



RNA-Associated Co-expression Network Identifies Novel Biomarkers for Digestive System Cancer

Zheng Chen^{1,2†}, Zijie Shen^{2†}, Zilong Zhang², Da Zhao^{1,2}, Lei Xu^{3*} and Lijun Zhang^{1*}

¹ School of Applied Chemistry and Biological Technology, Shenzhen Polytechnic, Shenzhen, China, ² Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China, ³ School of Electronic and Communication Engineering, Shenzhen Polytechnic, Shenzhen, China

OPEN ACCESS

Edited by:

Wei Jiang,
Nanjing University of Aeronautics
and Astronautics, China

Reviewed by:

Yi Xiong,
Shanghai Jiao Tong University, China
Lei Deng,
Central South University, China

*Correspondence:

Lei Xu
csleixu@szpt.edu.cn
Lijun Zhang
c7zlj@szpt.edu.cn

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 28 January 2021

Accepted: 25 February 2021

Published: 26 March 2021

Citation:

Chen Z, Shen Z, Zhang Z,
Zhao D, Xu L and Zhang L (2021)
RNA-Associated Co-expression
Network Identifies Novel Biomarkers
for Digestive System Cancer.
Front. Genet. 12:659788.
doi: 10.3389/fgene.2021.659788

Cancers of the digestive system are malignant diseases. Our study focused on colon cancer, esophageal cancer (ESCC), rectal cancer, gastric cancer (GC), and rectosigmoid junction cancer to identify possible biomarkers for these diseases. The transcriptome data were downloaded from the TCGA database (The Cancer Genome Atlas Program), and a network was constructed using the WGCNA algorithm. Two significant modules were found, and coexpression networks were constructed. CytoHubba was used to identify hub genes of the two networks. GO analysis suggested that the network genes were involved in metabolic processes, biological regulation, and membrane and protein binding. KEGG analysis indicated that the significant pathways were the calcium signaling pathway, fatty acid biosynthesis, and pathways in cancer and insulin resistance. Some of the most significant hub genes were *hsa-let-7b-3p*, *hsa-miR-378a-5p*, *hsa-miR-26a-5p*, *hsa-miR-382-5p*, and *hsa-miR-29b-2-5p* and *SECISBP2 L*, *NCOA1*, *HERC1*, *HIPK3*, and *MBNL1*, respectively. These genes were predicted to be associated with the tumor prognostic reference for this patient population.

Keywords: cancer, long non-coding RNA, WGCNA, MicroRNA, hub genes

INTRODUCTION

Cancer is one of the most common and lethal disease types in humans, and the molecular mechanisms governing cancer progression have not been elucidated to date (Liang et al., 2019; Yang et al., 2020; Zhang Z.M. et al., 2020). Tumors of the digestive system are a serious problem for human health, and esophageal tumors, colorectal tumors, gastric tumors and rectosigmoid-junction tumors are among the ten most common cancers worldwide. These cancers cause the most deaths, and the incidence rate is increasing year by year (Fitzmaurice et al., 2017, 2018, 2019; Bray et al., 2018; Doja et al., 2020; Hong et al., 2020; Li S. et al., 2020; Zeng et al., 2020a; Zhang L. et al., 2020). Therefore, the study of digestive system tumors has important practical significance.

Digestive system tumors are a common malignancy in the clinic. Colorectal cancer is the combination of colon and rectal cancer cases (Slattery et al., 2007; Ghosh and Yan, 2020). Colorectal cancer ranks third in men and second in women in the rankings of cancer incidence and is commonly diagnosed despite the increased death risk worldwide (Torre et al., 2015; Ozkan et al., 2019; Li J. et al., 2020). Colorectal cancer (CRC) is derived from the accumulation of epigenetic and genetic changes that lead to cancer-related deaths worldwide. With significant advancements in the early diagnosis of colorectal cancer, the cancer mortality rate has increased in the United States

(Rahman et al., 2015; Callahan et al., 2019; Ma et al., 2020). In European countries, colorectal cancer also has a high incidence, and the annual rate is increasing worldwide (Ferlay et al., 2015; Amer et al., 2020). Although the cancer-related death rate increases every year, treatments can be administered at the early stages of cancer (Pawa et al., 2011; Zhang Y. et al., 2019). Other digestive system tumors, such as esophageal cancer (ESCC), have been reported as the eighth most common cancer worldwide with a high cancer-related death rate (Ferlay et al., 2010; Abbas and Krasna, 2017). According to previous studies, almost 400,000 patients die from this disease among 450,000 diagnosed patients, the cancer-related death rate is approximately equal to 87.8% worldwide, and the 5-year survival rate is as low as approximately 25% (Rustgi and El-Serag, 2014, 2015). The most common cancer in Asia is gastric cancer (GC), which is the fifth ranking cancer leading to the third highest cancer-related death rate in the world. At stages I and IIA, the GC 5-year survival rate is 81.8–93.6%, but at stage IIIC, it is very low at only 17.9% (Amin et al., 2017; Lin et al., 2019).

Recently, network analysis has become a popular method for large-scale data (Cheng et al., 2019b; Zeng X. et al., 2019; Jin et al., 2020; Liu et al., 2020; Zeng et al., 2020b), and the WGCNA is a popular method for these coexpression network constructs (Jiang et al., 2010; Zeng W. et al., 2019; Iliopoulos et al., 2020). Five steps were included in this method: 1. a gene coexpression network was constructed by an adjacency matrix; 2. network modules were identified by the hierarchical cluster method; 3. modules with related phenotypes were analyzed; 4. module relationships were analyzed; and 5. key modules were found. The WGCNA method is widely used in biological networks and genomic data mining analysis, and also gives rise to network-based meta-analysis techniques (Langfelder et al., 2013).

The theoretical basis of the ceRNA (competitive endogenous RNA) hypothesis is that lncRNAs interact directly to regulate the expression of targeted genes and indirectly combine with miRNA sequences via the common miRNA response elements (MREs) of lncRNAs (Sardina et al., 2017). Cytoplasmic lncRNAs mainly affect mRNA reliability and translation mechanisms by binding miRNAs in the ceRNA regulatory network (Cao et al., 2018). Recently, analysis of lncRNA, miRNA, and mRNA networks has been identified to be involved in the progression of cancer (Zhang X. et al., 2017; Zeng et al., 2018), including endogenous cancer, hepatocellular carcinoma cancer and other malignant tumors, using the WGCNA method (Wang et al., 2010; Tay et al., 2014; Wei et al., 2014, 2018, 2019a,b; Thomson and Dinger, 2016; Jiang et al., 2018, 2019; Bai et al., 2019; Cheng et al., 2019a; Jin et al., 2019; Long et al., 2019; Shen et al., 2019). To date, there are few studies on digestive system malignant tumors. Therefore, integrated analysis of the regulatory functions of digestive system malignant tumors for lncRNA-miRNA-mRNA interaction networks requires large sample data and methods (Jiang et al., 2014; Zeng et al., 2016; Zou et al., 2016; Liu et al., 2017; Cheng et al., 2018; Cheng, 2019).

Competitive endogenous RNAs (ceRNAs) represent a novel gene expression regulation model and have attracted much attention from the academic community in recent years (Zhao et al., 2020). Compared with the miRNA regulation network, the

ceRNA regulation network is more sophisticated and complex because the ceRNA regulation network includes more RNA molecules, such as mRNAs, pseudogenes, long non-coding RNAs (lncRNAs) and miRNAs. The ceRNA network provides a new perspective for transcriptomic and biological research.

In our studies, we used RNA-seq data from TCGA for free-scale gene coexpression network construction of digestive system malignant tumors. This study was able to provide novel biomarkers for these malignant tumors. Therefore, potential biomarkers were identified by using bioinformatics methods for integrative analysis based on the large amount of RNA-seq. According to the results of the analysis of malignant tumors, *NCOA1*, *HERC1*, *HIPK3*, *MBNL1*, *hsa-let-7b-3p*, *hsa-miR-378a-5p*, and *hsa-miR-26a-5p* were predictive or prognostic factors for malignant tumors.

DATA AND METHODS

Date Collection

For data analysis in this study, *gdc-client* 1.5.0 *ww*¹ and *TCGAbiolinks* 2.16.3 *tools*² (Colaprico et al., 2016) were used from the TCGA database. A total of 437 samples of mRNA data, 388 samples of miRNA data and 385 samples of clinical data for colon cancer were obtained from TCGA, including colon cancer samples and normal samples. For colon cancer, there were over 500 individuals, including 480 tumor and 41 normal samples. Eighty-seven samples of mRNA data, 97 samples of miRNA data, and 87 samples of clinical data for ESCC were obtained from TCGA, including esophageal cancer (ESCC) samples and normal samples. There were 90 samples of mRNA data, 79 samples of miRNA data and 87 samples of clinical data for rectal cancer obtained from TCGA, including ESCC samples and normal samples. A total of 373 samples of mRNA data, 452 samples of miRNA data and 406 samples of clinical data for gastric cancer (GC) were obtained from TCGA, including ESCC samples and normal samples. Sixty-six samples of mRNA data, 63 samples of miRNA data and 63 samples of clinical data for rectosigmoid junction cancer were obtained from TCGA, including ESCC samples and normal samples. The guidelines of this study are from the TCGA website³.

