



# Construction of a novel ceRNA network and identification of lncRNA *ADAMTS9-AS2* and *PVT1* as hub regulators of miRNA and coding gene expression in gastric cancer

Song Wang<sup>1</sup>, Xing-Chuan Li<sup>2,3</sup>, Jia-Rui Zhu<sup>4</sup>, Zhi-Jie Ma<sup>5</sup>, Jun-Tao Ran<sup>1</sup>, Yong-Ning Zhou<sup>2,3</sup>

<sup>1</sup>Department of Radiotherapy, The First Hospital of Lanzhou University, Lanzhou, China; <sup>2</sup>Department of Gastroenterology, The First Hospital of Lanzhou University, Lanzhou, China; <sup>3</sup>Key Laboratory for Gastrointestinal Diseases of Gansu Province, Lanzhou University, Lanzhou, China; <sup>4</sup>Cuiying Biomedical Research Center, Lanzhou University Second Hospital, Lanzhou, China; <sup>5</sup>Department of Anesthesiology, Lanzhou University Second Hospital, Lanzhou, China

**Contributions:** (I) Conception and design: XC Li, S Wang, YN Zhou; (II) Administrative support: JT Ran, YN Zhou; (III) Provision of study materials or patients: XC Li, JR Zhu; (IV) Collection and assembly of data: XC Li, ZJ Ma; (V) Data analysis and interpretation: XC Li, S Wang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

**Correspondence to:** Jun-Tao Ran, Department of Radiotherapy, The First Hospital of Lanzhou University, Donggang West Road No. 1, Lanzhou 730000, China. Email: ranjt@lzu.edu.cn.

**Background:** Studies on the interactions of single long non-coding RNA, microRNA, and mRNA have many limitations; therefore, it is necessary to study the complex regulatory network of gastric cancer (GC) pathogenesis systematically.

**Methods:** In this study, gene and miRNA expression data for GC were downloaded from The Cancer Genome Atlas and used for transcriptome profiling, differential gene analysis, and construction of an lncRNA-miRNA-mRNA regulatory network in conjunction with an online database to identify the key genes and subnetworks in GC pathogenesis. Real-time quantitative polymerase chain reaction was used to detect the expression of hub lncRNAs in 54 paired GC and matched normal mucosal tissues.

**Results:** We constructed an lncRNA-miRNA-mRNA competitive endogenous RNA regulatory network containing 1,626 network nodes and 2,704 interactions. lncRNA *ADAMTS9-AS2* and *PVT1* were identified as key node genes in this competitive endogenous RNA network. Quantitative reverse transcription-polymerase chain reaction revealed *ADAMTS9-AS2* downregulation and *PVT1* upregulation in 54 pairs of GC and normal tissues adjacent to the cancer tissues.

**Conclusions:** This study systematically analysed the lncRNA-miRNA-mRNA regulatory network in GC and identified *ADAMTS9-AS2* and *PVT1* as key regulatory genes in this network, providing new understanding of GC pathogenesis and insights for its early diagnosis and treatment.

**Keywords:** Stomach neoplasm; long noncoding RNA; gene regulatory networks; *ADAMTS9-AS2*; *PVT1*

Submitted Jul 10, 2020. Accepted for publication Dec 04, 2020.

doi: 10.21037/tcr-20-2168

**View this article at:** <http://dx.doi.org/10.21037/tcr-20-2168>

## Introduction

Gastric cancer (GC), a malignant disease with the highest incidence among digestive tract, is the second leading cause of cancer-related mortality worldwide (1). GC prognosis depends primarily on tumour stage. Patients with GC usually present with advanced stages of the disease at initial

diagnosis, and the 5-year survival rate of these patients is significantly lower than that of patients with early diagnosis (2). Early detection and diagnosis are key to reducing GC-related mortality, and thus, new clinical biomarkers for early tumour diagnosis, monitoring treatment response, and predicting prognosis are urgently needed.

