

Received: 2019.01.27
Accepted: 2019.03.22
Published: 2019.04.05

Identification of Key Genes and Circular RNAs in Human Gastric Cancer

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

CEF 1 **Shuhong Hao**
B 2 **Junfeng Lv**
D 3 **Qiwei Yang**
F 4 **Ao Wang**
E 4 **Zhaoyan Li**
AG 5 **Yuchen Guo**
AG 3,4 **Guizhen Zhang**

1 Department of Hematology and Oncology, The Second Hospital of Jilin University, Changchun, Jilin, P.R. China
2 Department of Radiology, The Second Hospital of Jilin University, Changchun, Jilin, P.R. China
3 Medical Research Center, The Second Hospital of Jilin University, Changchun, Jilin, P.R. China
4 Department of Orthopedics, The Second Hospital of Jilin University, Changchun, Jilin, P.R. China
5 Department of Gastrointestinal Surgery, The First Hospital of Jilin University, Changchun, Jilin, P.R. China

Corresponding Authors: Guizhen Zhang, e-mail: zhangguizhenjlu@163.com, Yuchen Guo, e-mail: guoyuchen8688@live.com

Source of support: This work was supported by grants from the Education Department of Jilin Province, P.R.C. (No. JJKH20190059KJ), the Science and Technology Department of Jilin Province, P.R.C. (No. 20190201050JC), the National Natural Science Foundation of China (No. 81702195), and the Project of Application Demonstration Center of Precision Medicine for Molecular Diagnosis in Jilin Province (2016–2018, NDRC)

Background: Globally, gastric cancer (GC) is the third most common source of cancer-associated mortality. The aim of this study was to identify key genes and circular RNAs (circRNAs) in GC diagnosis, prognosis, and therapy and to further explore the potential molecular mechanisms of GC.


Material/Methods: Differentially expressed genes (DEGs) and circRNAs (DE circRNAs) between GC tissues and adjacent non-tumor tissues were identified from 3 mRNA and 3 circRNA expression profiles. Functional analyses were performed, and protein–protein interaction (PPI) networks were constructed. The significant modules and key genes in the PPI networks were identified. Kaplan-Meier analysis was performed to evaluate the prognostic value of these key genes. Potential miRNA-binding sites of the DE circRNAs and target genes of these miRNAs were predicted and used to construct DE circRNA–miRNA–mRNA networks.

Results: A total of 196 upregulated and 311 downregulated genes were identified in GC. The results of functional analysis showed that these DEGs were significantly enriched in a variety of functions and pathways, including extracellular matrix-related pathways. Ten hub genes (*COL1A1*, *COL3A1*, *COL1A2*, *COL5A2*, *FN1*, *THBS1*, *COL5A1*, *SPARC*, *COL18A1*, and *COL11A1*) were identified via PPI network analysis. Kaplan-Meier analysis revealed that 7 of these were associated with a poor overall survival in GC patients. Furthermore, we identified 2 DE circRNAs, *hsa_circ_0000332* and *hsa_circ_0021087*. To reveal the potential molecular mechanisms of circRNAs in GC, DE circRNA–microRNA–mRNA networks were constructed.

Conclusions: Key candidate genes and circRNAs were identified, and novel PPI and circRNA–microRNA–mRNA networks in GC were constructed. These may provide useful information for the exploration of potential biomarkers and targets for the diagnosis, prognosis, and therapy of GC.

MeSH Keywords: **Diagnosis • Gene Expression Profiling • Stomach Neoplasms**

Full-text PDF: <https://www.medscimonit.com/abstract/index/idArt/915382>

 3525

 10

 10

 63



Background

Globally, gastric cancer (GC) is the third most common cause of cancer-associated mortality [1]. In the United States, an estimated 26 240 cases of gastric cancer were diagnosed in 2018, and 10 800 patients died of the disease [2]. Because of the lack of specific signs of early GC, most patients (>70%) are diagnosed with advanced-stage GC, after the best opportunity for treatment [3]. Although considerable advancements in the diagnosis and treatment of GC have been achieved, the prognosis of GC patients remains poor, particularly for those with advanced disease [4]. Therefore, a deeper understanding of the mechanisms involved in GC progression and identification of potential biomarkers and targets for the diagnosis, prognosis, and therapy of GC are urgently needed.

Bioinformatics analyses, including the analysis of gene interaction networks, microarray expression profiles, and gene annotations, are being utilized as powerful tools for identifying potential diagnostic and treatment-relevant biomarkers of cancers [5]. For instance, by analyzing data from the Gene Expression Omnibus (GEO) database, Cao et al. identified 5 genes (*COL1A2*, *COL1A1*, *COL4A1*, *THBS2*, and *ITGA5*) as potential biomarkers and therapeutic targets for GC [6]. In addition, by analyzing data from GEO and The Cancer Genome Atlas (TCGA), Zhu et al. found that high expression of *CDK1* is a prognostic factor for hepatocellular carcinoma (HCC), making it a potential therapeutic target and biomarker for the diagnosis of HCC [7]. In particular, the method of integrated bioinformatics analysis can be used to overcome inaccuracies in sequencing due to small sample sizes [8].

Circular RNAs (circRNAs) are a novel class of endogenous non-coding RNAs that form a covalently closed continuous loop by back-splicing events via exon or intron circularization [9,10]. Owing to the development of high-throughput sequencing, researchers have discovered thousands of circRNAs, and emerging evidence has shown that circRNAs are involved in the progression of oncogenesis, invasion, and metastasis by regulating microRNAs (miRNAs) [10]. However, at present, most studies involving circRNAs have been limited to the sequencing of a few samples [11] or exploring the biological function of a single circRNA [12]. To the best of our knowledge, no researchers have yet used integrated analysis of multiple datasets to investigate GC-related circRNAs.

In this study, differentially expressed genes (DEGs) between human GC tissues and adjacent non-tumor tissues were identified via analysis of public GEO datasets. Next, to explore the roles of these genes, functional enrichment analyses and pathway enrichment analyses were performed. Then, protein-protein interaction (PPI) networks were successfully constructed. The key genes and significant modules in the networks were identified. Kaplan-Meier analysis was performed to evaluate

the prognostic value of these key genes. Furthermore, 3 additional circRNA expression profiles were analyzed to identify differentially expressed circRNAs (DE circRNAs) between GC and adjacent non-tumor tissues. Finally, DE circRNA-miRNA-mRNA networks were constructed. The present study provides useful information for the exploration of potential biomarkers and targets for the diagnosis, prognosis, and therapy of GC.

Material and Methods

Microarray data source

Microarray data were obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo>), a free database containing microarray profiles and next-generation sequencing data. Three mRNA expression profiles (GSE19826, GSE54129, and GSE79973), including data from 133 GC tissues and 46 adjacent non-tumor tissues, were downloaded to identify DEGs, while 3 circRNA expression profiles (GSE83521, GSE89143, and GSE93541), including data from 12 GC samples and 12 normal gastric samples, were downloaded to select DE circRNAs. Details for each dataset are provided in Table 1.

Identification of DEGs and DE circRNAs

The GEO2R tool (<https://www.ncbi.nlm.nih.gov/geo/geo2r>) was used to identify DEGs and DE circRNAs between GC and adjacent non-tumor tissues. Values of $|\log_2 \text{fold change}|$ ($|\log_2 \text{FC}|$) > 1 and $p < 0.05$ were set as cut-off criteria. Venn diagrams were constructed to determine intersections via VENNY 2.1 software (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>) [13].

GO analysis and KEGG analysis

GO analysis is widely used to provide gene annotation terms [14], and the KEGG database is widely used for pathway enrichment analysis [15]. In the present study, an online tool, the Database for Annotation, Visualization and Integration Discovery (DAVID 6.8) (<https://david.ncifcrf.gov>), was used to perform GO analysis [16]. KEGG analysis was performed using the online tool KOBAS 3.0 (<http://kobas.cbi.pku.edu.cn>) [17], with 4 databases, including KEGG Pathway (<https://www.genome.jp/kegg>), Reactome (<http://www.reactome.org>), BioCyc (<https://biocyc.org/>), and PANTHER (<http://www.pantherdb.org>), were used in the analysis. A value of $p < 0.05$ was selected as the threshold for significant GO and KEGG terms.

Construction of PPI and DE circRNA-miRNA-mRNA networks

Search Tool for the Retrieval of Interacting Genes (STRING) (<https://string-db.org/>) was used to construct PPI networks [18].

Table 1. Expression profile information.

Series	Type	First author	Publication year	Country	Platform	Sample	Number of samples (normal/tumor)
GSE19826	mRNA	Wang Q	2010	China	GPL570	Stomach	27 (15/12)
GSE54129	mRNA	Liu B	2017	China	GPL570	Stomach	132 (21/111)
GSE79973	mRNA	Shao Q	2016	China	GPL570	Stomach	20 (10/10)
GSE83521	circRNA	Zhang Y	2017	China	GPL19978	Stomach	12 (6/6)
GSE89143	circRNA	Shao Y	2017	China	GPL19978	Stomach	6 (3/3)
GSE93541	circRNA	Guo J	2017	China	GPL19978	Stomach	6 (3/3)

A combined interaction score >0.5 was used as the cut-off criterion for the extraction of PPI pairs. Cytoscape software version 3.6.0 (<http://www.cytoscape.org/>) was used for visualization of the PPI and DE circRNA–miRNA–mRNA networks [19]. The cytoHubba plug-in was used to screen hub genes, with a cut-off of degree ≥ 20 . The MCODE plug-in was used to identify significant modules of the PPI networks in Cytoscape, with cut-off criteria of degree ≥ 5 and k-core ≥ 4 .

Validation of gene expression

UALCAN (<http://ualcan.path.uab.edu/>), an online tool used to analyze gene expression data from The Cancer Genome Atlas (TCGA) database (including 34 normal gastric tissues and 415 primary tumors), was used to validate the expression of hub genes in the PPI networks. $P < 0.05$ was considered to indicate a statistically significant difference [20].

Survival analysis

The Kaplan-Meier plotter database (<http://kmplot.com/>), which is an online tool used to evaluate the prognostic values of genes in breast, ovarian, lung, and gastric cancer patients, was applied to analyze the associations between the hub genes in the PPI networks and overall survival [21]. Differences in the survival rates were statistically analyzed between patients with high and low expression levels of hub genes. The hazard ratio (HR) and its 95% confidence intervals were calculated. $P < 0.05$ was considered to indicate a statistically significant difference.

Prediction of circRNA and miRNA targets

The miRNA-binding sites of DE circRNAs were predicted using the Circular RNA Interactome online tool (<https://circinteractome.nia.nih.gov/>) [22]. Possible targets of miRNAs were predicted using the miRWalk 2.0 online tool (<http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2>), with a “predicted target module” [23]. Genes that appeared in 4 (miRWalk, miRanda, miRDB, and Targetscan) of the 12 software programs from

miRWalk v.2.0 were considered as possible target genes. Then, comparison of the list of potential target genes with the list of DEGs identified from the 3 gene expression profiles (GSE19826, GSE54129, and GSE79973) produced an intersection of genes used to construct the DE circRNA–miRNA–mRNA networks.