Computational Analysis of RNA-Seq Data

The level three data of these cancers from TCGA by Illumina HiSeq 2000 platform (Illumina Inc., San Diego, California, United States) including miRNA, lncRNA, and mRNA data were analyzed by Digital Gene Expression Data package (*edgeR* 3.30.3)⁴ and *limma* (*limma* 3.44.3)⁵ package in R. The genes were annotated gene symbols through the Ensembl database⁶

¹<https://github.com/NCI-GDC/gdc-client/>

²<http://www.bioconductor.org/packages/release/bioc/html/TCGAbiolinks.html>

³<http://cancergenome.nih.gov/publications/publicationguidelines>

⁴<http://www.bioconductor.org/packages/release/bioc/html/edgeR.html>

⁵<http://www.bioconductor.org/packages/release/bioc/html/limma.html>

⁶<https://uswest.ensembl.org/index.html>

(Zerbino et al., 2018). According to these analyses, absolute $\log_2(\text{fold change}) \geq 2.0$ and $\text{FDR} \leq 0.01$ were set to screen the significant differentially expressed genes of mRNAs (DEGs), lncRNAs (DELncRNAs) and miRNAs (DEmiRNAs) between tumor tissues and normal tissues. Non-cancer-specific expression genes were filtered, and upregulated or downregulated genes were saved. The package ggplot2 3.3.2⁷ was used in RStudio to generate volcano plots of upregulated or downregulated genes. Meanwhile, the VennDiagram R package⁸ was used for the Venn diagram of the datasets.

Construction of the Coexpression Network

The WGCNA package was used to analyze the data as described previously (Langfelder and Horvath, 2008). For the different conditions of tumor tissues, some samples may deviate from the actual situation. These tissue samples will affect the accuracy of the results. Therefore, these samples were removed before sample analysis. The function of goodSamplesGenes was used to verify if there were many missing values for each sample and to remove these samples from the total data. Then, the function hclust was used for cluster analysis of tissue samples. Appropriate soft thresholds need to be selected to build the network when the WGCNA method is performed. The soft threshold value should not only reach a scale-free fitting index of more than 0.9 but also requires greater mean connectivity of the network. Thus, 4 was selected for the soft threshold value. A good soft threshold is selected based on this rule. At the same time, the size of the gene module will be divided, and at least 30 genes will be identified for one module. If the correlation coefficient is greater than 0.75 for two modules, then these two modules will be merged into one module. To enhance the productivity of the modules, a cutoff (< 0.25) was selected to combine similar modules. After these modules were divided, the sample traits were associated with each module such that the most important associated modules for the traits were searched. Cytoscape 3.8.1 was used for the visualization of gene coexpression.

Gene Ontology and Pathway Enrichment Analysis

The DAVID online tool was applied for the GO (Gene Ontology) and KEGG analysis and functional annotation of DEmRNAs. The DAVID⁹ database website was used for annotation and GO enrichment. KEGG pathways and GO terms enriched adopt adjusted $P\text{-value} \leq 0.05$. WebGestalt¹⁰ is used for enrichment analysis that supports three well-established and complementary methods for enrichment analysis. The genes of the network were analyzed by the WebGestalt website. The R package of GPlot (version 1.0.2) was used for the GO analysis of the DEmRNAs of the five cancers.

⁷<https://cloud.r-project.org/package=ggplot2>

⁸<https://cran.r-project.org/web/packages/VennDiagram/>

⁹<http://david.abcc.ncifcrf.gov/>

¹⁰<http://www.webgestalt.org/>

Hub Gene Identification and Validation

To identify the gene connectivity, Pearson's correlation was used for the test. In general, hub genes existed in modules closely linked to traits ($\text{cor.geneTraitSignificance} > 0.2$), the modules with hub genes had high connectivity ($\text{cor.geneModuleMembership} > 0.8$), and the packages of cytoHubba and MCODE in Cytoscape 3.8.1 were used to search the hub genes.

RESULTS

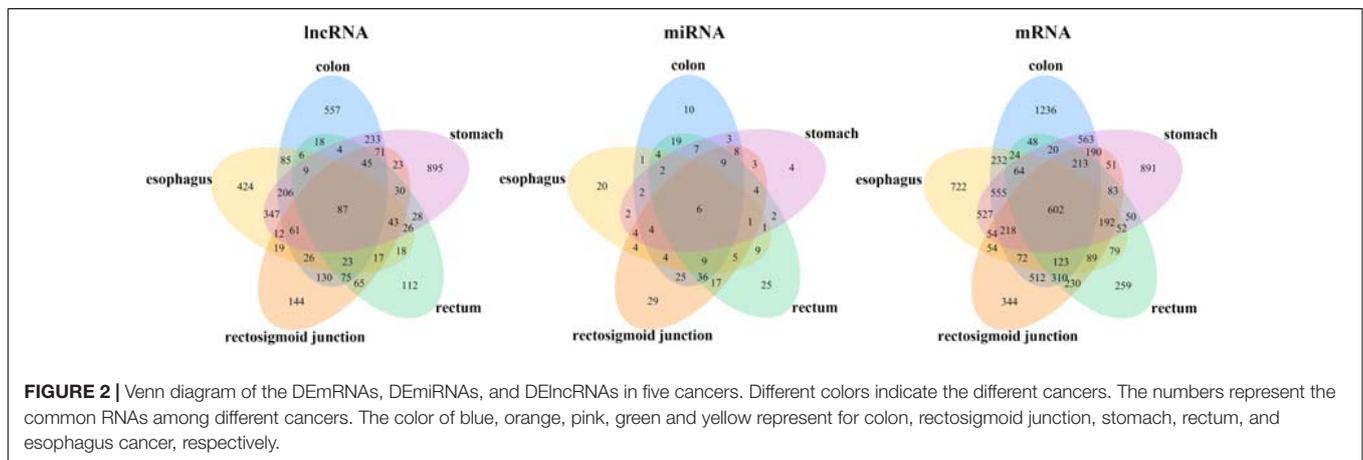
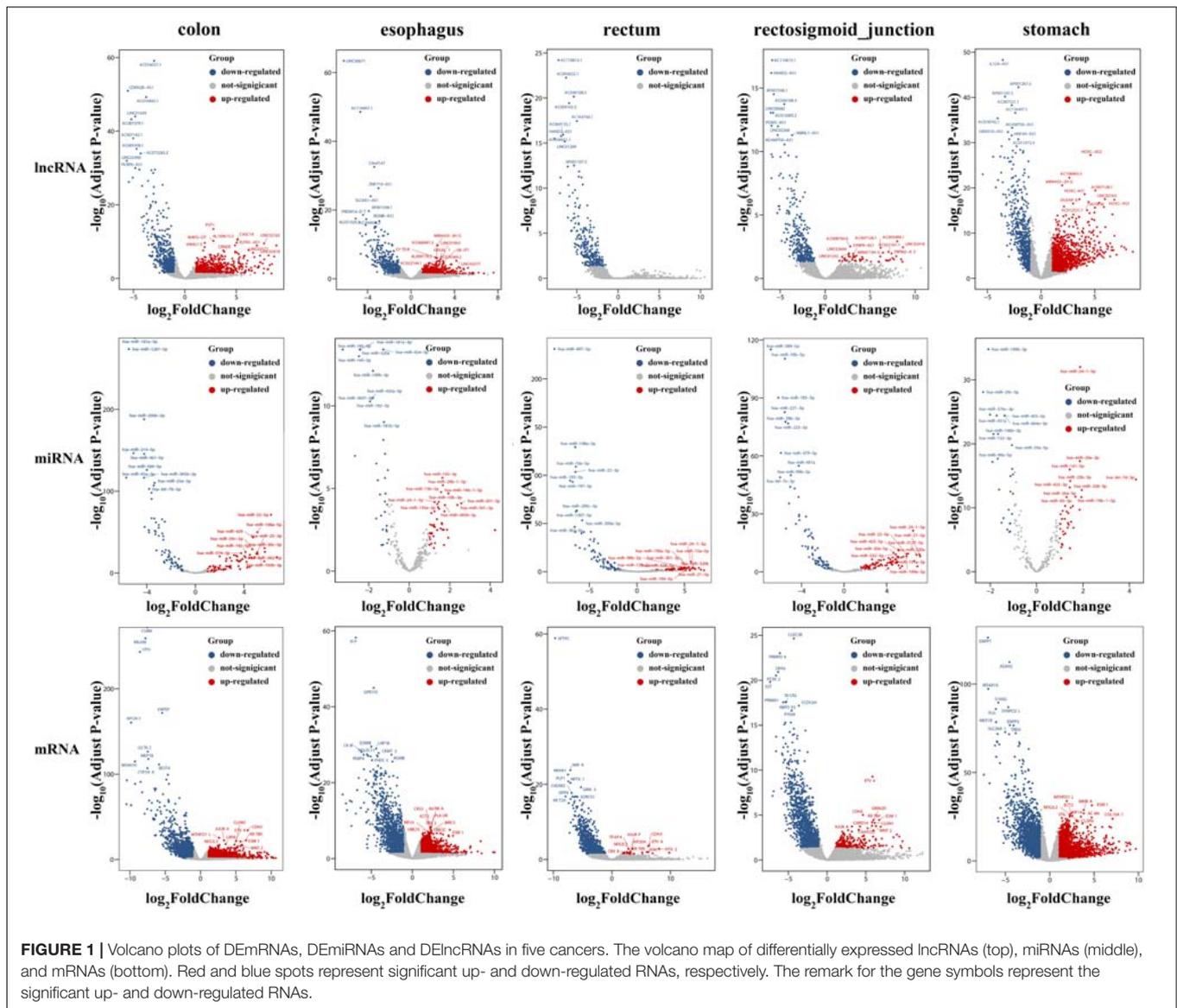
Identification and Validation of Differentially Expressed mRNAs, miRNAs, and lncRNAs