Long non-coding RNA (lncRNA) are molecules with lengths of 200 or more nucleotides. Accumulating research evidence show that lncRNAs play important regulatory roles in physiological and pathological processes as well as in cancer (3). Abnormal changes in the expression of lncRNAs have been found in almost all types of malignant tumours, indicating that they play important roles in the occurrence and development of tumours. Owing to their high tissue specificity, lncRNAs are stably expressed in various body fluids such as blood, urine, saliva, and interstitial fluid, and have broad prospects in the diagnosis, treatment, and prognosis of GC. MicroRNAs (miRNAs) is a class of small non-coding RNAs with lengths of 20–22 nucleotides and contain highly conserved sequences. MiRNAs are ubiquitous in multicellular organisms and participate in almost all life processes (4). MiRNAs are abnormally expressed in various cancers, including breast cancer, colon cancer, chronic lymphocytic leukaemia, lung cancer, and glioma (5–8). Abnormal expression of miR-158a, miR-27b, and miR-326 can predict the risk of metastasis in patients with colon cancer (9,10). The expression levels of miR-125b and miR-155 in the sera of patients with breast cancer can be used to evaluate the sensitivity of the patients to chemotherapy and determine their prognosis (11,12). MiRNA can affect target gene expression at both transcriptional and post-transcriptional levels and are involved in the occurrence and development of many tumours.

Abnormally expressed lncRNAs can regulate miRNAs by acting as competing endogenous RNA (ceRNA), indirectly inhibiting target genes and important signal pathways and affecting GC cell growth, metastasis, metabolism, and apoptosis (13–16). ceRNA networks are important for the interaction between lncRNA and miRNA. ceRNAs compete with miRNA response elements (MREs) for the same miRNA, thereby regulating the expression of target gene transcripts. In the regulatory mechanism of GC, the interactions between lncRNAs, miRNAs, and mRNAs are extremely complex and collectively form a complex regulatory network. Nevertheless, studying the interactions of single lncRNAs, miRNAs, and mRNAs has limitations, and it is necessary to study the complex regulatory network of GC pathogenesis systematically. Although researchers have constructed ceRNA networks of GC (17–19), differences in sample selection, gene selection criteria, and network databases result in the inconsistent of differential expressed genes, constructed networks and key genes. Furthermore, because lncRNA-mRNA interactions have not been extensively studied, the constructed ceRNA

networks were mainly based on lncRNA-miRNA and miRNA-mRNA interactions.

Considering these limitations, this study aimed to construct an lncRNA-miRNA-mRNA regulatory network based on the ceRNA hypothesis using data from The Cancer Genome Atlas (TCGA) public database and other online databases to integrate the interactions between differentially expressed lncRNAs, miRNAs, and mRNAs in GC, identify the key lncRNAs, and validate these results in GC and paired normal samples. The number of interactions and survival analyses were used to identify the key lncRNAs. The biological functions of subnetworks comprising lncRNAs and their target genes and their effects on prognosis were examined to understand GC pathogenesis from a systematic perspective. The expression of hub lncRNAs in GC and paired normal samples was validated using quantitative real-time polymerase chain reaction (qRT-PCR). We present the following article in accordance with the MDAR checklist (available at <http://dx.doi.org/10.21037/tcr-20-2168>).