Results

Identification of DEGs in GC

Three mRNA expression profiles including data from GC and adjacent non-tumor tissues were obtained from the NCBI GEO database (Table 1). In total, 2949, 3675, and 3251 DEGs from the GSE19826, GSE54129, and GSE79973 datasets, respectively, were extracted using GEO2R and met the thresholds of $|\log_2FC| > 1.0$ and $p < 0.05$ (Supplementary Table 1, Figure 1). By comparing the DEGs from the 3 cohort profile datasets, an intersection of 507 common DEGs was obtained, including 196 genes that were upregulated and 311 that were downregulated in GC tissues compared to levels in adjacent non-tumor tissues (Figure 2). A list of all 507 DEGs is provided in Supplementary Table 2.

GO analysis and KEGG analysis of DEGs

DEGs were functionally classified into the BP, CC, and MF categories. In the BP category, the top 3 most enriched terms were “collagen fibril organization”, “cell adhesion”, and “angiogenesis”. In the CC category, the top 3 most enriched terms were “extracellular space”, “proteinaceous extracellular matrix”, and “collagen trimer”. In the MF category, the top 3 most enriched terms were “extracellular matrix structural constituent”, “heparin binding”, and “extracellular matrix binding” (Supplementary Table 3, Figure 3A–3C). Moreover, the top 3 most enriched terms in the KEGG analysis were “extracellular matrix organization”, “degradation of the extracellular matrix”, and “collagen formation” (Supplementary Table 4, Figure 3D).

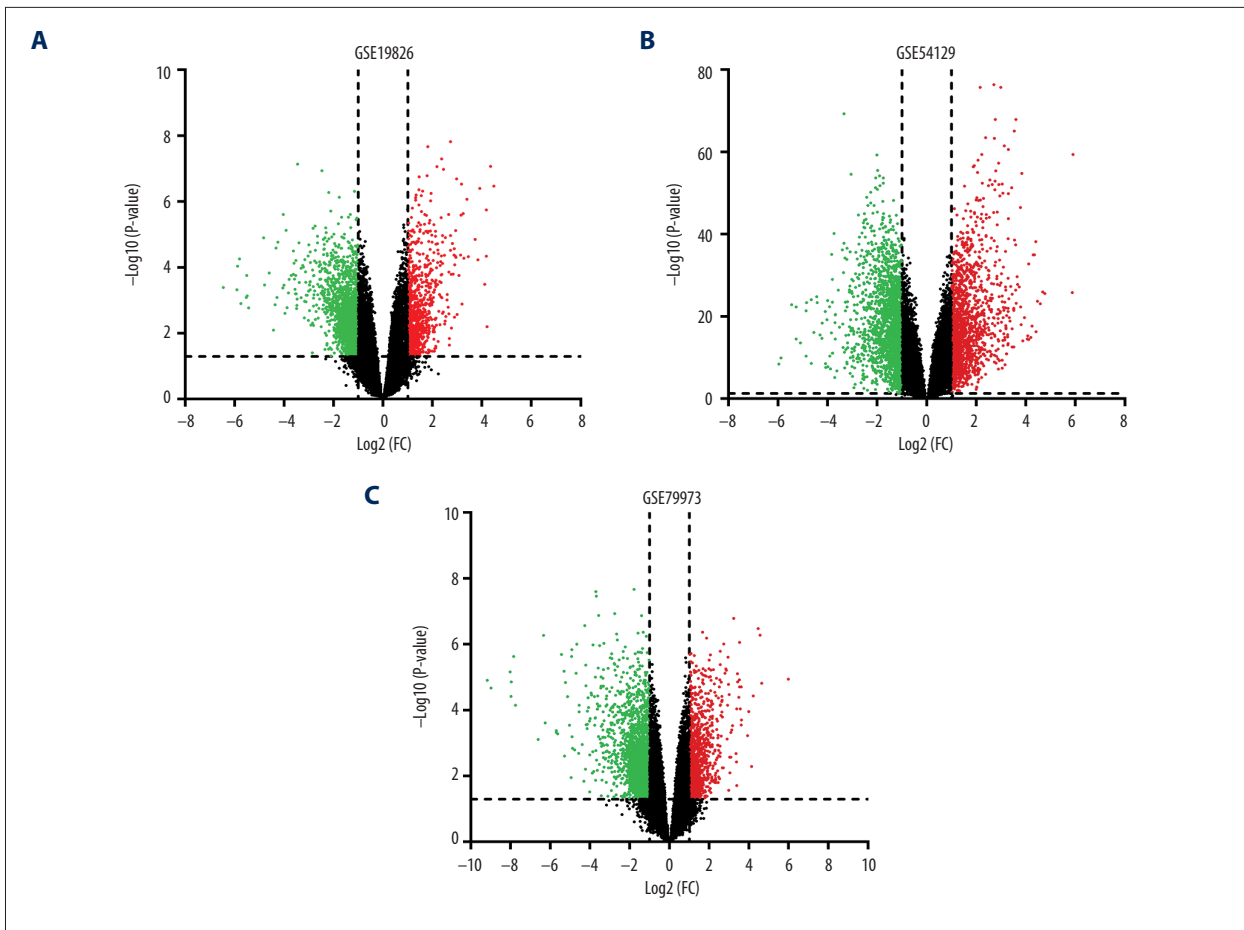


Figure 1. DEGs between GC and adjacent non-tumor tissues. **(A)** Volcano plot for DEGs in GSE19826. **(B)** Volcano plot for DEGs in GSE54129. **(C)** Volcano plot for DEGs in GSE79973. Red: $\log_2(\text{FC}) > 1$, $p < 0.05$; green: $\log_2(\text{FC}) < -1$, $p < 0.05$. DEG – differentially expressed gene; GC – gastric cancer; FC – fold change.

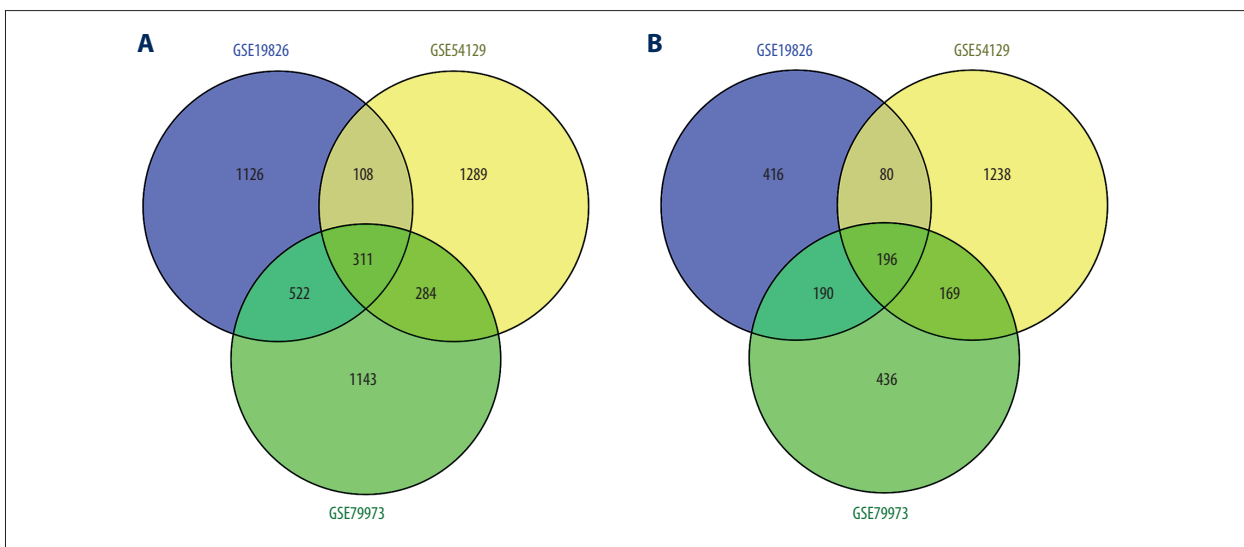


Figure 2. Venn diagram of DEGs in the 3 cohort profile datasets (GSE19826, GSE54129, and GSE79973). **(A)** Upregulated DEGs. **(B)** Downregulated DEGs. Different color areas represent different datasets. The overlapping areas are the common DEGs. DEG – differentially expressed gene.

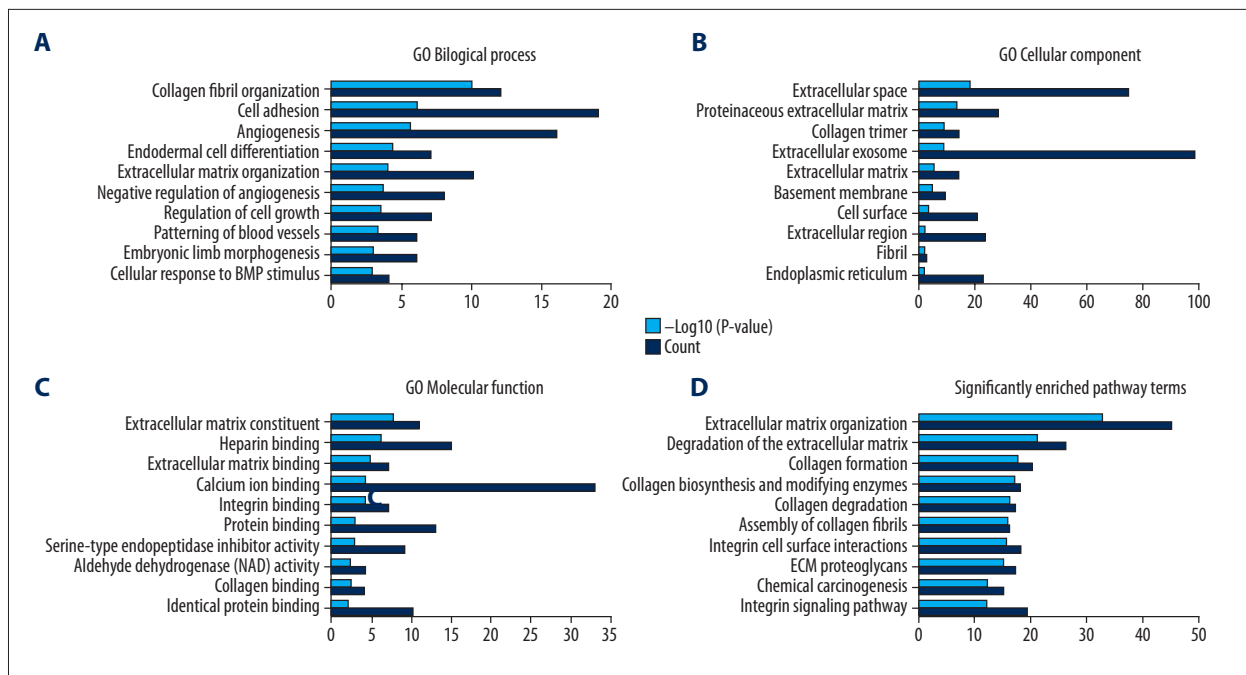


Figure 3. GO annotation and pathway enrichment analysis. (A) Top 10 terms in BP category. (B) Top 10 terms in CC category. (C) Top 10 terms in MF category. (D) Top 10 terms in pathway enrichment analysis. GO – gene ontology; DEG – differentially expressed gene; BP – biological process; CC – cellular component; MF – molecular function.