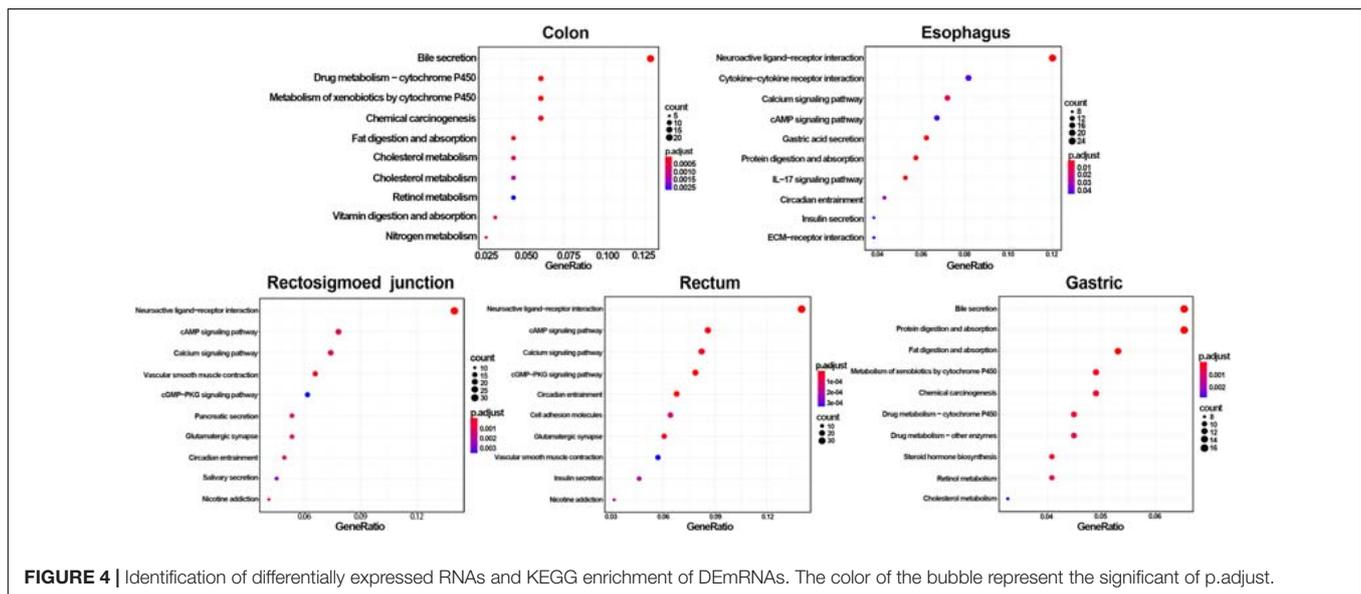
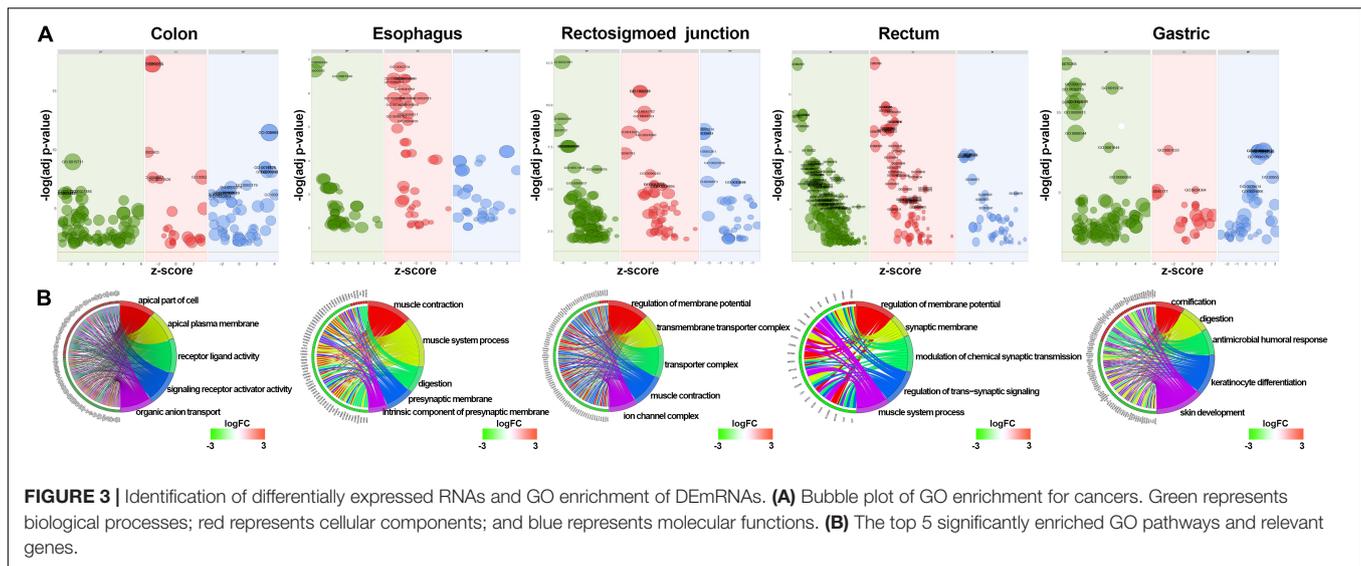
To identify the differentially expressed mRNAs, lncRNAs and miRNAs, the standard of $\log_2(\text{fold change}) \geq 1$ and $q\text{-value} < 0.05$ was used, and the EdgeR package of R (Law et al., 2016; Maza, 2016; Chen et al., 2017) was adopted to calculate separately for the normal samples and cancer samples of five cancers. For colon cancer, 4982 DEmRNAs, 149 DEmiRNAs and 1636 DELncRNAs were identified. There were 1409 differentially expressed genes in lncRNAs, 78 DEmiRNAs and 3659 differentially expressed mRNAs genes found between healthy and cancer-treated samples for esophageal cancer. A total of 3337 DEmRNAs, 168 DEmiRNAs and 871 DELncRNAs were identified between the rectosigmoid junction cancer tissues and matched normal control tissues; in addition, 2438 DEmRNAs, 156 DEmiRNAs and 607 DELncRNAs were identified for rectal cancer, and 4325 DEmRNAs, 62 DEmiRNAs and 2120 DELncRNAs were identified for gastric cancer compared with normal control samples. The distributions of differentially expressed DEmRNAs, DEmiRNAs and DELncRNAs were identified through volcano plots (Figure 1). A number of differences were found, and 87 DELncRNAs, 6 DEmiRNAs and 602 DEmRNAs were shared among the five cancers as shown in Figure 2.

GO and KEGG Pathway Analysis of DEmRNAs

GO and KEGG methods were adopted to investigate the annotation of DEmRNAs for the five cancers (Figures 3, 4). Standards of $\log_2(\text{fold change}) \geq 3$ and $q\text{-value} < 0.01$ were used for the analysis.

A total of 766 DEmRNAs were primarily enriched in biological processes (BP) such as organic anion transport for colon cancer (GO:0015711) and cell components (CC) such as the apical part of the cell (GO:0045177) and the apical plasma membrane (GO:0016324) as well as ligand activity (GO:0048018) and signaling receptor activator activity (GO:0030546) of molecular function (MF) receptor for colon cancer (Figure 3). In addition, KEGG enrichment analysis revealed that the significant pathways were "bile secretion," "drug metabolism - cytochrome P450," "chemical carcinogenesis," "metabolism of xenobiotics by cytochrome P450" and "retinol metabolism," which are associated with the progression of colon cancer (Figure 4).





To investigate the enrichment in esophageal cancer, 448 DEmRNAs were used and enriched in muscle contraction (GO:0006936), the muscle system process (GO:0003012), digestion (GO:0007586), the presynaptic membrane (GO:0042734), the intrinsic component of the presynaptic membrane (GO:0098889) of biological processes (BP) and cell components (CC) (Figure 3). In addition, “gastric acid secretion,” “neuroactive ligand-receptor interaction,” “protein digestion and absorption,” “IL-17 signaling pathway,” and “calcium signaling pathway” were the top 5 pathways of the KEGG analysis (Figure 4).