## Methods

### *Ethical statement*

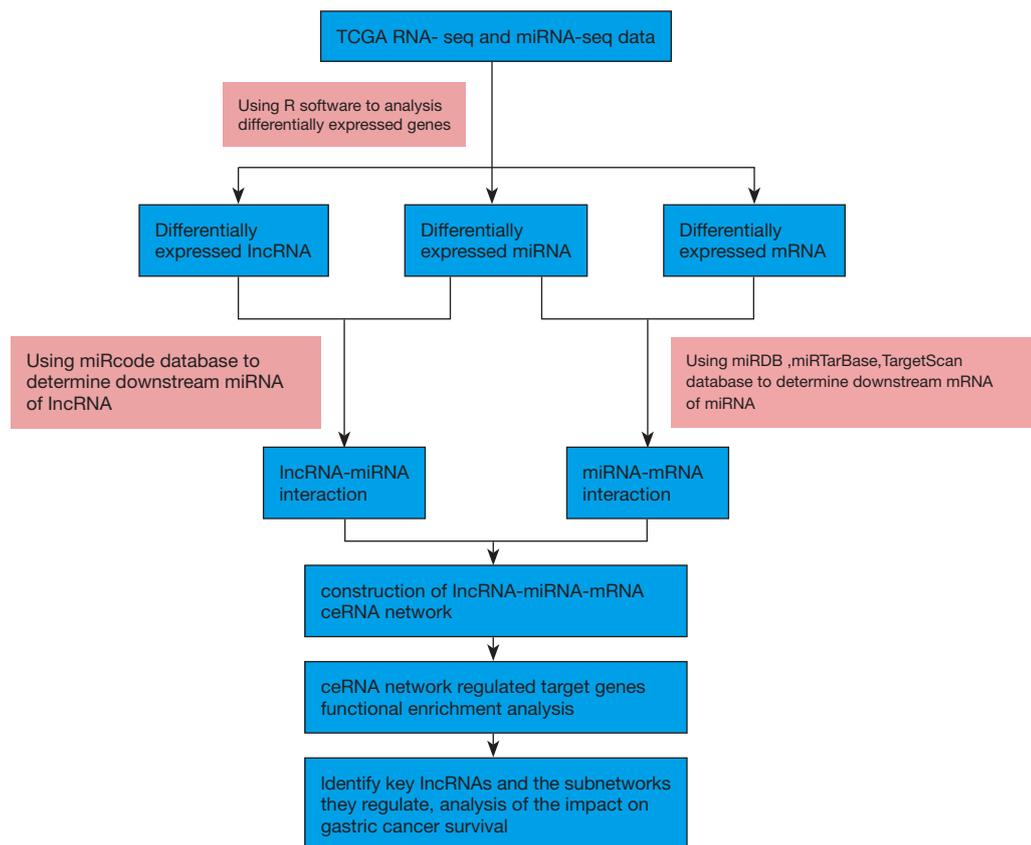
The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of Lanzhou University Second Hospital (2020A-031) and informed consent was taken from all the patients.

### *Organization of TCGA expression and clinical data*

Transcriptome profiling of gene and miRNA expression quantification data, as well as sample clinical information for GC, were downloaded from TCGA (<https://cancergenome.nih.gov/>). Overall, the transcriptome RNA data (mRNA and lncRNA gene expression data) from 343 GC tissue samples and 30 normal adjacent cancer tissue samples were included. miRNA expression data included 410 GC tissue samples and 42 normal adjacent cancer tissue samples. Simultaneously, the survival data of patients with GC were downloaded and collected. The data processing flowchart is shown in *Figure 1*.

### *Quantitative real-time PCR*

In total, 54 GC tissues and matched non-tumour gastric



**Figure 1** Flowchart representing analysis of TCGA data for gastric cancer. TCGA, The Cancer Genome Atlas.

tissues were collected from patients who underwent gastrectomy without radiotherapy and chemotherapy at the Lanzhou University Second Hospital from November to December 2019. The diagnosis of GC was made by two experienced pathologists. Samples were snap-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  before RNA isolation.

The qRT-PCR laboratory protocol was based on a previously described procedure (16). According to manufacturer's instructions, total RNA was extracted and used for cDNA synthesis using TRIzol (Invitrogen, CA, USA) and Takara PrimeScript RT reagent kit. Then, qRT-PCR was performed on a LightCycler 480 Real-time PCR system (Roche, Shanghai, China) using the following specific primers: ADAMTS9-AS2, forward, 5'-TTTACCATGCGCTGAGTGAG-3', reverse, 5'-AAAGTTGCGTCATGCTTCGG-3'; PVT1 forward, 5'-GCUUGGAGGCUGAGGAGUUTT-3'; reverse, 5'-AACUCCUCAGCCUCCAAGCTT-3'; GAPDH (internal control), forward,

5'-AGTGGCAAAGTGGAGATT-3'; reverse, 5'-GTGGAGTCATACTGGAACA-3'. Gene expression in each sample was measured thrice, and the average value was used as the gene expression value of the samples.