PPI network construction and modular analysis

A total of 253 DEGs (122 upregulated and 131 downregulated) among the 507 common DEGs were used to construct the PPI networks, which included 253 nodes and 588 edges (Figure 4). Using the cytoHubba plug-in, 10 genes – *COL1A1*, *COL3A1*, *COL1A2*, *COL5A2*, *FN1*, *THBS1*, *COL5A1*, *SPARC*, *COL18A1*, and *COL11A1* – were confirmed as hub genes (Table 2, Figure 5A). All of the hub genes were upregulated in the GC samples compared to levels in normal gastric samples. We further verified the expression of the 10 hub genes in the TCGA stomach adenocarcinoma database using UALCAN. The results showed that the expression levels of 9 of the 10 genes (all except *THBS1*), were significantly higher in primary tumors than those in the normal stomach tissues of GC patients from TCGA, suggesting that our results are highly credible (Supplementary Figure 1). Using the MCODE plug-in, 3 significant modules were screened from the PPI networks (Figure 5B–5D). KEGG enrichment analysis showed that Module 1 was mainly associated with collagen metabolism, including collagen biosynthesis and modifying enzymes, collagen formation, collagen degradation, and assembly of collagen fibrils. Module 2 was mainly associated with platelet function, including platelet degranulation; response to elevated platelet cytosolic Ca²⁺; and platelet activation, signaling, and aggregation. Module 3 was mainly associated with chemical carcinogenesis, – O linked glycosylation, metabolic pathways, and biological oxidation (Figure 5E, Supplementary Table 5).

Survival analysis of hub genes

Using the Kaplan-Meier plotter database, the prognostic value of the 10 hub genes was evaluated in 876 GC patients. The results indicated that 7 out of the 10 hub genes – *COL1A1*, *COL1A2*, *COL5A2*, *COL11A1*, *COL18A1*, *FN1*, and *SPARC* – were related to the OS of GC patients. High expression of each of the 7 genes was associated with poor OS in GC patients ($p < 0.05$). In contrast, the other 3 hub genes – *COL3A1*, *COL5A2*, and *THBS1* – were not related to the OS of GC patients ($p > 0.05$) (Figure 6).

Identification of DE circRNAs in GC

Three circRNA expression profiles including data from GC and adjacent non-tumor tissues were obtained from the GEO database (Table 1). In total, 150, 191, and 416 DE circRNAs from the GSE83521, GSE89143, and GSE93541 expression profiles, respectively, were extracted using GEO2R, meeting the thresholds of $|\log_2FC| > 1.0$ and $p < 0.05$ (Supplementary Table 6). By comparing the DE circRNAs from the 3 cohort profile datasets, we obtained 2 common DE circRNAs – *hsa_circ_0000332* and *hsa_circ_0021087* – both of which were downregulated in GC tissues compared to levels in adjacent non-tumor tissues (Figure 7A, 7B).

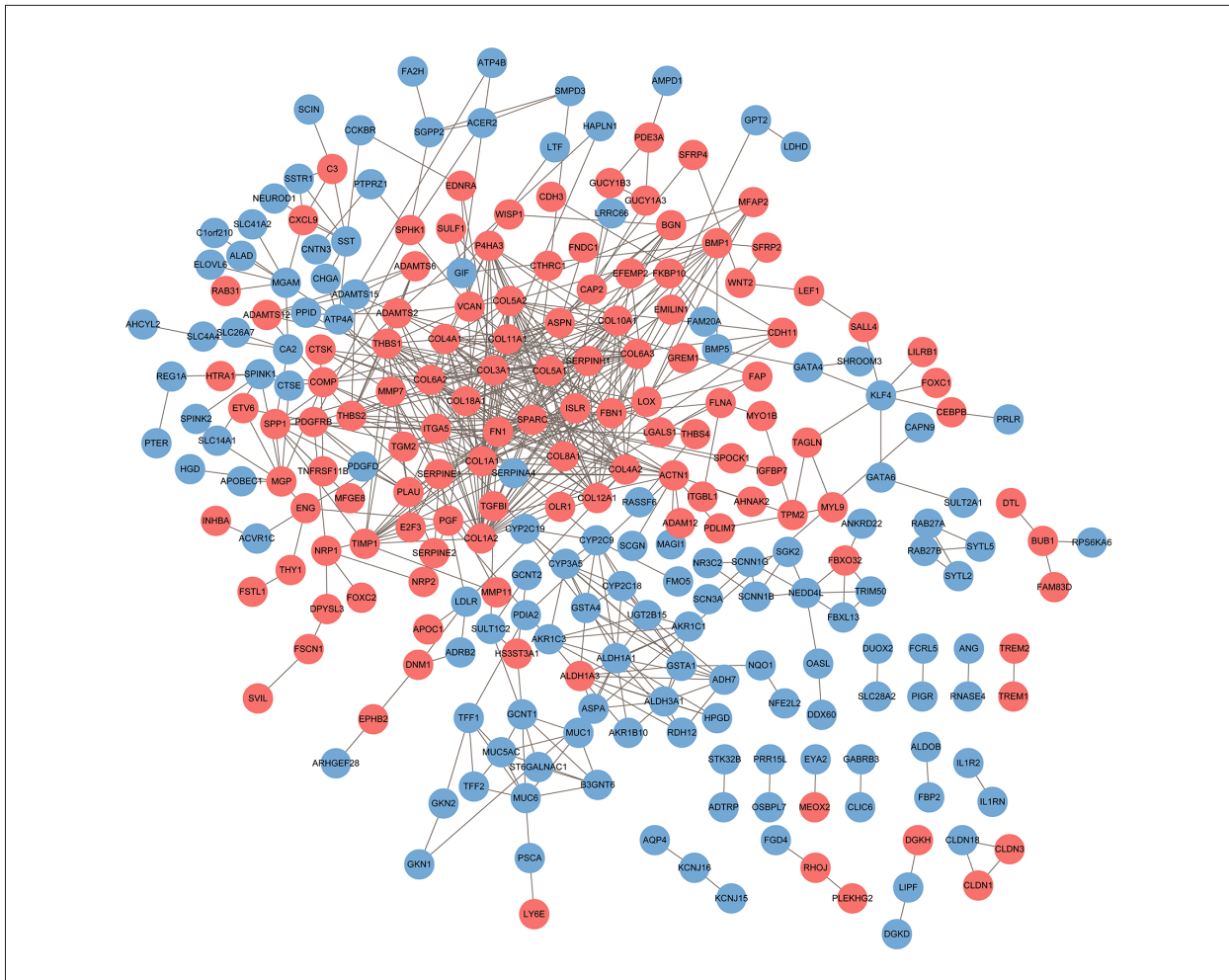


Figure 4. PPI networks of the DEGs. The network contains 253 nodes and 588 edges. Network nodes represent proteins (shown with gene names). The color of each node denotes the expression of genes in the GC samples compared to that in non-tumor samples (pink represents upregulated and blue represents downregulated). PPI – protein–protein interaction; DEG – differentially expressed gene.

Table 2. Hub genes in the protein–protein interaction networks identified by cytoHubba.

Gene symbol	Official full name	Degree	Expression in GC
<i>COL1A1</i>	Collagen type I alpha 1 chain	32	Up-regulation
<i>COL3A1</i>	Collagen type III alpha 1 chain	32	Up-regulation
<i>COL1A2</i>	Collagen type I alpha 2 chain	31	Up-regulation
<i>COL5A2</i>	Collagen type V alpha 2 chain	23	Up-regulation
<i>FN1</i>	Fibronectin 1	23	Up-regulation
<i>THBS1</i>	Thrombospondin 1	22	Up-regulation
<i>COL5A1</i>	Collagen type V alpha 1 chain	22	Up-regulation
<i>SPARC</i>	Secreted protein acidic and cysteine rich	22	Up-regulation
<i>COL18A1</i>	Collagen type XVIII alpha 1 chain	21	Up-regulation
<i>COL11A1</i>	Collagen type XI alpha 1 chain	20	Up-regulation

GC – gastric cancer.