Rectosigmoid junction cancer was also analyzed by GO and KEGG enrichment; 623 DEmRNAs were screened out, and the top 5 GO enrichments were the regulation of membrane potential (GO:0042391), the transmembrane transporter complex (GO:1902495), the transporter complex

(GO:1990351), muscle contraction (GO:0006936), and the ion channel complex (GO:0034702) of biological processes (BP) and cell components (CC) (Figure 3). The top 5 KEGG pathways were “neuroactive ligand-receptor interaction,” “nicotine addiction,” “vascular smooth muscle contraction,” “pancreatic secretion,” and “cAMP signaling pathway” (Figure 4).

Rectal cancer was also analyzed by GO and KEGG enrichment; 691 DEmRNAs were screened out, and the top 5 GO enrichments were the regulation of membrane potential (GO:0042391), the regulation of transsynaptic signaling (GO:0099177), the synaptic membrane (GO:0097060), the modulation of chemical synaptic transmission (GO:0050804), and the muscle system process (GO:0003012) of biological processes (BP) and cell components (CC) (Figure 3). The top 5 KEGG pathways were “neuroactive ligand-receptor

interaction,” “circadian entrainment,” “cGMP-PKG signaling pathway,” “glutamatergic synapse” and “cAMP signaling pathway” (Figure 4).

Gastric cancer was also subjected to GO and KEGG enrichment analyses. A total of 637 DEMRNAs were screened out, and the top 5 GO enrichments were cornification (GO:0070268), digestion (GO:0007586), antimicrobial humoral response (GO:0019730) (Ji et al., 2019; Nadia and Ramana, 2020), keratinocyte differentiation (GO:0030216), and skin development (GO:0043588) of biological processes (BP) (Figure 3). The top 5 KEGG pathways were “Protein digestion and absorption,” “Bile secretion,” “Fat digestion and absorption,” “Chemical carcinogenesis,” and “Metabolism of xenobiotics by cytochrome P450” (Figure 4).

Construction of the WGCNA Network

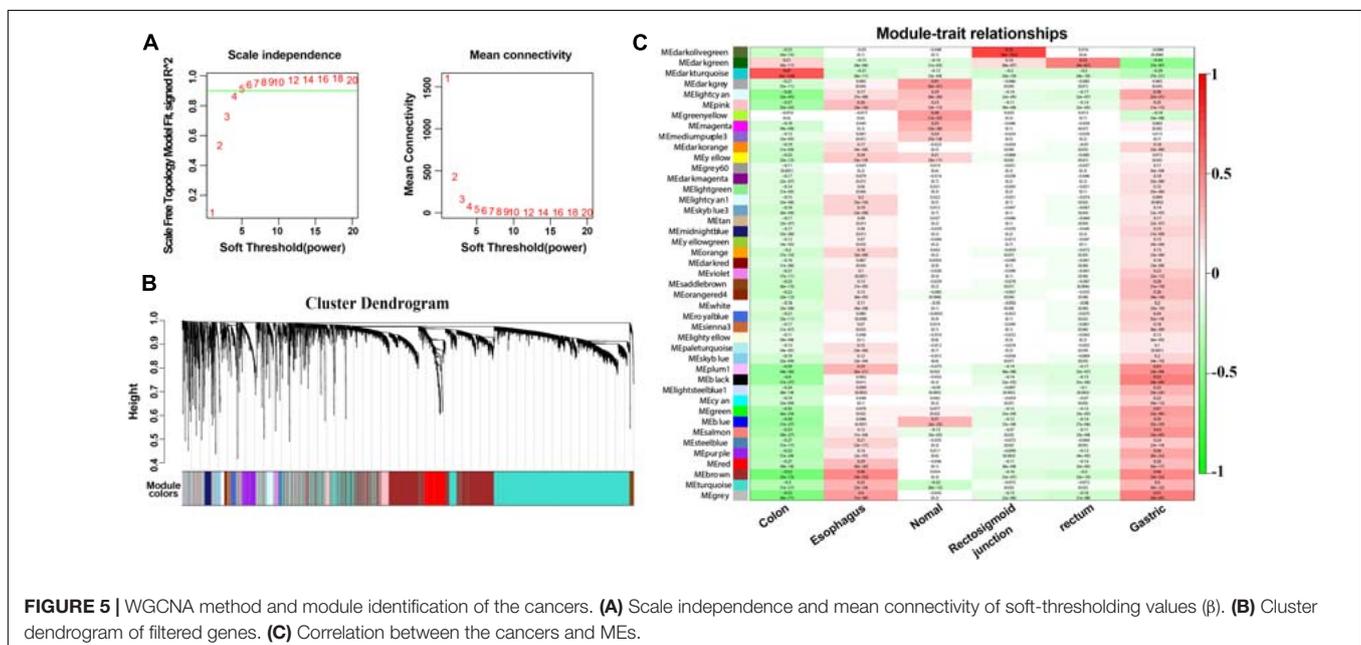
To explore further the key biological genes related to cancers, the WGCNA method was applied to select the DEMRNAs, DEMiRNAs and DELncRNAs. For further analysis, a soft threshold ($\beta = 4$) was adopted to guarantee a scale-free network with high scale independence and low mean connectivity (near 0) (Figure 5A). DEGs from the five cancers were divided into several modules by cluster analysis (Figure 5B). Approximately 43 modules were generated for the five cancers. The module trait relationship is shown in Figure 5C. The dark green module related to the colon, rectosigmoid junction and rectal tumors were the deepest ($\text{cor} = 0.21, 0.16, \text{ and } 0.52$ $P = 4\text{E-}11, 6\text{e-}07, \text{ and } 4\text{e-}67$). For the other two cancers, the esophageal and gastric cancers, the brown module was chosen according to the correlation ($\text{cor} = 0.46$ and 0.46 $P = 4\text{E-}52, 3\text{e-}52$). Dark green and brown modules were selected for further analysis (Figure 5C). The high correlation and high P-value indicated

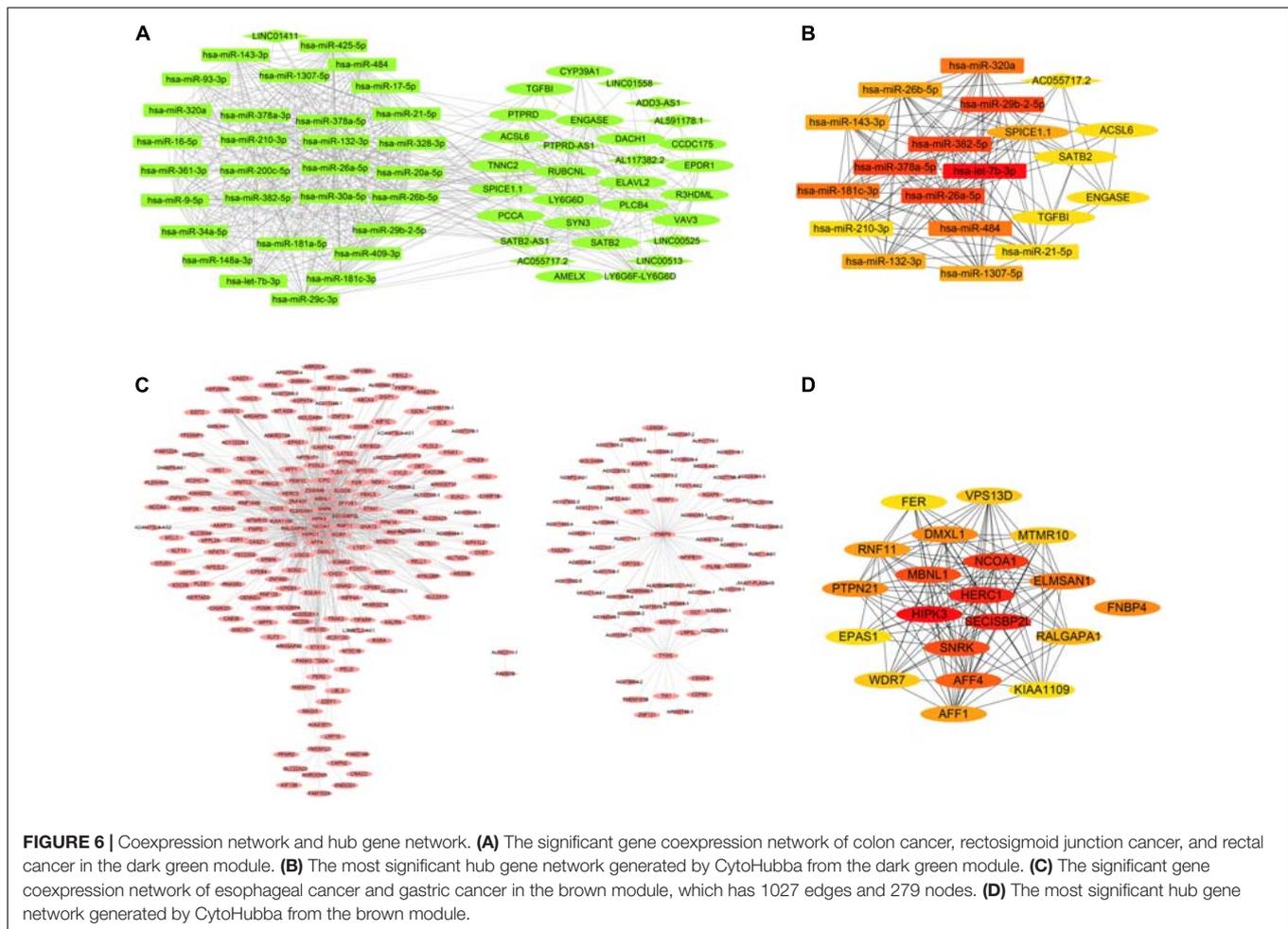
that these modules are suitable for hub gene identification in these cancers.