#### *Differentially expressed RNA analysis*

The mRNA and miRNA expression data count values of each sample were checked. Values  $<1$  were considered missing values. If multiple probes corresponded to the same RNA, the average count value of multiple probe expressions was considered. The gene expression levels were then normalised and analysed using the limma package and EdgeR package in R (version 3.6.1). If the genes (including lncRNA, miRNA and mRNA) from the GC and adjacent normal tissue groups satisfied  $|\log_2\text{foldChange}| > 1$  and  $P < 0.05$ , they were considered to be differentially expressed. The volcano maps and heat maps of differentially expressed genes were plotted using R software.

**Table 1** Count of differentially expressed RNA in The Cancer Genome Atlas (TCGA) gastric cancer dataset

Gene category	Differentially expressed gene count	Upregulation gene count	Downregulation gene count
LncRNA	2,684	1,899	785
miRNA	240	180	60
mRNA	4,285	2,169	2,116

### Construction of ceRNA interaction network

The miRcode database (<http://www.mircode.org/>) was used to determine the interactions between differentially expressed lncRNAs and miRNAs. The miRNA-mRNA interaction relationship was examined using miRDB (<http://www.mirdb.org/>), miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw/>), and TargetScan (<http://www.targetscan.org/>) databases. The above interactions were integrated and imported into Cytoscape software for visualising the ceRNA interaction network.

### Identification of key lncRNAs in the GC lncRNA-miRNA-mRNA regulatory network

The key lncRNAs in the ceRNA network were identified based on the number of downstream miRNAs interacting with lncRNA and their correlation with survival in GC. Univariate Cox survival analysis was used to determine the effect of the key lncRNAs on overall survival (OS) in patients with GC. To strengthen the reliability of the potential prognostic values of the key lncRNAs, the UALCAN (<http://ualcan.path.uab.edu/analysis.html>) and GEPIA (<http://gepia.cancer-pku.cn/>) online tools were used to investigate the influence of the key lncRNAs on the OS of patients with GC.  $P < 0.05$  indicated significant influence on OS in GC.

### Gene function enrichment analysis

To further understand the biological functions of the key genes in the lncRNA-miRNA-mRNA regulatory network, R software was used to perform KEGG (Kyoto Encyclopaedia of Genes and Genomes) signalling pathway and GO (Gene Ontology) function enrichment analysis.  $P < 0.05$  indicated that the enrichment was statistically significant.

### Statistical analysis

Survival curves were drawn using the Kaplan-Meier method and compared using the log-rank test. All tests were two-sided, with  $P < 0.05$  indicating a statistically significant

difference. The data analysis period was from September 1 to November 24, 2019.