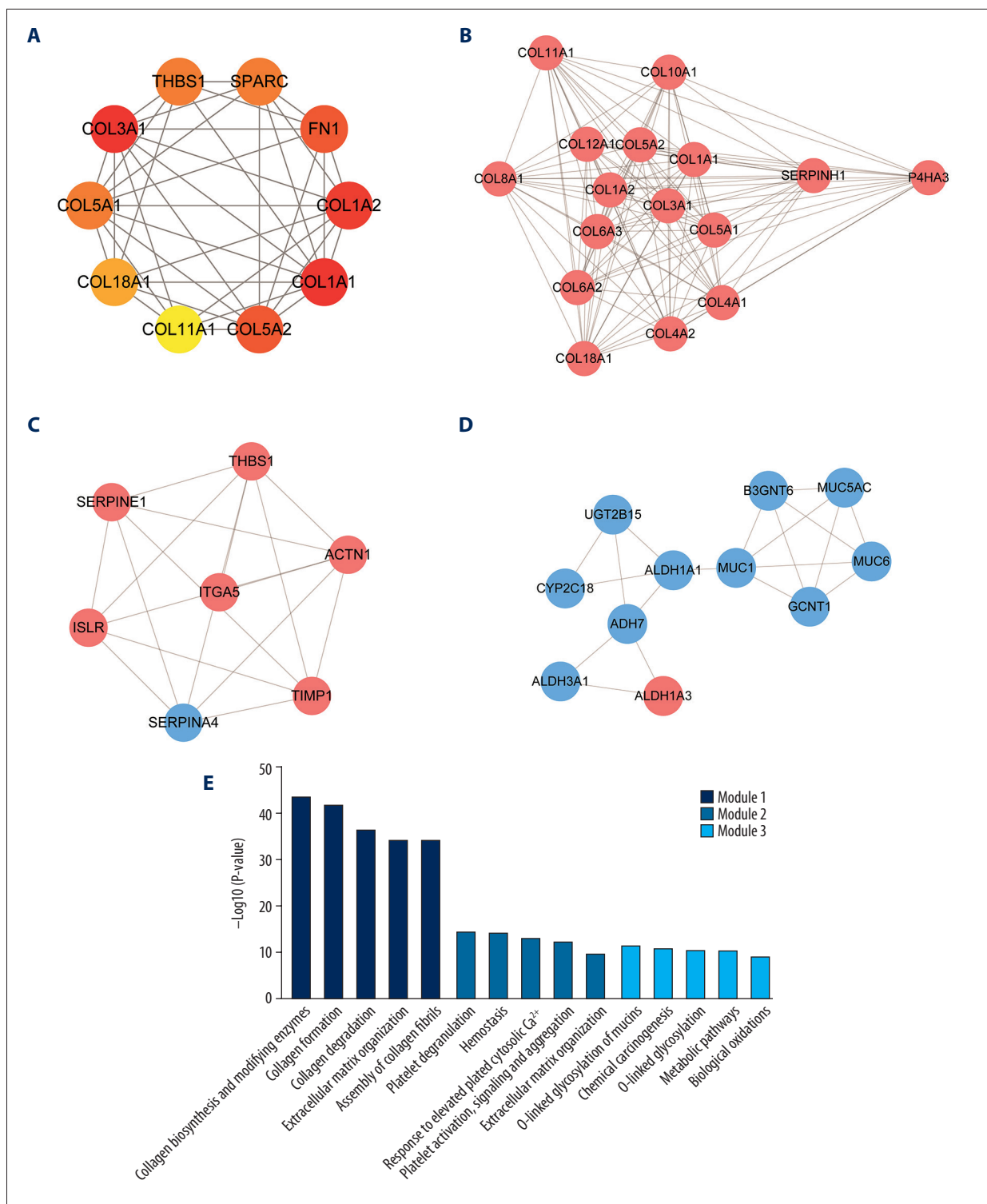


Figure 5. Hub genes and module analysis of the PPI networks. **(A)** Hub genes identified by cytoHubba plug-in (interaction degree ≥ 20). Node color denotes interaction degree (red for high degree, orange for intermediate degree, and yellow for low degree). **(B)** Module 1, identified by MCODE plug-in, contains 16 genes. **(C)** Module 2 contains 7 genes. **(D)** Module 3 contains 11 genes. The color of each node denotes the expression of genes in the GC samples compared to that in normal gastric samples (pink for upregulation and blue for downregulation). **(E)** Pathway enrichment analysis of 3 significant modules in PPI networks. PPI, protein-protein interaction.

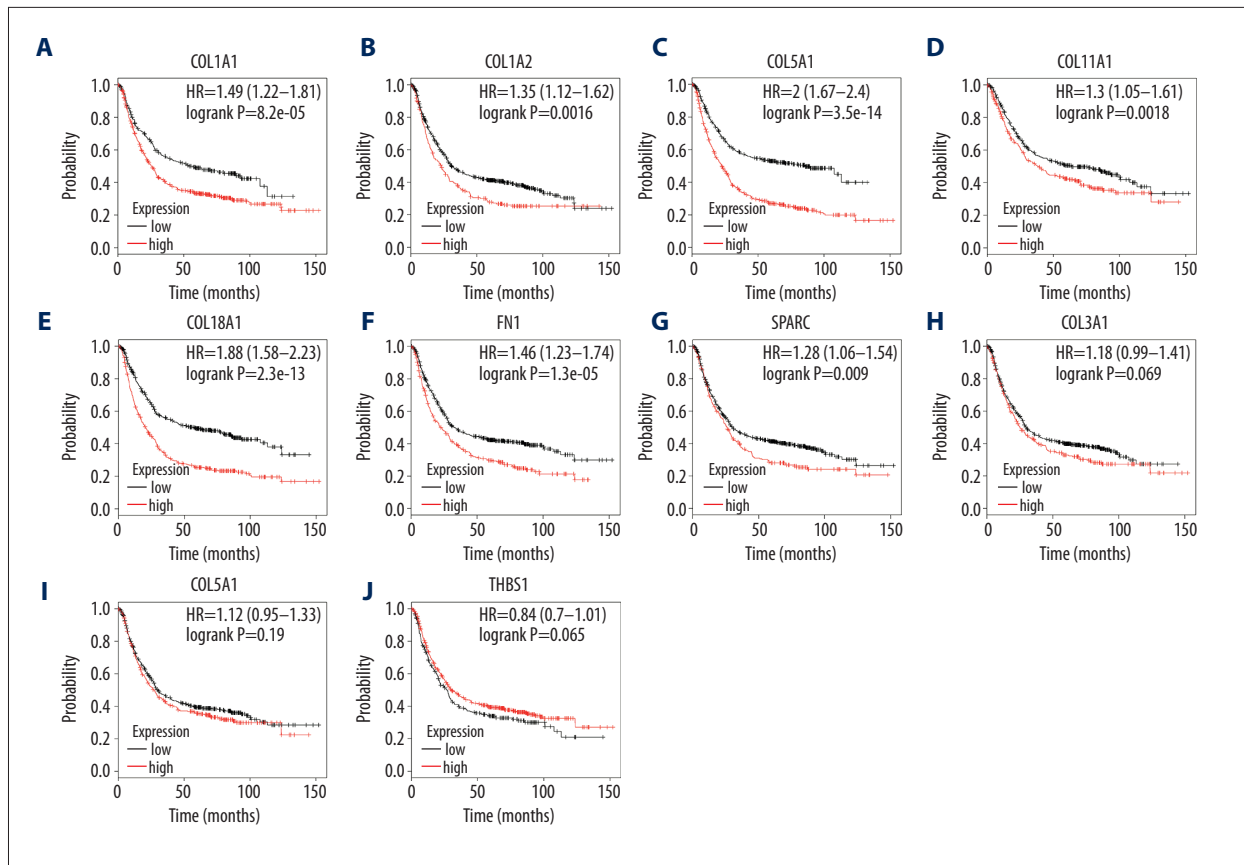


Figure 6. Association of the 10 hub genes with the OS of GC patients, analyzed by Kaplan-Meier survival plots. (A–G) High expression of *COL1A1*, *COL1A2*, *COL5A2*, *COL11A1*, *COL18A1*, *FN1*, and *SPARC* was associated with poor OS of GC patients. (H–J) Expression of *COL3A1*, *COL5A2*, and *THBS1* was not related to OS of GC patients. OS, overall survival; GC, gastric cancer.

Prediction of targets for circRNAs and miRNAs and construction of DE circRNA–miRNA–mRNA networks

Using the Circular RNA Interactome online tool to predict potential miRNAs capable of binding to the 2 DE circRNAs, we obtained a total of 13 miRNAs, 2 of which were for hsa_circ_0000332, while the remaining 11 were for hsa_circ_0021087. Using the miRWalk 2.0 online tool to predict the targets of these 13 miRNAs, we obtained 1563 genes. Then, by comparing the 1563 predicted targets with the 507 common DEGs identified from the gene expression profiles, we found that none of the predicted targets of 5 of the 13 miRNAs were included in the list of DEGs, while there were 42 predicted targets of 8 of the 13 miRNAs that were present on the list of DEGs (Table 3, Figure 7C). Based on these results, we constructed circRNA–miRNA–mRNA networks with the 2 circRNAs, 8 miRNAs, and 42 target genes using Cytoscape 3.6.0 software (Figure 8). From the networks, we found that hsa_circ_0021087, which interacted with 7 miRNAs and indirectly interacted with 39 target genes, may play a more important circRNA role in the GC network.

KEGG analysis of genes in the DE circRNA–miRNA–mRNA networks

KEGG analysis of the 42 genes in the DE circRNA–miRNA–mRNA networks was performed. The results indicated that these 42 genes were primarily enriched in the following terms: “signal transduction”, “extracellular matrix organization”, and “non-integrin membrane-ECM interactions” (Supplementary Table 7, Figure 9).

Discussion

Numerous studies have been conducted to understand the causes and underlying mechanisms of GC occurrence and development over the past several decades; however, the incidence and mortality of GC remain very high globally [1,4]. With the development of next-generation sequencing, large-scale high-throughput molecular profiling studies are now providing molecular insights into the complex nature of this disease [24]. The overall aim of this study was to identify key genes and circRNAs for GC diagnosis, prognosis, and therapy

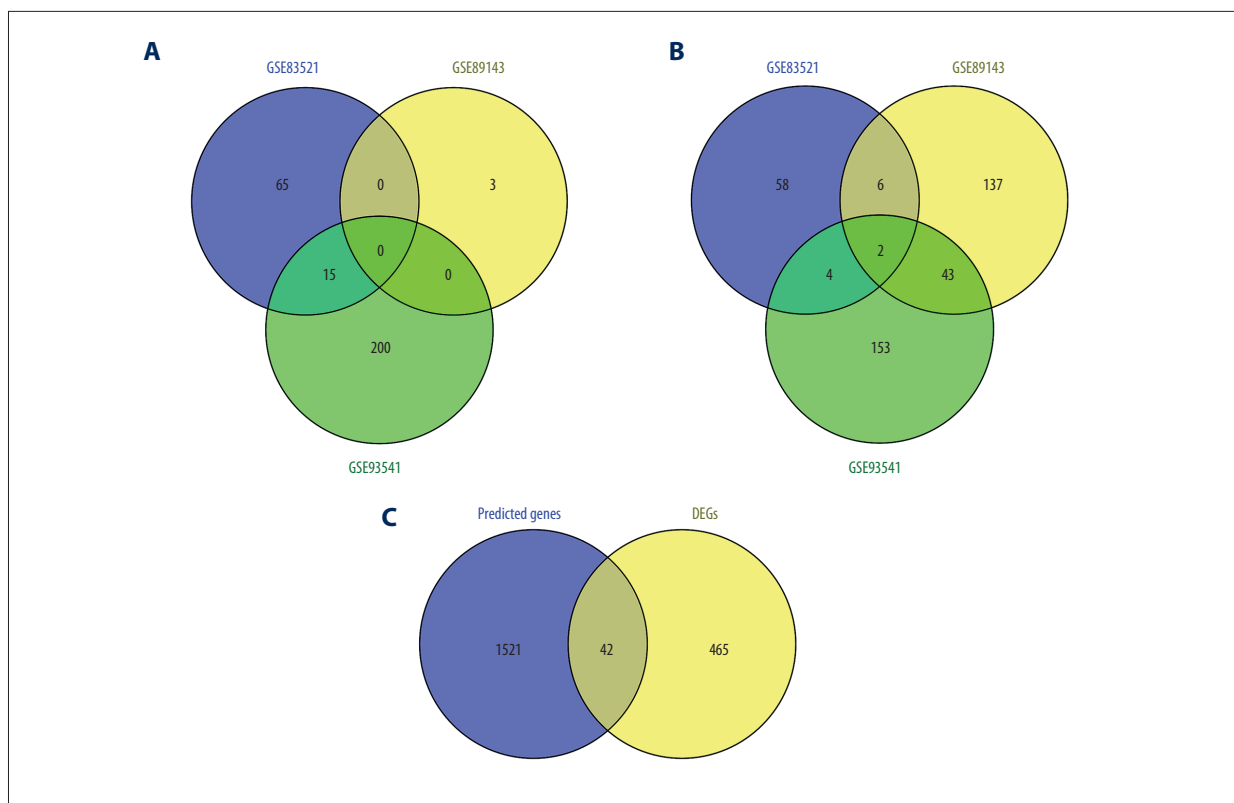


Figure 7. Venn diagrams of (A) common upregulated and (B) common downregulated DE circRNAs in the 3 cohort profile datasets (GSE83521, GSE89143, and GSE93541). Different colored areas represent different datasets. The overlapping areas indicate common DE circRNAs. (C) Predicted target genes overlapping DEGs identified from the gene expression profiles. DE circRNAs, differentially expressed circRNAs; DEGs, differentially expressed genes.