Identification of the Hub Genes in the Five Cancers

The dark green module was selected for the edges signifying the correlations in colon cancer, rectosigmoid junction cancer and rectal cancer with the filter criterion of a weight value greater than 0.02 by the WGCNA algorithm. A total of 629 edges and 61 nodes were obtained and input into Cytoscape 3.8.1 (Figure 6A). A network was generated by Cytoscape software, and the hub genes were obtained from the network by applying CytoHubba. The most significant hub gene network was discovered by CytoHubba, as shown in Figure 6B. The identified hub miRNAs included *hsa-let-7b-3p*, *hsa-miR-378a-5p*, *hsa-miR-26a-5p*, *hsa-miR-382-5p*, and *hsa-miR-29b-2-5p*. In addition, the brown module was selected for the edges signifying the correlations in esophageal cancer and gastric cancer with the filter criterion of a weight value greater than 0.12 by the WGCNA algorithm. A total of 1027 edges and 279 nodes were obtained and input into Cytoscape 3.8.1 (Figure 6C). The most significant hub genes were discovered by CytoHubba as shown in Figure 6D. The identified hub genes were *SECISBP2 L*, *NCOA1*, *HERC1*, *HIPK3*, and *MBNL1*.

We discover that most of the genes in dark green module that were significantly correlated with colon cancer, rectosigmoid junction cancer and rectal cancer were miRNAs. This dominance of miRNAs in dark green module suggests the possibility that lncRNAs play a significant role in these cancers through regulation of coding genes in key pathways. miRNA can act by guiding histone modifiers and chromatin modifiers to regulate transcription and play crucial roles in cell differentiation that ultimately determine cell fate.





Functional Annotation of the Module of Interest and the Network Genes

GO and KEGG analyses of the two modules of dark green and brown with 61 and 279 genes, respectively, in WebGestalt are shown in **Figure 7**. These were involved in molecular and functional biological processes and cellular components. For the dark green module, the genes were associated with metabolic process, membrane and protein binding, membrane among others (**Figure 7A**), and the genes in the brown module were enriched in biological regulation, membrane, and protein binding among others (**Figure 7B**). Analysis of KEGG pathways for the dark green module, the calcium signaling pathway, fatty acid biosynthesis and the chemokine signaling pathway were significant among all the pathways (**Figure 7C**), and pathways in cancer, insulin resistance and renal cell carcinoma were significant pathways in the brown module (**Figure 7D**).

DISCUSSION

Cancer of the digestive system is a common cancer worldwide. Recently, RNAs have been reported to be related to cancers. The WGCNA method was applied to analyze the related modules

with different tumors and to identify the important genes of these tumors.

ESCC remains a serious burden on the health system worldwide, and some research has suggested that RNAs are associated with ESCC. For example, 2,046 circRNAs were frequently altered in ESCC tissues (Song et al., 2019). Several induced circRNAs were identified in radioresistant ESCC cells compared to normal ESCC cells (Su et al., 2016). Eight lncRNAs could be practical and reliable prognostic tools for esophageal cancers (Li W. et al., 2020). Additionally, the special functions of RNAs have been clearly demonstrated. For the analysis in gastric cancer (GC), 9 miRNAs, 41 lncRNAs, and 10 mRNAs were suggested and significant for GC (Pan et al., 2019). In addition, 15 lncRNAs were identified to be significantly related to the clinical features of colon cancer (Qian et al., 2019). In rectal cancer, COL1A1 and MZB1, as the key genes of rectal cancer, can interact with other genes correlated with shorter survival for patients (Wu et al., 2019). However, these studies presented research on only one cancer for the digestion system. Therefore, there is still an urgent need to identify the hub genes in digestive system cancers.

In this study, data from the TCGA website were adopted, and DE mRNAs, 168 DE miRNAs, and 871 DE lncRNAs were identified for the different tumors with the RNA-sequencing

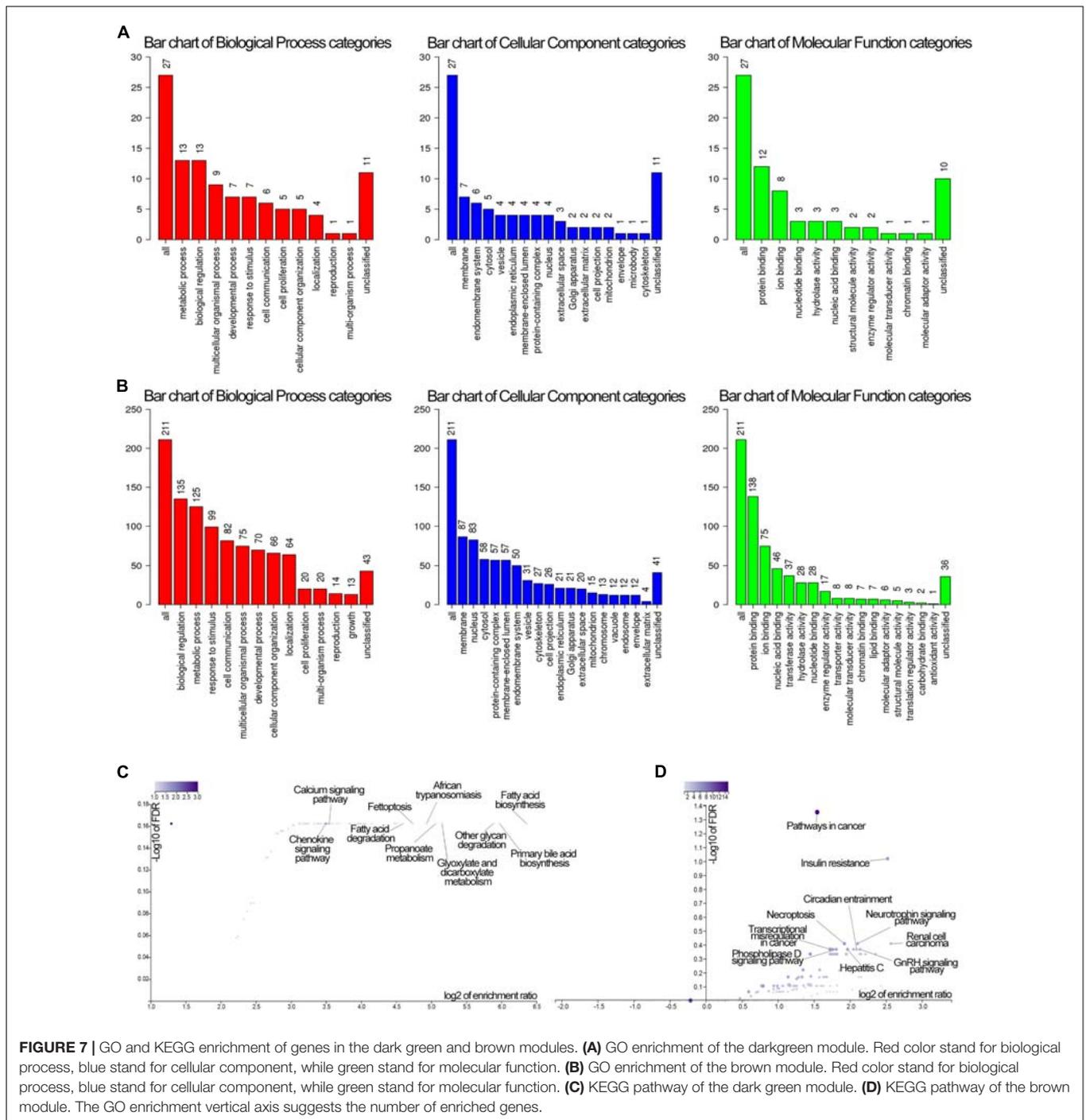


FIGURE 7 | GO and KEGG enrichment of genes in the dark green and brown modules. **(A)** GO enrichment of the darkgreen module. Red color stand for biological process, blue stand for cellular component, while green stand for molecular function. **(B)** GO enrichment of the brown module. Red color stand for biological process, blue stand for cellular component, while green stand for molecular function. **(C)** KEGG pathway of the dark green module. **(D)** KEGG pathway of the brown module. The GO enrichment vertical axis suggests the number of enriched genes.

method and integrated bioinformatics analysis. Based on the WGCNA algorithm, the modules associated with different tumors were generated under the appropriate thresholds. Therefore, a network related to the different cancers was constructed in this study. Based on the network analysis, two modules, dark green and brown, were selected for further research. CytoHubba from Cytoscape 3.8.1 was used for this analysis, and several hub genes were chosen for the prediction.

