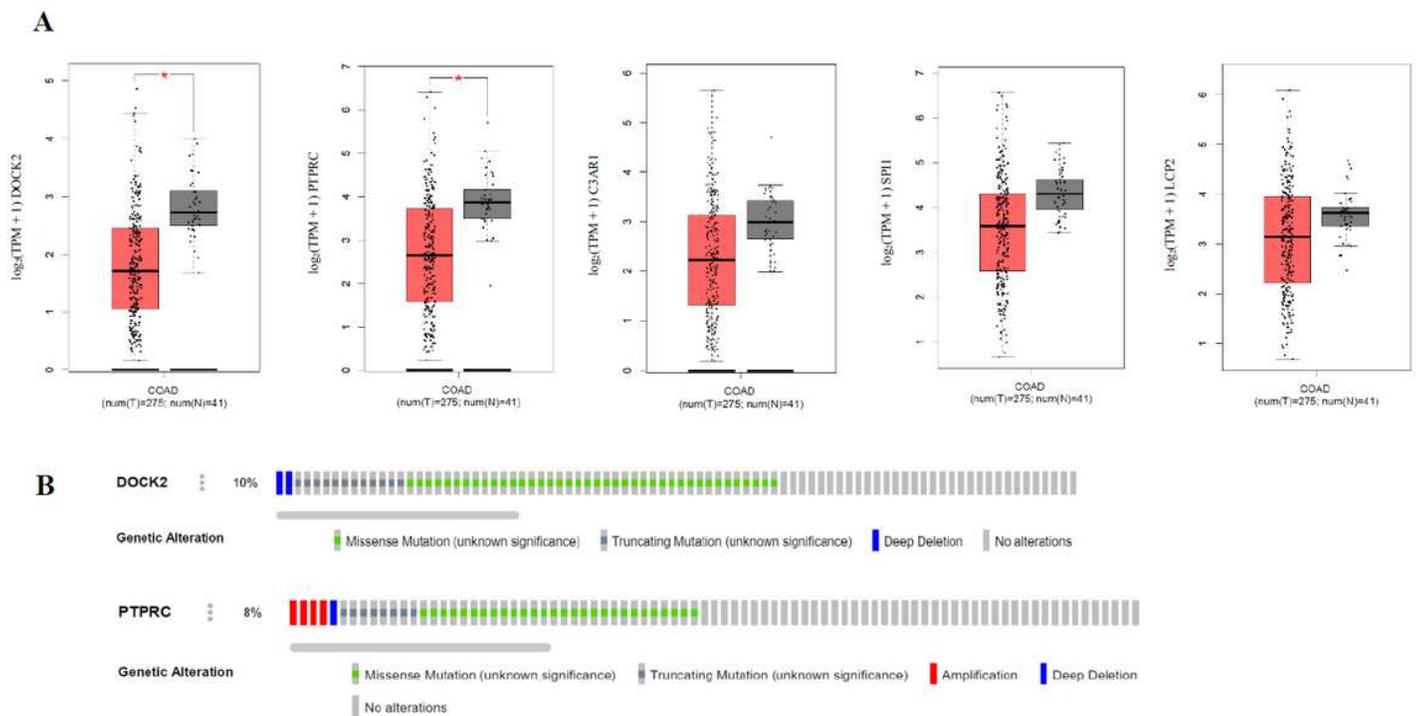


Figure 7

A: Protein-Protein Interaction (PPI) network. Data were retrieved from String based on 11 differentially expressed (DE) genes and their co-expressed genes in CRC and visualized by Cytoscape. The color of nodes were determined based on betweenness centrality and their size based on degree; darker color and bigger nodes have more betweenness centrality and degree. Also for determining of the edges width, combined score was considered. B: The main module is extracted from PPI by MCODE cytoscape app. C: Functional annotation for the five hub nodes including SPI1, DOCK2, PTPRC, C3AR1, and LCP2. The

graphs were achieved from Enrichr and sorted based on combined score. The longer bar and brighter color represent term that is more



**Figure 8**

A: Expression analysis of five hub genes of PPI network between tumors and normal samples. The boxplots were retrieved from GEPIA based on TCGA data including expression of DOCK2, PTPRC, C3AR1, SPI1, and LCP2 from left to right, respectively. Tumors and normal samples present in red and gray, respectively. Red star indicates significant difference between tumors and normal tissues (P.value  $\leq 0.01$ ).

B: Genetic alterations of DOCK2 and PTPRC in 526 colorectal tumor samples (the two hubgenes with significant difference expression between tumors and normal tissues). Figures were extracted from cBioPortal based on TCGA, PanCancer atlas.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfigures.pdf](#)
- [supplementary.xlsx](#)