Table 3. Prediction of targets for circRNAs and miRNAs.

DE circRNAs	Predicted miRNAs binding to DE circRNAs	Predicted targets for miRNAs overlapped with DEGs
hsa_circ_0000332	hsa-miR-1292-5p	None
	hsa-miR-370-3p	<i>CLIC6, NDOR1, PRLR</i>
	hsa-miR-1224-3p	<i>TFCP2L1</i>
	hsa-miR-1276	<i>COL4A2, FMNL3, FGD6, PAIP2B, GRAMD1C, SLC14A1, DUOX2</i>
	hsa-miR-1304-5p	<i>THY1, DPYSL3, LY6E, FAM189A2, CMTM4, DNER, SYTL2, KIAA1324, SCNN1G, ACVR1C, SLC4A4, PRSS12</i>
	hsa-miR-184	None
hsa_circ_0021087	hsa-miR-450b-3p	<i>ADAM12, TGM2, C4orf19, MAP7D2</i>
	hsa-miR-769-3p	<i>ADAM12, C4orf19, LDLR</i>
	hsa-miR-517a-3p	None
	hsa-miR-517c-3p	None
	hsa-miR-623	<i>NRP2, ELOVL6, LIPH</i>
	hsa-miR-638	None
	hsa-miR-1827	<i>SERPINH1, EFEMP2, IGF2BP3, FGD6, THBS1, MXRA7, AKR7A3, METTL7A, SLC26A9, IQGAP2, KIAA1324, PLAC8, IL1RN</i>

DE circRNAs – differentially expressed circRNAs.

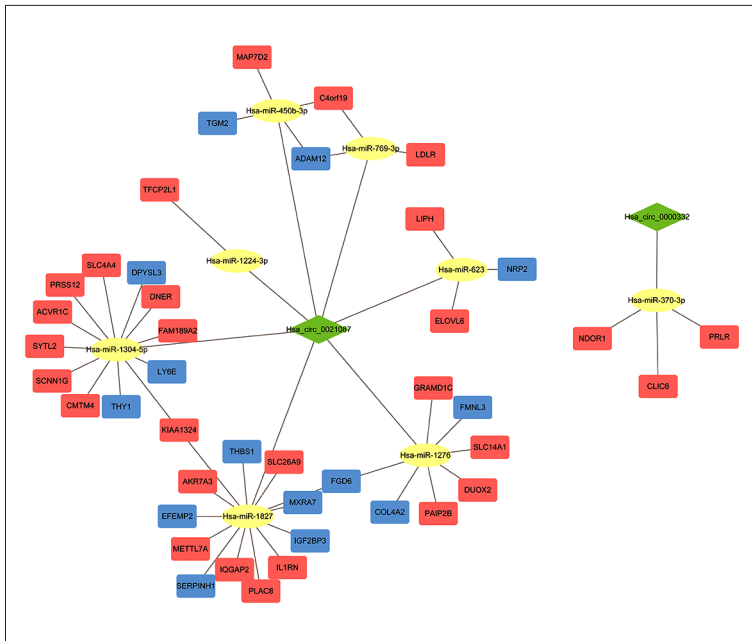


Figure 8. DE circRNA-miRNA-mRNA networks. The network contains 52 nodes and 54 edges. The green diamond nodes represent circRNAs; yellow ellipsoidal nodes, miRNAs; orange square nodes, upregulated genes; blue square nodes, downregulated genes.

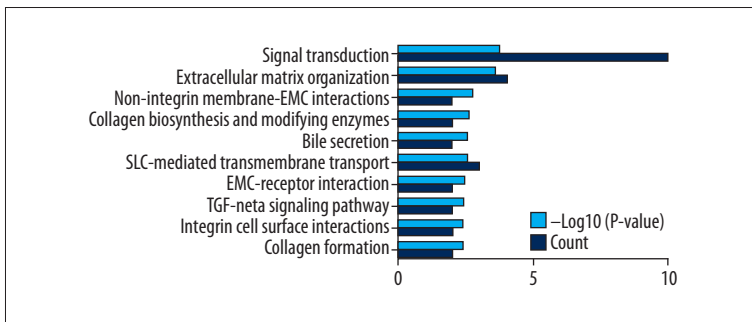


Figure 9. Pathway enrichment analysis of the 42 genes in the DE circRNA-miRNA-mRNA networks.

and to further explore the potential mechanisms of GC by integrated profiling analysis.

Initially, 3 public GEO datasets were analyzed to identify DEGs between human GC tissues and adjacent non-tumor tissues. A total of 507 DEGs (311 upregulated and 196 downregulated) were identified. To further elucidate the underlying functions of these DEGs, GO analysis was performed. The results indicated several abnormally modified GO terms closely associated with cancer, such as cell adhesion, angiogenesis, negative regulation of angiogenesis, and regulation of cell growth. The results of KEGG analysis indicated that the most significantly enriched pathways were extracellular matrix (ECM)-related pathways, such as ECM organization, degradation of the ECM, and ECM proteoglycans. The ECM is a noncellular component that is composed of 2 major classes of molecules: glycosaminoglycans and fibrous proteins, which include collagen, elastin, fibronectin, and laminin [25]. As an important member of the tumor microenvironment, the ECM secreted by cancer and/or stromal cells affects the dissemination process through modulation of the interaction between cancer and endothelial cells

and plays important roles in tumor angiogenesis, metastasis, and chemotherapy resistance [26]. Furthermore, other significantly enriched pathways, including integrin-related pathways (integrin cell surface interactions and integrin signaling pathway) and collagen-related pathways (collagen formation, collagen biosynthesis and modifying enzymes, collagen degradation, and assembly of collagen fibrils) are also related to the ECM. Integrins are a family of essential cell surface adhesion receptors with the ability to sense the content and stiffness of the surrounding ECM and transfer this information to cytoplasmic signaling pathways [27]. Collagen, as a key structural component of the ECM, is known to regulate cell migration, proliferation, differentiation, and survival by signaling through cell surface receptors, including integrins [28]. Thus, the results of the present study suggest that pathways associated with the ECM play an important role in the occurrence of GC and could be potential therapeutic targets.

PPI networks were constructed, and the most closely related genes were identified: *COL1A1*, *COL3A1*, *COL1A2*, *COL5A2*, *FN1*, *THBS1*, *COL5A1*, *SPARC*, *COL18A1*, and *COL11A1*. Among these,

COL1A1, *COL3A1*, *COL1A2*, *COL5A2*, *COL5A1*, *COL18A1*, and *COL11A1* belong to the collagen family, which comprises 28 members (I–XXVIII) [29]. Each collagen consists of 3 polypeptide chains (α chains) numbered with Arabic numerals [29]. For instance, collagen type V $\alpha 2$ (*COL5A2*) encodes the pro- $\alpha 2$ chains of type V collagen, and collagen type III $\alpha 1$ (*COL3A1*) encodes the pro- $\alpha 1$ chains of type III collagen. As key structural components of the ECM, collagens have been found to be overexpressed in multiple cancers to provide a rigid matrix that facilitates tumor growth [25]. The expression of *COL1A1* and *COL1A2* has been confirmed by some researchers [30,31] to be upregulated in GC tissues compared to levels in normal tissues and may predict an adverse prognosis in GC patients [32]. Furthermore, Zhang et al. found that *COL1A1* promotes CRC metastasis and may serve as an oncoprotein by regulating key genes in the WNT/PCP signaling pathway [33]. Ao et al. found that the silencing of *COL1A2* inhibits the proliferation and promotes the apoptosis of GC cells through the PI3k-Akt signaling pathway [34]. Overexpression of *COL3A1* has been confirmed in many cancers, including bladder cancer [35], glioblastoma [36], and GC [37]. Boguslawska et al. found that the expression of some genes related to adhesion and ECM remodeling, including *COL1A1*, *COL5A1*, and *COL11A1*, was upregulated in renal cell carcinoma compared with levels in controls [38]. *COL5A1* is also reported to be involved in breast cancer [39], oral squamous cell carcinoma [40], lung adenocarcinoma [41], and GC [42]. *COL5A2* has previously been found to be associated with the pathological processes of osteosarcoma [43], bladder cancer [44], and GC [45]. The role of *COL18A1* is discrepant in different cancers. Fang et al. found that *COL18A1* may be involved in bladder cancer progression, as its expression was increased with the progression of pathological stage [46]. However, a study by Qiu et al. found that the expression level of *COL18A1* was negatively correlated with metastasis of melanoma, indicating that *COL18A1* can inhibit the proliferation and metastasis of melanoma [47]. Nevertheless, there are few published investigations of the expression of *COL18A1* in GC. Furthermore, other hub genes may additionally be involved in the metastasis of GC. Secreted protein acidic and rich in cysteine (SPARC) is regarded as a collagen chaperone because it binds to the procollagens and its existence is essential for collagen deposition in tissues [29,48]. SPARC overexpression has been demonstrated in some cancers, such as pancreatic cancer [49], esophageal squamous cell carcinoma [50], and GC [51]. In contrast, some studies have shown that SPARC expression is reduced in bladder cancer [52] and acute leukemia [53]. Fibronectin (FN) is a glycoprotein component of the ECM involved in cell adhesion, migration, metastasis, proliferation, and differentiation [54]. Recent evidence showed that FN is overexpressed in GC cells and that microRNA-200c can suppress the proliferation of GC cells by selectively downregulating FN expression [55]. Thrombospondin 1 (THBS1) is an extracellular glycoprotein that plays multiple roles

in the ECM and cell–cell interactions [56]. Numerous studies have confirmed a role for THBS1 in cell invasion and migration. However, conflicting results have been obtained in different cell types [57]. In GC cells, Huang et al. demonstrated a marked overexpression of THBS1 when compared to levels in adjacent normal tissues. Nevertheless, THBS1 was correlated with well- and moderately-differentiated tumors, clearly indicating a less aggressive tumor biology [56].