Twenty hub genes were associated with the staging of esophageal cancer and gastric cancer, such as *SECISBP2 L*, *NCOA1*, *HERC1*, *HIPK3*, and *MBNL1*. *SECISBP2 L*, which encodes the SECIS binding protein 2 (SBP2), is a selenoprotein that exists in yeast, fungi and higher plants (Donovan and Copeland, 2009). However, the molecular mechanism of this protein has not been elucidated. *NCOA1* is nuclear receptor coactivator 1 (also known as SRC-1) belonging to the SRC gene family (Oñate et al., 1995; Xu et al., 2009). The abnormal

expression level of SRCs has triggered cancers, such as breast, endometrial and ovarian cancers. The *SRC-3* and *SRC-1* genes have high expression levels in breast cancer. *SRC-1* regulates polyoma enhancer activator 3 (PEA3) and promotes breast-to-lung metastasis via transcription factors (Redmond et al., 2009; Walsh et al., 2012; Qin et al., 2015). HERC1 is a HECT and RLD domain containing E3 ubiquitin protein ligase family member 1 in humans. E3 ubiquitin ligases play a critical role in catalyzing ubiquitin transfer from E2 enzymes to the substrate in the ubiquitylation system. Similar to most E3 ubiquitin ligases, HERC1 has been shown to be a potential target for cancer therapy (Schneider et al., 2018; Dong et al., 2020). Homeodomain-interacting protein kinase 3 (*HIPK3*) is encoded by the *HIPK3* gene in humans. In NSCLC tissues, low expression of *HIPK3* was associated with poor survival rates. Therefore, *HIPK3* is considered a valuable biomarker for the survival of NSCLC patients (Liu et al., 2018). Muscleblind-like splicing regulator 1 (*MBNL1*), encoded by the *MBNL1* gene, is an RNA splicing protein and essential for MLL-rearranged leukemia cell growth (Itskovich et al., 2020). The isoforms of the *MBNL1* protein affect cancer development and are targets for drug development (Li Z. et al., 2020). The results of the network indicated that these genes are also related to esophageal cancer and gastric cancer, and further research is necessary to determine the molecular mechanisms of these genes.

We eventually obtained 20 miRNAs associated with the staging of esophageal cancer and gastric cancer (Jiang et al., 2009), including *hsa-let-7b-3p*, *hsa-miR-378a-5p*, *hsa-miR-26a-5p*, *hsa-miR-382-5p*, and *hsa-miR-29b-2-5p*. Previous studies indicate that gastric cancer stem-like cells are derived from MKN-45- by the expression change of *hsa-let-7b-3p* (Salehi et al.,

2018). Moreover, *microRNA-378a-5p* increased the expression level in melanoma cells and is a novel positive regulator of melanoma progression (Tang et al., 2018; Tupone et al., 2020). The expression of *hsa-miR-26a-5p* was downregulated in breast cancer tissues and appeared to have a poor prognosis for patients (Huang et al., 2019). However, overexpression of *hsa-miR-382-5p* increased oral squamous cell carcinoma cell invasion and migration (Sun et al., 2019). These studies suggested that miRNAs are also biomarkers for various cancers; however, the prediction of the interaction network indicated that these miRNAs are also related to colon cancer, rectosigmoid junction cancer and rectal cancer (Jiang et al., 2013).

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>.

AUTHOR CONTRIBUTIONS

LX and LZ designed the research. ZC and ZS performed the research. ZC, ZS, ZZ, and DZ analyzed the data. ZC wrote the manuscript. All authors read and approved the manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 62002244, 62002242, and 62001311).

REFERENCES

- Abbas, G., and Krasna, M. (2017). Overview of esophageal cancer. *Annals of Cardiothoracic Surgery* 6, 131–136. doi: 10.21037/acs.2017.03.03
- Amer, A., Nagah, A., Tian, T., and Zhang, X. (2020). Mutation Mechanisms of Breast Cancer among the Female Population in China. *Current Bioinformatics* 15, 253–259. doi: 10.2174/1574893615666191220141548
- Amin, M. B., Greene, F. L., Edge, S. B., Compton, C. C., Gershenwald, J. E., Brookland, R. K., et al. (2017). The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging. *CA: A Cancer Journal for Clinicians* 67, 93–99. doi: 10.3322/caac.21388
- Bai, Y., Long, J., Liu, Z., Lin, J., Huang, H., Wang, D., et al. (2019). Comprehensive analysis of a ceRNA network reveals potential prognostic cytoplasmic lncRNAs involved in HCC progression. *Journal of Cellular Physiology* 234, 18837–18848. doi: 10.1002/jcp.28522
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68, 394–424. doi: 10.3322/caac.21492
- Callahan, K., Ponce, C., Cross, C., Sy, F., and Pinheiro, P. (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.0221337
- Cao, Z., Pan, X., Yang, Y., Huang, Y., and Shen, H. B. (2018). The lncLocator: a subcellular localization predictor for long non-coding RNAs based on a stacked ensemble classifier. *Bioinformatics* 34, 2185–2194. doi: 10.1093/bioinformatics/bty085
- Chen, Y., Pal, B., Visvader, J. E., and Smyth, G. K. (2017). Differential methylation analysis of reduced representation bisulfite sequencing experiments using edgeR. *F1000Res* 6, 2055. doi: 10.12688/f1000research.13196.2
- Cheng, L. (2019). Computational and Biological Methods for Gene Therapy. *Current Gene Therapy* 19, 210–210.
- Cheng, L., Hu, Y., Sun, J., Zhou, M., and Jiang, Q. (2018). DincRNA: a comprehensive web-based bioinformatics toolkit for exploring disease associations and ncRNA function. *Bioinformatics* 34, 1953–1956. doi: 10.1093/bioinformatics/bty002
- Cheng, L., Wang, P., Tian, R., Wang, S., Guo, Q., Luo, M., et al. (2019a). LncRNA2Target v2.0: a comprehensive database for target genes of lncRNAs in human and mouse. *Nucleic Acids Res* 47, D140–D144. doi: 10.1093/nar/gky1051
- Cheng, L., Zhao, H., Wang, P., Zhou, W., Luo, M., Li, T., et al. (2019b). Computational Methods for Identifying Similar Diseases. *Molecular therapy. Nucleic acids* 18, 590–604.
- 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
- Doja, M. N., Kaur, I., and Ahmad, T. (2020). Current State of the Art for Survival Prediction in Cancer Using Data Mining Techniques. *Current Bioinformatics* 15, 174–186. doi: 10.2174/1574893614666190902152142
- Dong, Y.-M., Bi, J.-H., He, Q.-E., and Song, K. (2020). ESDA: An Improved Approach to Accurately Identify Human snoRNAs for