## Results

### Differential expression of lncRNAs, miRNAs, and mRNAs

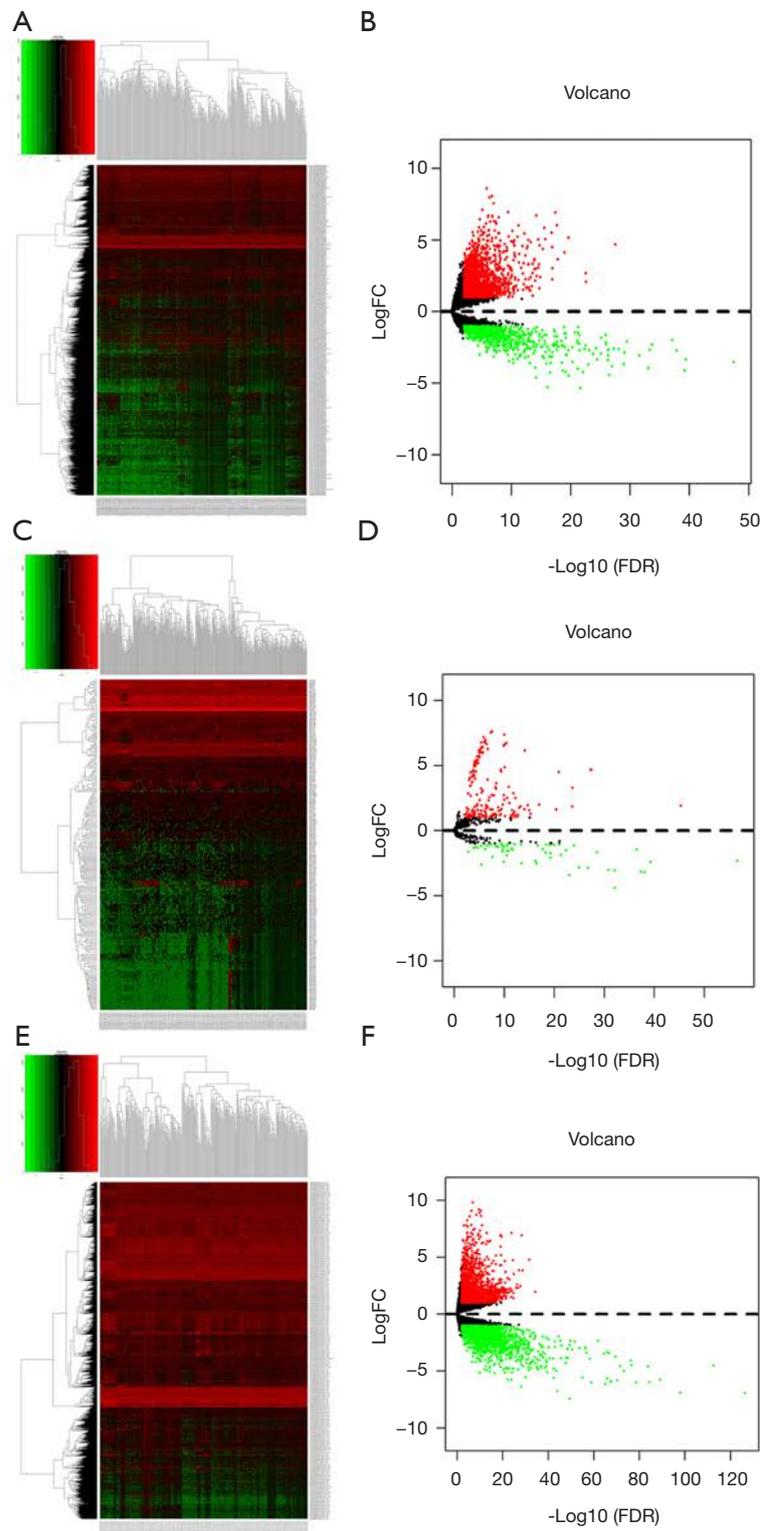
Differential analysis of genes in GC tissues and adjacent normal tissues from TCGA data revealed a total of 2,684 differentially expressed lncRNAs, of which 1,899 were upregulated and 785 were downregulated; 240 differentially expressed miRNAs, of which 180 were upregulated and 60 were downregulated; and 4,285 differentially expressed mRNAs, including 2,169 upregulated and 2,116 downregulated mRNAs (Table 1). The heat maps and volcano maps of differentially expressed lncRNAs, miRNAs, and mRNAs are shown in Figure 2.