In the present study, we also evaluated the prognostic value of the hub genes in GC using Kaplan-Meier analysis. The results showed that high expression of *COL1A1*, *COL1A2*, *COL5A2*, *COL11A1*, *COL18A1*, *FN1*, and *SPARC* was related to poor OS in patients with GC. These results suggest that these genes may provide novel tools for prognostication, as well as drug targets for GC patients.

Additionally, the 3 most significant modules were extracted from the PPI network. Module 1 was mainly associated with the ECM and collagen, which have been thoroughly discussed above. Interestingly, we found that the significantly enriched pathways and most of the corresponding genes of Module 2 were mainly associated with platelets and hemostasis. It has been demonstrated that cancer increases the risk of thrombosis by 4-fold, indicating a close relationship between cancer and platelet activation [58]. Platelets induce cancer cell plasticity and result in the enhancement of extravasation and survival during the process of circulating tumor cell dissemination [58,59]. Similarly, tumor cells induce platelet aggregation, suggesting that the interaction between platelets and cancers is reciprocal [58,59]. Mikami et al. showed that platelets promoted the proliferation of GC cells [60]. Hiroto found that platelets contribute to creating a favorable microenvironment for resistance to chemotherapy [61].

Increasing evidence has shown that circRNAs are abundant and stable RNAs in mammalian cells and are important regulators of some biological processes related to cancer, including cell invasion and metastasis [62]. The most well-known function of circRNAs is their ability to “sponge up” miRNAs, thus serving as competitive inhibitors and suppressing the ability of the miRNA to bind its mRNA targets [63]. However, at present, the biological functions of most circRNAs are unclear, and studies of the underlying functional pathways and the co-expression networks among circRNAs, miRNAs, and mRNAs in GC are limited. In the present study, 3 circRNA expression profiles were analyzed to identify DE circRNAs between human GC and adjacent non-tumor tissues. Two DE circRNAs were identified, including hsa_circ_0000332 (111 nt), whose parent gene is pyruvate carboxylase (*PC*), located at chr11:66626276–66626387, and hsa_circ_0021087 (340 nt), whose parent gene is LIM domain only 1 (*LMO1*), located at chr11:8248521–8252051. Investigations of the biological functions of these 2 circRNAs

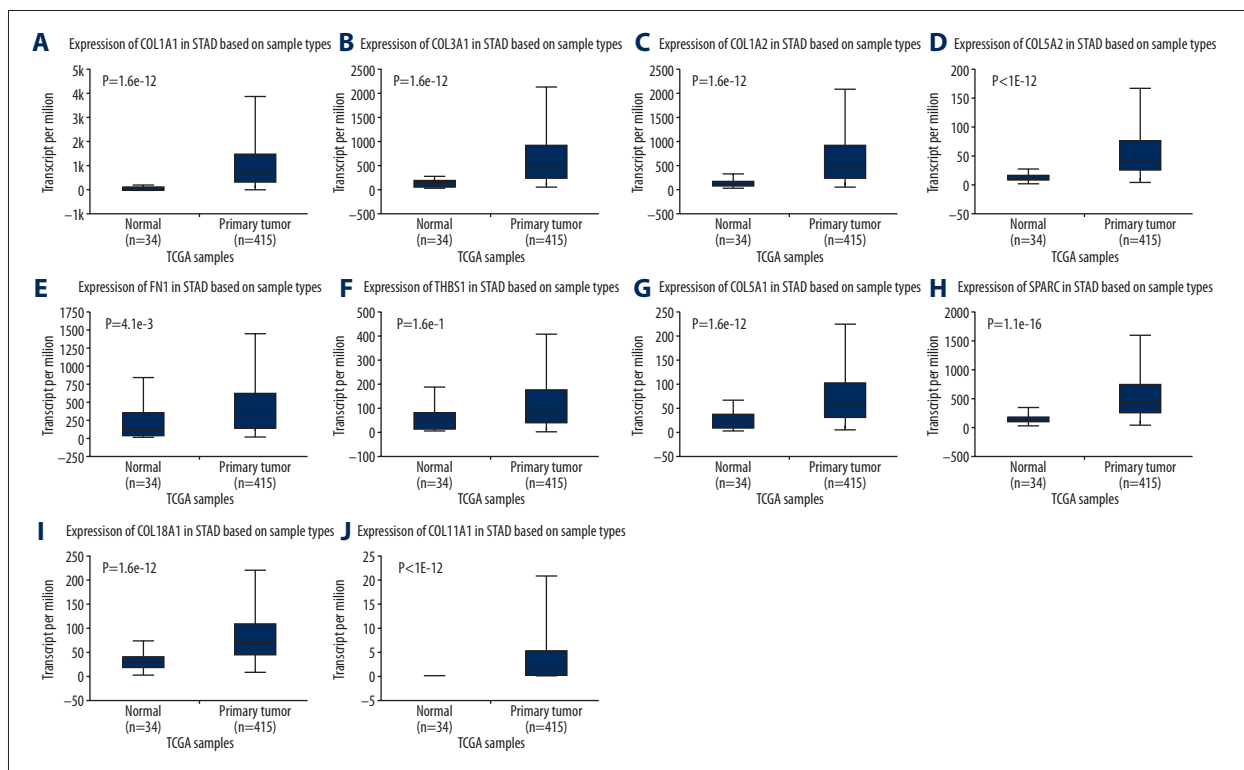
have not been reported. Additionally, to reveal the comprehensive circRNA-related regulatory network in GC, DE circRNA-miRNA-mRNA networks were constructed, showing that hsa_circ_0021087 in particular may play a more critical circRNA role in GC. The construction of circRNA-related regulation networks in GC may provide valuable information for exploration of the potential mechanisms of circRNAs in GC.

The present study was limited by the fact that all results were derived from bioinformatics analysis and lacked corresponding experiments. Functional research is necessary to explore the potential mechanisms of the DEGs and DE circRNAs in GC. The roles of DEGs and DE circRNAs in prognosis and therapy require further investigation in large clinical samples.

Conclusions

Using integrated bioinformatical analysis, we identified 507 DEGs in GC. PPI network analysis identified 10 hub genes: *COL1A1*, *COL3A1*, *COL1A2*, *COL5A2*, *FN1*, *THBS1*, *COL5A1*, *SPARC*, *COL18A1*, and *COL11A1*. Kaplan-Meier analysis revealed that 7 of the hub genes were associated with poor OS in GC patients. GO and KEGG analysis indicated that these genes are enriched in a variety of functions and pathways, among which ECM-related pathways may be the most important in GC. Furthermore, we identified 2 DE circRNAs: hsa_circ_0000332 and hsa_circ_0021087. To reveal the comprehensive circRNA-related regulatory network in GC, DE circRNA-miRNA-mRNA networks were constructed. Although further validation of the present results is required, this study provides useful information for exploration of potential biomarkers and targets for the diagnosis, prognosis, and therapy of GC.

Supplementary Files



Supplementary Figure 1. (A–J) Validation of the expression of 10 hub genes in the TCGA database. STAD – stomach adenocarcinoma.

Supplementary Table 1. Number of differentially expressed genes in each dataset.

	Up-regulation	Down-regulation	Total
GSE19826	882	2067	2949
GSE54129	1683	1992	3675
GSE79973	991	2260	3251
Common DEGs	196	311	507

DEGs – differentially expressed genes.

Supplementary Table 2. Common differentially expressed genes among all three gene expression profiles.