- Precision Cancer Therapy. *Current Bioinformatics* 15, 34–40. doi: 10.2174/1574893614666190424162230
- Donovan, J., and Copeland, P. R. (2009). Evolutionary history of selenocysteine incorporation from the perspective of SECIS binding proteins. *BMC Evolutionary Biology* 9:229. doi: 10.1186/1471-2148-9-229
- Ferlay, J., Shin, H.-R., Bray, F., Forman, D., and Parkin, D. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN. *International journal of cancer. Journal international du cancer* 127, 2893–2917. doi: 10.1002/ijc.25516
- 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
- Fitzmaurice, C., Abate, D., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdel-Rahman, O., et al. (2019). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 5, 1749–1768. doi: 10.1001/jamaoncol.2019.2996
- Fitzmaurice, C., Akinyemiju, T. F., Al Lami, F. H., Alam, T., Alizadeh-Navaei, R., Allen, C., et al. (2018). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2016: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 4, 1553–1568. doi: 10.1001/jamaoncol.2018.2706
- Fitzmaurice, C., Allen, C., Barber, R. M., Barregard, L., Bhutta, Z. A., Brenner, H., et al. (2017). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* 3, 524–548. doi: 10.1001/jamaoncol.2016.5688
- Ghosh, A., and Yan, H. (2020). Stability Analysis at Key Positions of EGFR Related to Non-small Cell Lung Cancer. *Current Bioinformatics* 15, 260–267. doi: 10.2174/1574893614666191212112026
- Hong, Z., Zeng, X., Wei, L., and Liu, X. J. B. (2020). Identifying enhancer-promoter interactions with neural network based on pre-trained DNA vectors and attention mechanism. *Bioinformatics* 36, 1037–1043.
- Huang, Z.-M., Ge, H.-F., Yang, C.-C., Cai, Y., Chen, Z., Tian, W.-Z., et al. (2019). MicroRNA-26a-5p inhibits breast cancer cell growth by suppressing RNF6 expression. *The Kaohsiung Journal of Medical Sciences* 35, 467–473. doi: 10.1002/kjm2.12085
- Iliopoulos, A. C., Beis, G., Apostolou, P., and Papisotiriou, I. (2020). Complex Networks, Gene Expression and Cancer Complexity: A Brief Review of Methodology and Applications. *Current Bioinformatics* 15, 629–655. doi: 10.2174/1574893614666191017093504
- Itskovich, S. S., Guranathan, A., Clark, J., Burwinkel, M., Wunderlich, M., Berger, M. R., et al. (2020). MBNL1 regulates essential alternative RNA splicing patterns in MLL-rearranged leukemia. *Nature Communications* 11, 2369. doi: 10.1038/s41467-020-15733-8
- Ji, J., Tang, J., Xia, K.-J., and Jiang, R. (2019). LncRNA in Tumorigenesis Microenvironment. *Current Bioinformatics* 14, 640–641. doi: 10.2174/157489361407190917161654
- Jiang, L., Ding, Y., Tang, J., and Guo, F. (2018). MDA-SKF: Similarity Kernel Fusion for Accurately Discovering miRNA-Disease Association. *Frontiers in Genetics* 9:1–13. doi: 10.3389/fgene.2018.00618
- Jiang, L., Xiao, Y., Ding, Y., Tang, J., and Guo, F. (2019). FKL-Spa-LapRLS: an accurate method for identifying human microRNA-disease association. *BMC Genomics* 19:11–25.
- Jiang, Q., Hao, Y., Wang, G., Juan, L., Zhang, T., Teng, M., et al. (2010). Prioritization of disease microRNAs through a human phenome-microRNAome network. *BMC Syst Biol* 4(Suppl. 1):S2. doi: 10.1186/1752-0509-4-S1-S2
- Jiang, Q., Wang, G., Jin, S., Li, Y., and Wang, Y. (2013). Predicting human microRNA-disease associations based on support vector machine. *Int J Data Min Bioinform* 8, 282–293.
- Jiang, Q., Wang, J., Wang, Y., Ma, R., Wu, X., and Li, Y. (2014). TF2LncRNA: identifying common transcription factors for a list of lncRNA genes from ChIP-Seq data. *Biomed Res Int* 2014, 317642. doi: 10.1155/2014/317642
- Jiang, Q., Wang, Y., Hao, Y., Juan, L., Teng, M., Zhang, X., et al. (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res* 37, D98–D104. doi: 10.1093/nar/gkn714
- Jin, S., Zeng, X., Fang, J., Lin, J., Chan, S. Y., Erzurum, S. C., et al. (2019). A network-based approach to uncover microRNA-mediated disease comorbidities and potential pathobiological implications. *NPJ systems biology and applications* 5, 1–11.
- Jin, S., Zeng, X., Xia, F., Huang, W., and Liu, X. (2020). Application of deep learning methods in biological networks. *Briefings in Bioinformatics* 2, bbaa043. doi: 10.1093/bib/bbaa043
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559
- Langfelder, P., Mischel, P. S., and Horvath, S. (2013). When Is Hub Gene Selection Better than Standard Meta-Analysis? *PLoS One* 8:e61505. doi: 10.1371/journal.pone.0061505
- Law, C., Alhamdoosh, M., Su, S., Smyth, G., and Ritchie, M. (2016). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research* 5, 1408. doi: 10.12688/f1000research.9005.1
- Li, J., Chang, M., Gao, Q., Song, X., and Gao, Z. (2020). Lung Cancer Classification and Gene Selection by Combining Affinity Propagation Clustering and Sparse Group Lasso. *Current Bioinformatics* 15, 703–712. doi: 10.2174/1574893614666191017103557
- Li, W., Liu, J., and Zhao, H. (2020). Identification of a nomogram based on long non-coding RNA to improve prognosis prediction of esophageal squamous cell carcinoma. *Aging* 12, 1512–1526. doi: 10.18632/aging.102697
- Li, Z., Zhang, T., Lei, H., Wei, L., Liu, Y., Shi, Y., et al. (2020). Research on Gastric Cancer's Drug-resistant Gene Regulatory Network Model. *Current Bioinformatics* 15, 225–234. doi: 10.2174/1574893614666190722102557
- Li, S., Jiang, L., Tang, J., Gao, N., and Guo, F. (2020). Kernel Fusion Method for Detecting Cancer Subtypes via Selecting Relevant Expression Data. *Front. Genet.* 11:979. doi: 10.3389/fgene.2020.00979
- Liang, C., Changlu, Q., He, Z., Tongze, F., and Xue, Z. (2019). gutMDisorder: a comprehensive database for dysbiosis of the gut microbiota in disorders and interventions. *Nucleic Acids Research* 48, 7603.
- Lin, M., Li, X., Guo, H., Ji, F., Ye, L., Ma, X., et al. (2019). Identification of Bone Metastasis-associated Genes of Gastric Cancer by Genome-wide Transcriptional Profiling. *Current Bioinformatics* 14, 62–69. doi: 10.2174/1574893612666171121154017
- Liu, X., Hong, Z., Liu, J., Lin, Y., Rodríguez-Patón, A., Zou, Q., et al. (2020). Computational methods for identifying the critical nodes in biological networks. *Briefings in Bioinformatics* 21, 486–497.
- Liu, Y., Qian, L., Yang, J., Huang, H., Feng, J., Li, X., et al. (2018). The expression level and prognostic value of HIPK3 among non-small-cell lung cancer patients in China. *OncoTargets and therapy* 11, 7459–7469. doi: 10.2147/OTT.S166878
- Liu, Y., Zeng, X., He, Z., and Zou, Q. (2017). Inferring MicroRNA-Disease Associations by Random Walk on a Heterogeneous Network with Multiple Data Sources. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 14, 905–915. doi: 10.1109/tcbb.2016.2550432
- Long, J., Bai, Y., Yang, X., Lin, J., Yang, X., Wang, D., et al. (2019). Construction and comprehensive analysis of a ceRNA network to reveal potential prognostic biomarkers for hepatocellular carcinoma. *Cancer Cell International* 19, 90. doi: 10.1186/s12935-019-0817-y
- Ma, X., Xi, B., Zhang, Y., Zhu, L., Sui, X., Tian, G., et al. (2020). A Machine Learning-based Diagnosis of Thyroid Cancer Using Thyroid Nodules Ultrasound Images. *Current Bioinformatics* 15, 349–358. doi: 10.2174/1574893614666191017091959
- Maza, E. (2016). In Papyro Comparison of TMM (edgeR), RLE (DESeq2), and MRN Normalization Methods for a Simple Two-Conditions-Without-Replicates RNA-Seq Experimental Design. *Front Genet* 7:164. doi: 10.3389/fgene.2016.00164
- Nadia, and Ramana, J. (2020). The Human OncoBiome Database: A Database of Cancer Microbiome Datasets. *Current Bioinformatics* 15, 472–477. doi: 10.2174/1574893614666190902152727
- Oñate, S. A., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (1995). Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* 270, 1354–1357. doi: 10.1126/science.270.5240.1354