### Construction of the lncRNA-miRNA-mRNA regulatory network in GC

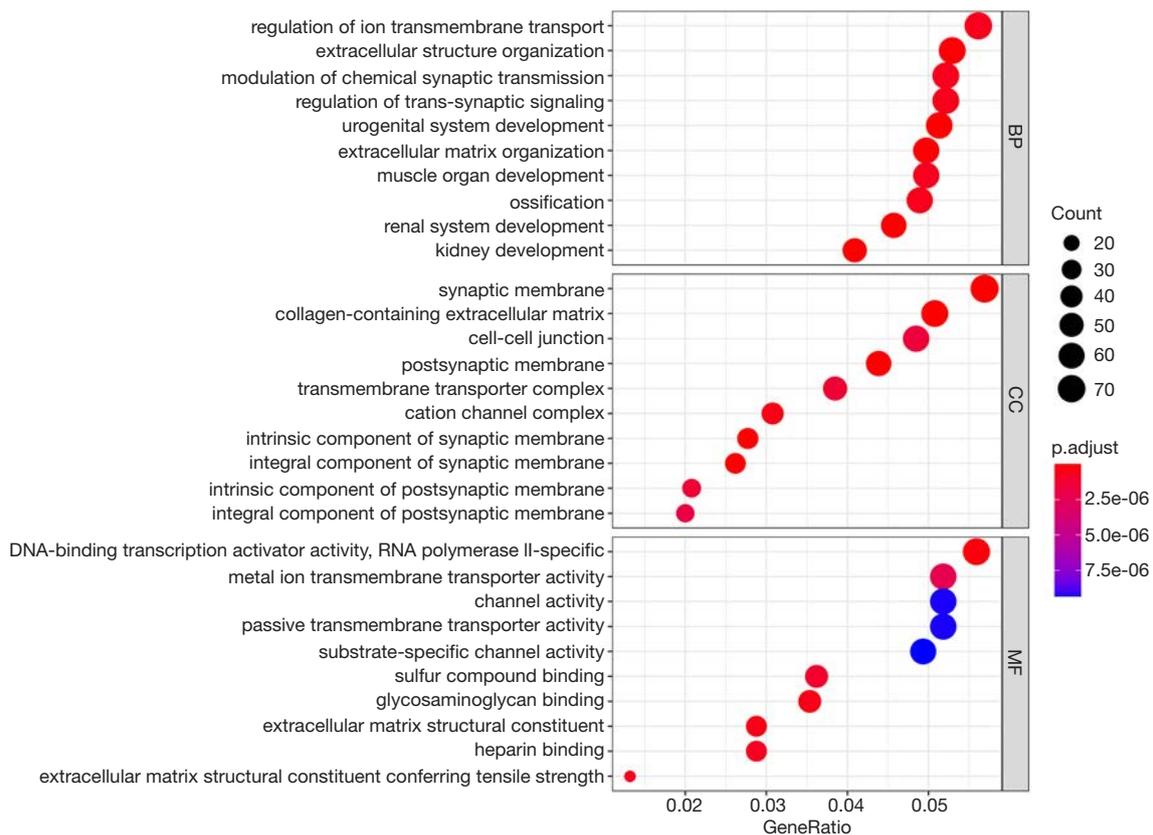
The miRcode database matched 245 differentially expressed lncRNAs with 34 differentially expressed miRNAs, resulting in 1,355 pairs of lncRNA-miRNA interactions. The three databases, miRDB, miRTarBase, and TargetScan, yielded 1,349 pairs of miRNA-mRNA interactions (matching in at least two databases), including 29 differentially expressed miRNAs (database annotated differentially expressed miRNA mature miRNAs) and 1,349 differentially expressed mRNAs. All lncRNA-miRNA relationship pairs and miRNA-mRNA relationships were integrated to construct the GC lncRNA-miRNA-mRNA regulatory network including 245 lncRNAs, 34 miRNAs, 1,349 mRNAs, and 2,704 interactions (Table S1, Figure S1).

### Gene enrichment analysis of GC using the lncRNA-miRNA-mRNA regulatory network

GO and KEGG signal pathway enrichment analyses of the target genes in the GC lncRNA-miRNA-mRNA regulatory



**Figure 2** Heat map and volcano map of differentially expressed lncRNAs, miRNAs, and mRNAs in TCGA gastric cancer tissues and adjacent normal tissues. lncRNA heat map (A), volcano map (B); miRNA heat map (C), volcano map (D); mRNA heat map (E), volcano map (F); red indicates gene upregulation, green indicates gene downregulation. TCGA, The Cancer Genome Atlas.