Common DEGs	Gene symbol
Up-regulated DEGs	<i>COL1A, COL5A1, COL1A1, INHBA, COL6A3, MFAP2, SPARC, SERPINH1, COL8A1, HOXA13, FAP, COL18A1, COL5A2, COL3A1, CDH11, EDNRA, COL6A2, PRRX1, VCAN, HOXA10, SERPINE2, ADAMTS2, SULF1, FN1, NID2, PDGFRB, BGN, COL12A1, SEPT11, THY1, RAB31, THBS2, COL4A1, COL4A2, LGALS1, TIMP1, FBN1, IGFBP4, BICC1, CTSK, APOC1, HEYL, TGFB1, P4HA3, HOXB7, PLXDC2, BMP1, RCN3, COL10A1, GPNMB, TMEM158, ADAMTS12, GUCY1A3, PDLIM7, ACTN1, SFRP4, EFEMP2, SERPINF1, FSTL1, ASPN, IGFBP7, CLEC11A, AMIGO2, SALL4, SFRP2, CLDN1, CTHRC1, CST1, CDH3, FNDC1, PGF, SPHK1, ALDH1A3, WISP1, GUCY1B3, HES4, MGP, DOK5, C3, RARRES1, FLNA, CEBPB, SULF2, NRP1, ELOVL5, ISLR, LEF1, ZAK, PMEPA1, AEBP1, TNFRSF11B, SVIL, SPP1, GDDP5, OLFML2B, LOX, EMILIN1, PLEKHO1, HTRA1, GPR161, DGKH, ANTXR1, GXYL2, PLAU, FSCN1, SRPX2, IGF2BP3, EPHB2, ANOS1, CEMIP, FBXO32, COMP, COL11A1, THBS4, PLEKHG2, E2F3, ADAM12, DNM1, ITGBL1, MFG8, RIN3, NRP2, CTSL, CPXM1, ENG, NRK, MMP11, S100A10, DPYSL3, FMNL3, MYO1B, FKBP10, CST4, KLHL23, SPOCK1, FGD6, CCDC80, TNFSF4, BAG2, MLLT11, THBS1, TGM2, PLPPR4, CST2, ITGA5, MIR100HG, PLA2G7, VAC14, ZNF423, CRISPLD1, GINS4, NT5DC2, MMP7, FXYD5, GREM1, FOXC2, CLDN3, FAM83D, FJX1, OLR1, DUXAP10, KLK10, HS3ST3A1, ADAMTS6, SLCO2B1, LY6E, BOC, AHNAC2, HOXC10, SNX10, BUB1, DTL, FOXC1, IL32, MYL9, CXCL8, NOX4, CAP2, WNT2, ETV6, LILRB1, PDE3A, RHOJ, TREM1, SERPINE1, YBX2, TPM2, STC2, MXRA7, TREM2, PLEKHA4, MEOX2, CXCL9, SLC1A3, TAGLN, HOXC6</i>
Down-regulated DEGs	<i>MUC5AC, C8orf49, ZNF57, UBL3, MAG1, AKR7A3, PSAPL1, MFSD4A, GATA4, SMIM6, ALDH6A1, DPCR1, TRIM50, NFE2L2, CLIC6, KLK11, FAM107B, TMPRSS2, MUC6, METTL7A, RASSF6, VSIG1, SGPP2, RAB27A, CWH43, WIPF3, ELOVL6, SVIP, PELI2, C1orf116, FMO5, ADTRP, HOMER2, SHROOM3, SMIM5, ELL2, AQP4, EPN3, CYSTM1, PPFIBP2, PBLD, FA2H, RNASE4, CAPN9, ATP11B, BEX5, C1orf210, FAM189A2, VSIG2, PTER, RDH12, CA2, CXCL17, PXMP2, PLPP5, SLC16A7, TPCN2, KCNE2, CLDN18, TMEM220, DGKD, CMTM4, SLC41A2, KCNJ15, SIDT2, ERO1B, DNER, ANKRD22, DHRS7, PAIP2B, CYP3A5, MYZAP, FBXL13, SSTR1, FGD4, FAM20A, CA9, ALAD, RAB27B, SOSTDC1, PDGFD, LIPH, TMED6, ALDH3A1, UBE2QL1, TMEM92, MECOM, ARHGFE28, RNASE1, SULT1B1, PDIA2, CPA2, ADAM28, VSTM2A, ADGRG2, PLLP, TMEM171, TMEM116, KLHD7A, LDHD, ADRB2, SH3BGL2, TFF1, EPB41L4B, PGC, IRS4, CNTN3, B4GALNT3, ATP4A, ARHGFE37, GCNT2, PPID, SGK2, CCKBR, MAL, ATP4B, RASSF10, SLC26A9, CYP3A43, HPGD, NEDD4L, NECTIN3, PTPRZ1, AKR1C3, IGFBP2, FAM46C, SHISA6, MUM1L1, HMG2N2P46, FBP2, ANG, KIAA1804, SPINK2, GRAMD1C, SCGB2A1, GUCA2B, C16orf89, GKN1, AKR1B10, KCNJ16, PPP1R36, GATA6-AS1, ALDOB, SST, PCAT18, LIPF, GATA5, TMEM125, NTN4, ENPP6, GPT2, BCL2L15, TFF2, IQGAP2, PTPRN2, SYTL2, RPS6KA6, KIAA1324, PSCA, OASL, GABRB3, ASPA, SCNN1G, SLC7A8, SMPD3, LYPD6B, CYP2C18, ALDH1A1, PLAC8, AHCYL2, GIF, ASCL1, ESRRG, RAB11B-AS1, CHGA, PIGR, ACVR1C, NOSTRIN, SPINK7, ADH7, TSPAN12, ACKR4, GKN2, NRG4, PKIB, CAPN13, RNF128, NKX2-3, ATP8A1, ZBTB7C, SH3RF2, SCNN1B, GDDP1, SYTL5, TEX9, UGT2B15, HMGNS5, B3GNT6, SULT2A1, GOLM1, PIWIL2, ACER2, CYP2C9, SOWAHA, ST6GALNAC1, GSTA4, ODAM, PRR15L, TMEM266, SCGN, SLC28A2, TMEM27, SELENBP1, ANKRD29, ATP13A4, HAPLN1, MAMDC2, C4orf19, REG1A, UPK1B, SCN3A, AMPD1, SLC4A4, FAM3B, SMIM24, ANXA10, IL1R2, AKR1C1, SLC14A1, HHIP, CTSE, LRRC66, KCNMB2, STYK1, GCNT1, NDOR1, IRX3, PRDM16, PER3, ZNF793-AS1, TCEA3, SCIN, NR3C2, DUSP4, SLC26A7, UNC5B-AS1, OSBPL7, XK, ARL14, FER1L4, MAP7D2, TMEM38A, ADAMTS15, CAPN8, TPH1, GATA6, SULT1C2, NQO1, KLF4, ITPKA, DDX60, SLC44A4, KBTBD12, RHBDL2, MUC1, KAZALD1, MST1L, DUOX2, EYA2, NEUROD1, TFCP2L1, PDZD3, PIK3C2G, RMST, JCHAIN, STK32B, HRASLS2, RFX6, TCAIM, GAS2, MRAP2, CYT1P, SERPINA4, CDHR3, GCKR, POU2AF1, RGS7, CYP2C19, HGD, SPINK1, PRSS12, PRLR, APOBEC1, TRPA1, GSTA1, NCKAP5, TRIQK, LRRC31, STX19, KRT20, BMP5, ATP2A3, LTF, IGH, ADH1C, IL1RN, LDLR, MGAM, TOX3, ABCC13, FCRL5</i>

DEGs – differentially expressed genes.

Supplementary Table 3. Gene ontology annotation of differentially expressed genes in gastric cancer.

Term	ID	P-value	Count
Top 10 terms in BP category			
Collagen fibril organization	GO: 0030199	1.04E-10	12
Cell adhesion	GO: 0007155	9.74E-07	19
Angiogenesis	GO: 0001525	3.16E-06	16
Endodermal cell differentiation	GO: 0035987	5.32E-05	7
Extracellular matrix organization	GO: 0030198	1.15E-04	10
Negative regulation of angiogenesis	GO: 0016525	2.75E-04	8
Regulation of cell growth	GO: 0001558	3.59E-04	7
Patterning of blood vessels	GO: 0001569	5.67E-04	6
Embryonic limb morphogenesis	GO: 0030326	0.001311	6
Cellular response to BMP stimulus	GO: 0071773	0.001439	4
Top 10 terms in CC category			
Extracellular space	GO: 0005615	4.06E-20	79
Proteinaceous extracellular matrix	GO: 0005578	2.82E-15	30
Collagen trimer	GO: 0005581	1.53E-10	15
Extracellular exosome	GO: 0070062	3.76E-10	104
Extracellular matrix	GO: 0031012	1.98E-06	15
Basement membrane	GO: 0005604	5.98E-06	10
Cell surface	GO: 0009986	1.81E-04	22
Extracellular region	GO: 0005576	0.003658	25
Fibril	GO: 0043205	0.003947	3
Endoplasmic reticulum	GO: 0005783	0.005962	24
Top 10 terms in MF category			
Extracellular matrix structural constituent	GO: 0005201	2.84E-08	11
Heparin binding	GO: 0008201	7.82E-07	15
Extracellular matrix binding	GO: 0050840	1.76E-05	7
Calcium ion binding	GO: 0005509	5.56E-05	33
Integrin binding	GO: 0005178	7.75E-05	7
Protein binding	GO: 0005515	0.001686	13
Serine-type endopeptidase inhibitor activity	GO: 0004867	0.001864	9
Aldehyde dehydrogenase (NAD) activity	GO: 0004029	0.005447	4
Collagen binding	GO: 0005518	0.006676	4
Identical protein binding	GO: 0042802	0.009949	10

BP – biological process; CC – cellular component; MF – molecular function.

Supplementary Table 4. Significantly enriched pathway terms of differentially expressed genes in gastric cancer.

Term	Database	ID	P-value	Count
Extracellular matrix organization	Reactome	R-HSA-1474244	2.06E-33	45
Degradation of the extracellular matrix	Reactome	R-HSA-1474228	7.14E-22	26
Collagen formation	Reactome	R-HSA-1474290	2.36E-18	20
Collagen biosynthesis and modifying enzymes	Reactome	R-HSA-1650814	9.37E-18	18
Collagen degradation	Reactome	R-HSA-1442490	9.00E-17	17
Assembly of collagen fibrils	Reactome	R-HSA-2022090	1.69E-16	16
Integrin cell surface interactions	Reactome	R-HSA-216083	3.51E-16	18
ECM proteoglycans	Reactome	R-HSA-3000178	7.22E-16	17
Chemical carcinogenesis	KEGG PATHWAY	hsa05204	6.76E-13	15
Integrin signalling pathway	PANTHER	P00034	9.35E-13	19

Supplementary Table 5. Significantly enriched pathway terms of three significant modules in protein–protein interaction networks.

Term	Database	ID	P-value	Count
Module 1				
Collagen biosynthesis and modifying enzymes	Reactome	R-HSA-1650814	2.76E-44	16
Collagen formation	Reactome	R-HSA-1474290	1.42E-42	16
Collagen degradation	Reactome	R-HSA-1442490	4.31E-37	14
Extracellular matrix organization	Reactome	R-HSA-1474244	1.01E-34	16
Assembly of collagen fibrils	Reactome	R-HSA-2022090	1.42E-34	13
Module 2				
Platelet degranulation	Reactome	R-HSA-114608	9.13E-15	6
Response to elevated platelet cytosolic Ca ²⁺	Reactome	R-HSA-76005	1.14E-14	6
Hemostasis	Reactome	R-HSA-109582	1.98E-13	7
Platelet activation, signaling and aggregation	Reactome	R-HSA-76002	8.57E-13	6
Extracellular matrix organization	Reactome	R-HSA-1474244	4.51E-10	5
Module 3				
O-linked glycosylation of mucins	Reactome	R-HSA-913709	6.69E-12	5
Chemical carcinogenesis	KEGG PATHWAY	hsa05204	2.04E-11	5
O-linked glycosylation	Reactome	R-HSA-5173105	8.07E-11	5
Metabolic pathways	KEGG PATHWAY	hsa01100	1.42E-10	8
Biological oxidations	Reactome	R-HSA-211859	1.49E-09	5

Supplementary Table 6. Number of differentially expressed circRNAs from each expression profile dataset.

	Up-regulation	Down-regulation	Total
GSE83521	80	70	150
GSE89143	3	188	191
GSE93541	215	201	416
Common DE circRNAs	0	2	2

DE circRNAs – differentially expressed circRNAs.

Supplementary Table 7. Significantly enriched pathway terms of genes in the DE circRNA–miRNA–mRNA networks.