- Ozkan, A., Isgor, S. B., Sengul, G., and Isgor, Y. G. (2019). Benchmarking Classification Models for Cell Viability on Novel Cancer Image Datasets. *Current Bioinformatics* 14, 108–114. doi: 10.2174/1574893614666181120093740
- Pan, H., Guo, C., Pan, J., Guo, D., Song, S., Zhou, Y., et al. (2019). Construction of a Competitive Endogenous RNA Network and Identification of Potential Regulatory Axis in Gastric Cancer. *Frontiers in Oncology* 9:912. doi: 10.3389/fonc.2019.00912
- 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
- Qian, W., Feng, Y., Li, J., Peng, W., Gu, Q., Zhang, Z., et al. (2019). Construction of ceRNA networks reveals differences between distal and proximal colon cancers. *Oncol Rep* 41, 3027–3040. doi: 10.3892/or.2019.7083
- Qin, L., Xu, Y., Xu, Y., Ma, G., Liao, L., Wu, Y., et al. (2015). NCOA1 promotes angiogenesis in breast tumors by simultaneously enhancing both HIF1 α - and AP-1-mediated VEGFa transcription. *Oncotarget* 6, 23890–23904. doi: 10.18632/oncotarget.4341
- Rahman, R., Schmalz, 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
- Redmond, A. M., Bane, F. T., Stafford, A. T., McIlroy, M., Dillon, M. F., Crotty, T. B., et al. (2009). Coassociation of estrogen receptor and p160 proteins predicts resistance to endocrine treatment; SRC-1 is an independent predictor of breast cancer recurrence. *Clin Cancer Res* 15, 2098–2106. doi: 10.1158/1078-0432.Ccr-08-1649
- Rustgi, A., and El-Serag, H. (2014). Esophageal Carcinoma. *The New England journal of medicine* 371, 2499–2509. doi: 10.1056/NEJMr1314530
- Rustgi, A., and El-Serag, H. B. (2015). Esophageal carcinoma. *N Engl J Med* 372, 1472–1473. doi: 10.1056/NEJMc1500692
- Salehi, Z., Akrami, H., and Sisakhtnezhad, S. (2018). The effect of placenta growth factor knockdown on hsa-miR-22-3p, hsa-let-7b-3p, hsa-miR-451b, and hsa-miR-4290 expressions in MKN-45-derived gastric cancer stem-like cells. *Middle East Journal of Cancer* 9, 113–122.
- Sardina, D. S., Alaimo, S., Ferro, A., Pulvirenti, A., and Giugno, R. (2017). A novel computational method for inferring competing endogenous interactions. *Brief Bioinform* 18, 1071–1081. doi: 10.1093/bib/bbw084
- Schneider, T., Martinez-Martinez, A., Cubillos-Rojas, M., Bartrons, R., Ventura, F., and Rosa, J. L. (2018). The E3 ubiquitin ligase HERC1 controls the ERK signaling pathway targeting C-RAF for degradation. *Oncotarget* 9, 31531–31548. doi: 10.18632/oncotarget.25847
- Shen, C., Jiang, L., Ding, Y., Tang, J., and Guo, F. (2019). LPI-KTASLP: Prediction of lncRNA-Protein Interaction by Semi-Supervised Link Learning with Multivariate Information. *IEEE Access* 7, 13486–13496.
- Slattery, M. L., Wolff, R. K., Herrick, J. S., Caan, B. J., and Potter, J. D. (2007). IL6 genotypes and colon and rectal cancer. *Cancer Causes & Control* 18, 1095–1105. doi: 10.1007/s10552-007-9049-x
- Song, J., Lu, Y., Sun, W., Han, M., Zhang, Y., and Zhang, J. (2019). Changing expression profiles of lncRNAs, circRNAs and mRNAs in esophageal squamous carcinoma. *Oncology Letters* 18, 5363–5373. doi: 10.3892/ol.2019.10880
- Su, H., Lin, F., Deng, X., Shen, L., Fang, Y., Fei, Z., et al. (2016). Profiling and bioinformatics analyses reveal differential circular RNA expression in radioresistant esophageal cancer cells. *Journal of translational medicine* 14, 225–225. doi: 10.1186/s12967-016-0977-7
- Sun, L.-P., Xu, K., Cui, J., Yuan, D.-Y., Zou, B., Li, J., et al. (2019). Cancer-associated fibroblast-derived exosomal miR-382-5p promotes the migration and invasion of oral squamous cell carcinoma. *Oncology reports* 42, 1319–1328. doi: 10.3892/or.2019.7255
- Tang, W., Wan, S., Yang, Z., Teschendorff, A. E., and Zou, Q. (2018). Tumor origin detection with tissue-specific miRNA and DNA methylation markers. *Bioinformatics* 34, 398–406. doi: 10.1093/bioinformatics/btx622
- Tay, Y., Rinn, J., and Pandolfi, P. P. (2014). The multilayered complexity of ceRNA crosstalk and competition. *Nature* 505, 344–352. doi: 10.1038/nature12986
- Thomson, D. W., and Dinger, M. E. (2016). Endogenous microRNA sponges: evidence and controversy. *Nature Reviews Genetics* 17, 272–283. doi: 10.1038/nrg.2016.20
- Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. *CA Cancer J Clin* 65, 87–108. doi: 10.3322/caac.21262
- Tupone, M. G., D'Aguanno, S., Di Martile, M., Valentini, E., Desideri, M., Trisciuglio, D., et al. (2020). microRNA-378a-5p is a novel positive regulator of melanoma progression. *Oncogenesis* 9, 22. doi: 10.1038/s41389-020-0203-6
- Walsh, C. A., Qin, L., Tien, J. C.-Y., Young, L. S., and Xu, J. (2012). The function of steroid receptor coactivator-1 in normal tissues and cancer. *International journal of biological sciences* 8, 470–485. doi: 10.7150/ijbs.4125
- Wang, J., Liu, X., Wu, H., Ni, P., Gu, Z., Qiao, Y., et al. (2010). CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res* 38, 5366–5383. doi: 10.1093/nar/gkq285
- Wei, L., Liao, M., Gao, Y., Ji, R., He, Z., and Zou, Q. (2014). Improved and Promising Identification of Human MicroRNAs by Incorporating a High-Quality Negative Set. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 11, 192–201. doi: 10.1109/tcbb.2013.146
- Wei, L., Su, R., Wang, B., Li, X., Zou, Q., and Gao, X. (2019a). Integration of deep feature representations and handcrafted features to improve the prediction of N-6-methyladenosine sites. *Neurocomputing* 324, 3–9. doi: 10.1016/j.neucom.2018.04.082
- Wei, L., Xing, P., Shi, G., Ji, Z., and Zou, Q. (2019b). Fast Prediction of Protein Methylation Sites Using a Sequence-Based Feature Selection Technique. *Ieee-Acm Transactions on Computational Biology and Bioinformatics* 16, 1264–1273. doi: 10.1109/tcbb.2017.2670558
- Wei, L., Zhou, C., Chen, H., Song, J., and Su, R. (2018). ACPred-FL: a sequence-based predictor using effective feature representation to improve the prediction of anti-cancer peptides. *Bioinformatics* 34, 4007–4016. doi: 10.1093/bioinformatics/bty451
- Wu, W., Yang, Z., Long, F., Luo, L., Deng, Q., Wu, J., et al. (2019). COL1A1 and MZB1 as the hub genes influenced the proliferation, invasion, migration and apoptosis of rectum adenocarcinoma cells by weighted correlation network analysis. *Bioorganic Chemistry* 95, 103457. doi: 10.1016/j.bioorg.2019.103457
- Xu, J., Wu, R. C., and O'Malley, B. W. (2009). Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat Rev Cancer* 9, 615–630. doi: 10.1038/nrc2695
- Yang, L., Gao, H., Wu, K., Zhang, H., Li, C., and Tang, L. (2020). Identification of Cancerlectins By Using Cascade Linear Discriminant Analysis and Optimal g-gap Tripeptide Composition. *Current Bioinformatics* 15, 528–537. doi: 10.2174/1574893614666190730103156
- Zeng, W., Wang, F., Ma, Y., Liang, X., and Chen, P. (2019). Dysfunctional Mechanism of Liver Cancer Mediated by Transcription Factor and Non-coding RNA. *Current Bioinformatics* 14, 100–107. doi: 10.2174/1574893614666181119121916
- Zeng, X., Zhu, S., Liu, X., Zhou, Y., Nussinov, R., and Cheng, F. J. B. (2019). deepDR: a network-based deep learning approach to in silico drug repositioning. *Bioinformatics* 35, 5191–5198.
- Zeng, X., Lin, Y., He, Y., Lv, L., and Min, X. (2020a). Deep collaborative filtering for prediction of disease genes. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 17, 1639–1647.
- Zeng, X., Zhu, S., Lu, W., Liu, Z., Huang, J., Zhou, Y., et al. (2020b). Target identification among known drugs by deep learning from heterogeneous networks. *Chemical Science* 11, 1775–1797.
- Zeng, X., Zhang, X., and Zou, Q. (2016). Integrative approaches for predicting microRNA function and prioritizing disease-related microRNA using biological interaction networks. *Briefings in Bioinformatics* 17, 193–203. doi: 10.1093/bib/bbv033
- Zeng, X. X., Liu, L., Lu, L. Y., and Zou, Q. (2018). Prediction of potential disease-associated microRNAs using structural perturbation method. *Bioinformatics* 34, 2425–2432. doi: 10.1093/bioinformatics/bty112
- Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Research* 46, D754–D761. doi: 10.1093/nar/gkx1098
- Zhang, L., He, Y., Song, H., Wang, X., Lu, N., Sun, L., et al. (2020). Elastic Net Regularized Softmax Regression Methods for Multi-subtype Classification in Cancer. *Current Bioinformatics* 15, 212–224. doi: 10.2174/1574893613666181112141724

- Zhang, Z. M., Tan, J. X., Wang, F., Dao, F. Y., Zhang, Z. Y., and Lin, H. (2020). Early Diagnosis of Hepatocellular Carcinoma Using Machine Learning Method. *Front Bioeng Biotechnol* 8:254. doi: 10.3389/fbioe.2020.00254
- Zhang, X., Zou, Q., Rodriguez-Paton, A., and Zeng, X. (2017). Meta-path methods for prioritizing candidate disease miRNAs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 16, 283–291.
- Zhang, Y., Kou, C., Wang, S., and Zhang, Y. (2019). Genome-wide Differential-based Analysis of the Relationship between DNA Methylation and Gene Expression in Cancer. *Current Bioinformatics* 14, 783–792. doi: 10.2174/1574893614666190424160046
- Zhao, T., Hu, Y., Peng, J., and Cheng, L. (2020). DeepLGP: a novel deep learning method for prioritizing lncRNA target genes. *Bioinformatics* 36, 4466–4472. doi: 10.1093/bioinformatics/btaa428
- Zou, Q., Li, J., Song, L., Zeng, X., and Wang, G. (2016). Similarity computation strategies in the microRNA-disease network: a survey. *Briefings in Functional Genomics* 15, 55–64. doi: 10.1093/bfgp/elv024

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Chen, Shen, Zhang, Zhao, Xu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.