**Figure 3** Analysis of GO function enrichment in the gastric cancer lncRNA-miRNA-mRNA regulatory network. P<0.05 was considered statistically significant. BP, biological process; MF, molecular function; CC, cellular component; GO: Gene Ontology.

network were performed using R software. The GO enrichment analysis results showed that the target genes of the regulatory network were significantly enriched in the following biological processes (BP): the regulation of ion transmembrane transport, extracellular structure organization, and modulation of chemical synaptic transmission. The enriched molecular function (MF) categories were DNA-binding transcription activator activity, RNA polymerase II-specific, metal ion transmembrane transporter activity, and channel activity. The significantly enriched cell component (CC) categories were synaptic membrane, collagen-containing extracellular matrix, and cell-cell junction (Figure 3). The KEGG pathways were mainly enriched in the PI3K-Akt signalling pathway (Figure 4).

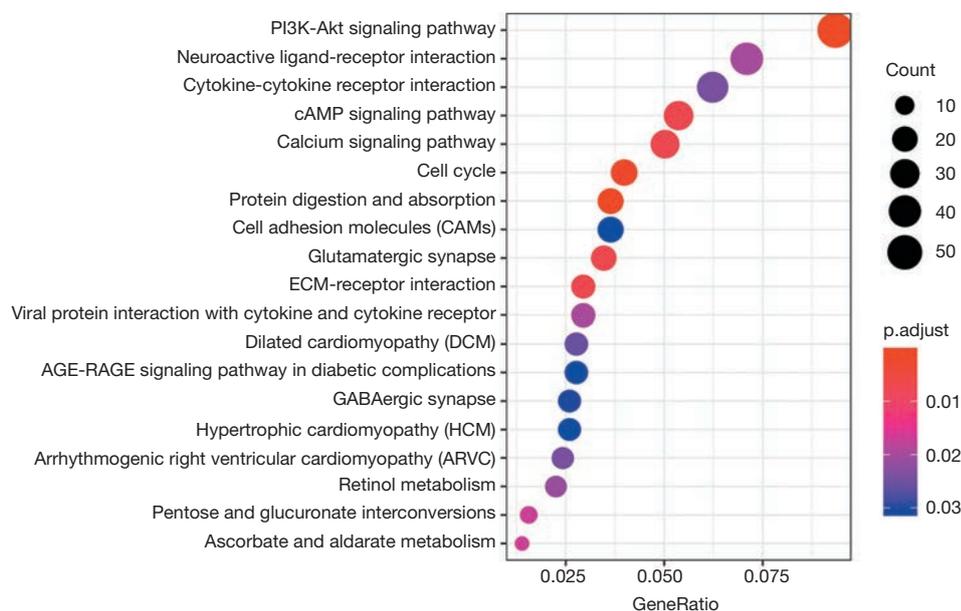
**Key lncRNAs survival analysis**

A total of 12 lncRNAs were selected (Table 2) according to the following hub lncRNA criteria: (I) more than 15 downstream miRNAs interacted with lncRNAs in the GC ceRNA regulatory network and (II) were related to OS in

GC. Univariate Cox survival analysis (Figure 5) showed that *ADAMTS9-AS2* and *PVT1* are related to survival in GC and that patients with highly expressed *ADAMTS9-AS2* had poor prognosis. Although the Ualcan online database did not include survival analysis data for *ADAMTS9-AS2*, these were found in the GEPIA online database. Survival analysis for *ADAMTS9-AS2* in the GEPIA online database indicated poor prognosis. On the contrary, patients with low expression of *PVT1* in GC showed poor prognosis (P<0.05), consistent with the survival analysis results in the Ualcan online database (Figure 5). To further study the role of the lncRNAs *ADAMTS9-AS2* and *PVT1* in the lncRNA-miRNA-mRNA regulatory network of GC, we extracted and analysed these two lncRNAs and their associated regulatory networks.

**ADAMTS9-AS2 related regulatory network**

As shown in Figure 6, *ADAMTS9-AS2* has regulatory relationships with 19 miRNAs and indirect relationships with 974 mRNAs. Functional enrichment analysis of



**Figure 4** Enrichment analysis of KEGG signal pathways in the lncRNA-miRNA-mRNA regulatory network of gastric cancer. P<0.05 was considered statistically significant. KEGG, Kyoto Encyclopaedia of Genes and Genomes.