Term	Database	ID	P-value	Count
Signal transduction	Reactome	R-HSA-162582	0.000188	10
Extracellular matrix organization	Reactome	R-HSA-1474244	0.000262	4
Non-integrin membrane-ECM interactions	Reactome	R-HSA-3000171	0.001912	2
Collagen biosynthesis and modifying enzymes	Reactome	R-HSA-1650814	0.002438	2
Bile secretion	KEGG PATHWAY	hsa04976	0.002724	2
SLC-mediated transmembrane transport	Reactome	R-HSA-425407	0.002854	3
ECM-receptor interaction	KEGG PATHWAY	hsa04512	0.003587	2
TGF-beta signaling pathway	KEGG PATHWAY	hsa04350	0.003756	2
Integrin cell surface interactions	Reactome	R-HSA-216083	0.003842	2
Collagen formation	Reactome	R-HSA-1474290	0.004105	2

References:

- Ren J, Niu G, Wang X et al: Effect of age on prognosis of gastric signet-ring cell carcinoma: A SEER database analysis. *Med Sci Monit*, 2018; 24: 8524–32
- Siegel RL, Miller KD, Jemal A: Cancer statistics, 2018. *Cancer J Clin*, 2018; 68(1): 7–30
- Song Z, Wu Y, Yang J et al: Progress in the treatment of advanced gastric cancer. *Tumour Biol*, 2017; 39(7): 1010428317714626
- Chen W, Lu C, Hong J: TRIM15 exerts anti-tumor effects through suppressing cancer cell invasion in gastric adenocarcinoma. *Med Sci Monit*, 2018; 24: 8033–41
- Liu W, Ouyang S, Zhou Z et al: Identification of genes associated with cancer progression and prognosis in lung adenocarcinoma: Analyses based on microarray from Oncomine and The Cancer Genome Atlas databases. *Mol Genet Genomic Med*, 2019; 7(2): e00528
- Cao L, Chen Y, Zhang M et al: Identification of hub genes and potential molecular mechanisms in gastric cancer by integrated bioinformatics analysis. *Peer J*, 2018; 6: e5180
- Zhu Q, Sun Y, Zhou Q et al: Identification of key genes and pathways by bioinformatics analysis with TCGA RNA sequencing data in hepatocellular carcinoma. *Mol Clin Oncol*, 2018; 9(6): 597–606
- Yan P, He Y, Xie K et al: In silico analyses for potential key genes associated with gastric cancer. *Peer J*, 2018; 6: e6092
- Shao Y, Li J, Lu R et al: Global circular RNA expression profile of human gastric cancer and its clinical significance. *Cancer Med*, 2017; 6(6): 1173–80
- Li J, Ni S, Zhou C, Ye M: The expression profile and clinical application potential of hsa_circ_0000711 in colorectal cancer. *Cancer Manag Res*, 2018; 10: 2777–84
- Huang YS, Jie N, Zou KJ, Weng Y: Expression profile of circular RNAs in human gastric cancer tissues. *Mol Med Rep*, 2017; 16(3): 2469–76
- Li P, Chen H, Chen S et al: Circular RNA 0000096 affects cell growth and migration in gastric cancer. *Br J Cancer*, 2017; 116(5): 626–33
- Oliveros JC: Venny. An interactive tool for comparing lists with Venn's diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html>
- Ashburner M, Ball CA, Blake JA et al: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, 2000; 25(1): 25–29
- Kanehisa M, Goto S: KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 2000; 28(1): 27–30
- Huang DW, Sherman BT, Tan Q et al: The DAVID Gene Functional Classification Tool: A novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol*, 2007; 8(9): R183
- Xie C, Mao X, Huang J et al: KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res*, 2011; 39(Web Server issue): W316–22
- Szklarczyk D, Franceschini A, Wyder S et al: STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*, 2015; 43(Database issue): D447–52
- Shannon P, Markiel A, Ozier O et al: Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res*, 2003; 13(11): 2498–504
- Chandrashekar DS, Bashel B, Balasubramanya SAH et al: UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia*, 2017; 19(8): 649–58
- Lánczky A, Nagy Á, Bottai G et al: miRpower: A web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res Treat*, 2016; 160(3): 439–46
- Dudekula DB, Panda AC, Grammatikakis I et al: CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA Biol*, 2016; 13(1): 34–42
- Dweep H, Gretz N: miRWalk2.0: A comprehensive atlas of microRNA-target interactions. *Nat Methods*, 2015; 12(8): 697

24. Chia NY, Tan P: Molecular classification of gastric cancer. *Ann Oncol*, 2016; 27(5): 763–69
25. Liu X, Xu Y, Zhou Q et al: PI3K in cancer: Its structure, activation modes and role in shaping tumor microenvironment. *Future Oncol*, 2018; 14(7): 665–74
26. Su Y, Zhang M, Zhang L et al: Construction of an miRNA-mRNA regulatory network in colorectal cancer with bioinformatics methods. *Anticancer Drugs*, 2018 [Epub ahead of print]
27. Alanko J, Ivaska J: Endosomes: Emerging platforms for integrin-mediated FAK signalling. *Trends Cell Biol*, 2016; 26(6): 391–98
28. Islam MS, Ciavattini A, Petraglia F et al: Extracellular matrix in uterine leiomyoma pathogenesis: A potential target for future therapeutics. *Hum Reprod Update*, 2018; 24(1): 59–85
29. Ricard-Blum S: The collagen family. *Cold Spring Harb Perspect Biol*, 2011; 3(1): a004978
30. Wang Q, Yu J: MiR-129-5p suppresses gastric cancer cell invasion and proliferation by inhibiting COL1A1. *Biochem Cell Biol*, 2018; 96(1): 19–25
31. Zhuo C, Li X, Zhuang H et al: Elevated THBS2, COL1A2, and SPP1 expression levels as predictors of gastric cancer prognosis. *Cell Physiol Biochem*, 2016; 40(6): 1316–24
32. Li J, Ding Y, Li A: Identification of COL1A1 and COL1A2 as candidate prognostic factors in gastric cancer. *World J Surg Oncol*, 2016; 14(1): 297
33. Zhang Z, Wang Y, Zhang J et al: COL1A1 promotes metastasis in colorectal cancer by regulating the WNT/PCP pathway. *Mol Med Rep*, 2018; 17(4): 5037–42
34. Ao R, Guan L, Wang Y, Wang JN: Silencing of COL1A2, COL6A3, and THBS2 inhibits gastric cancer cell proliferation, migration, and invasion while promoting apoptosis through the PI3k-Akt signaling pathway. *J Cell Biochem*, 2018; 119(6): 4420–34
35. Yuan L, Shu B, Chen L et al: Overexpression of COL3A1 confers a poor prognosis in human bladder cancer identified by co-expression analysis. *Oncotarget*, 2017; 8(41): 70508–20
36. Long H, Liang C, Zhang X et al: Prediction and analysis of key genes in glioblastoma based on bioinformatics. *Biomed Res Int*, 2017; 2017: 7653101
37. Chen Z, Soutto M, Rahman B et al: Integrated expression analysis identifies transcription networks in mouse and human gastric neoplasia. *Genes Chromosomes Cancer*, 2017; 56(7): 535–47
38. Boguslawska J, Kedzierska H, Poplawski P et al: Expression of genes involved in cellular adhesion and extracellular matrix remodeling correlates with poor survival of patients with renal cancer. *J Urol*, 2016; 195(6): 1892–902
39. Chai F, Liang Y, Zhang F et al: Systematically identify key genes in inflammatory and non-inflammatory breast cancer. *Gene*, 2016; 575(2 Pt 3): 600–14
40. Li G, Li X, Yang M et al: Prediction of biomarkers of oral squamous cell carcinoma using microarray technology. *Sci Rep*, 2017; 7: 42105
41. Liu W, Wei H, Gao Z et al: COL5A1 may contribute the metastasis of lung adenocarcinoma. *Gene*, 2018; 665: 57–66
42. Zhao X, Cai H, Wang X, Ma L: Discovery of signature genes in gastric cancer associated with prognosis. *Neoplasma*, 2016; 63(2): 239–45
43. Wu D, Chen K, Bai Y et al: Screening of diagnostic markers for osteosarcoma. *Mol Med Rep*, 2014; 10(5): 2415–20
44. Zeng XT, Liu XP, Liu TZ, Wang XH: The clinical significance of COL5A2 in patients with bladder cancer: A retrospective analysis of bladder cancer gene expression data. *Medicine (Baltimore)*, 2018; 97(10): e0091
45. Wang Y: Transcriptional regulatory network analysis for gastric cancer based on mRNA microarray. *Pathol Oncol Res*, 2017; 23(4): 785–91
46. Fang ZQ, Zang WD, Chen R et al: Gene expression profile and enrichment pathways in different stages of bladder cancer. *Genet Mol Res*, 2013; 12(2): 1479–89
47. Qiu T, Wang H, Wang Y et al: Identification of genes associated with melanoma metastasis. *Kaohsiung J Med Sci*, 2015; 31(11): 553–61
48. Park H, Lee Y, Lee H et al: The prognostic significance of cancer-associated fibroblasts in pancreatic ductal adenocarcinoma. *Tumour Biol*, 2017; 9(10): 1010428317718403
49. Gundewar C, Sasor A, Hilmersson KS et al: The role of SPARC expression in pancreatic cancer progression and patient survival. *Scand J Gastroenterol*, 2015; 50(9): 1170–49
50. Che Y, Luo A, Wang H et al: The differential expression of SPARC in esophageal squamous cell carcinoma. *Int J Mol Med*, 2006; 17(6): 1027–33
51. Wang J, Gao P, Song Y et al: Prognostic value of gastric cancer-associated gene signatures: Evidence based on a meta-analysis using integrated bioinformatics methods. *J Cell Mol Med*, 2018; 22(11): 5743–47
52. Said N, Frierson HF, Sanchez-Carbayo M et al: Loss of SPARC in bladder cancer enhances carcinogenesis and progression. *J Clin Invest*, 2013; 123(2): 751–66
53. DiMartino JF, Lacayo NJ, Varadi M et al: Low or absent SPARC expression in acute myeloid leukemia with MLL rearrangements is associated with sensitivity to growth inhibition by exogenous SPARC protein. *Leukemia*, 2006; 20(3): 426–32
54. Xiao J, Yang W, Xu B et al: Expression of fibronectin in esophageal squamous cell carcinoma and its role in migration. *BMC Cancer*, 2018; 18(1): 976
55. Zhang H, Sun Z, Li Y et al: MicroRNA-200c binding to FN1 suppresses the proliferation, migration and invasion of gastric cancer cells. *Biomed Pharmacother*, 2017; 88: 285–92
56. Huang T, Wang L, Liu D et al: FGF7/FGFR2 signal promotes invasion and migration in human gastric cancer through upregulation of thrombospondin-1. *Int J Oncol*, 2017; 50(5): 1501–12
57. Roberts DD: Regulation of tumor growth and metastasis by thrombospondin-1. *FASEB J*, 1996; 10(10): 1183–91
58. Wang S, Li Z, Xu R: Human cancer and platelet interaction, a potential therapeutic target. *Int J Mol Sci*, 2018; 19(4): pii: E1246
59. Gasic GJ, Gasic TB, Galanti N et al: Platelet-tumor-cell interactions in mice. The role of platelets in the spread of malignant disease. *Int J Cancer*, 1973; 11(3): 704–18
60. Mikami J, Kurokawa Y, Takahashi T et al: Antitumor effect of antiplatelet agents in gastric cancer cells: An *in vivo* and *in vitro* study. *Gastric Cancer*, 2016; 19(3): 817–26
61. Saito H, Fushida S, Miyashita T et al: Potential of extravasated platelet aggregation as a surrogate marker for overall survival in patients with advanced gastric cancer treated with preoperative docetaxel, cisplatin and S-1: A retrospective observational study. *BMC Cancer*, 2017; 17(1): 294
62. Sun H, Xi P, Sun Z et al: Circ-SFMBT2 promotes the proliferation of gastric cancer cells through sponging miR-182-5p to enhance CREB1 expression. *Cancer Manag Res*, 2018; 10: 5725–34
63. Zhong S, Wang J, Hou J et al: Circular RNA hsa_circ_0000993 inhibits metastasis of gastric cancer cells. *Epigenomics*, 2018; 10(10): 1301–13