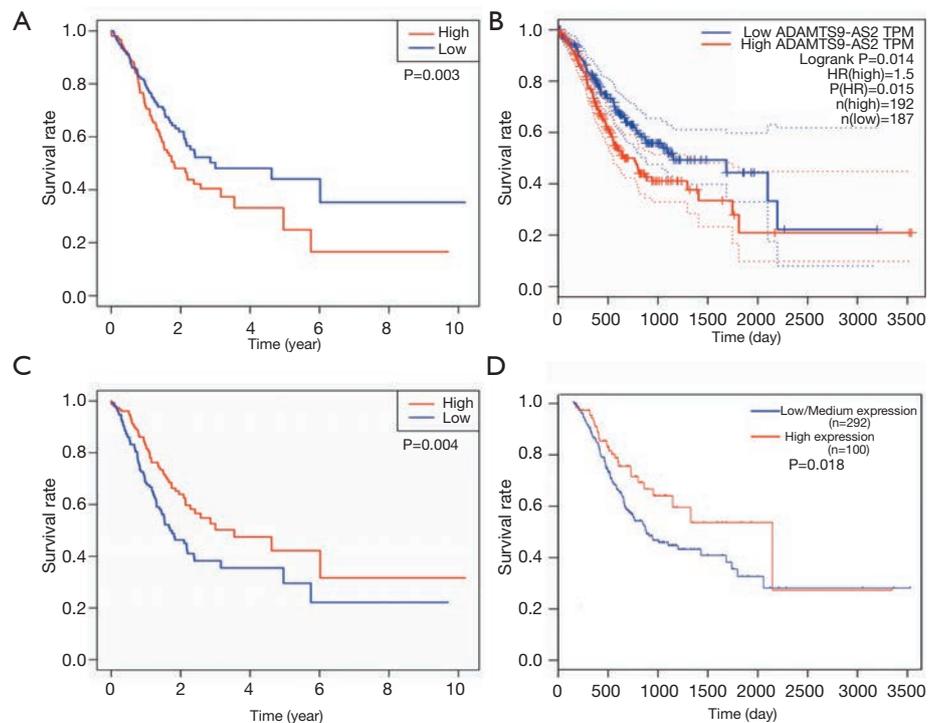
**Table 2** Hub lncRNA and survival analysis of gastric cancer lncRNA-miRNA-mRNA regulatory network

lncRNA	Direct regulated miRNA count	Indirectly regulated mRNA count	Survival analysis (P value)
SNHG14	28	1,319	0.068
KCNQ1OT1	26	1,198	0.367
MUC19	25	1,098	0.209
<i>ADAMTS9-AS2</i>	19	974	0.003
<i>PVT1</i>	19	769	0.004
PCBP1	19	934	0.605
LINC00461	19	973	0.175
AC010336.1	17	835	0.165
AGAP11	17	886	0.392
HOTTIP	16	740	0.249
TTY10	15	543	0.130
PCAT1	15	686	0.553

Italic values indicate *ADAMTS9-AS2* and *PVT1* are hub lncRNA and related to survival in gastric cancer.

*ADAMTS9-AS2* regulated mRNA revealed significant enrichment in the following BP: extracellular matrix organisation, epithelial tube morphogenesis, and mesenchymal cell proliferation (Figure 7A). The significantly enriched MF categories were mainly

transcription factor activity, proximal promoter DNA-binding transcription activator activity, and DNA-binding transcription activity (Figure 7B). The significantly enriched CC categories were extracellular matrix, collagen-containing extracellular matrix, and receptor complex



**Figure 5** TCGA dataset and online database for univariate cox survival analysis of ADAMTS9-AS2 and PVT1 in gastric cancer. (A) Patients with high expression of ADAMTS9-AS2 have poor prognosis; (B) Patients with low expression of PVT1 have poor prognosis. TCGA dataset survival analysis results for ADAMTS9-AS2 and PVT1 are consistent with online database survival analysis results (C and D). TCGA, The Cancer Genome Atlas.

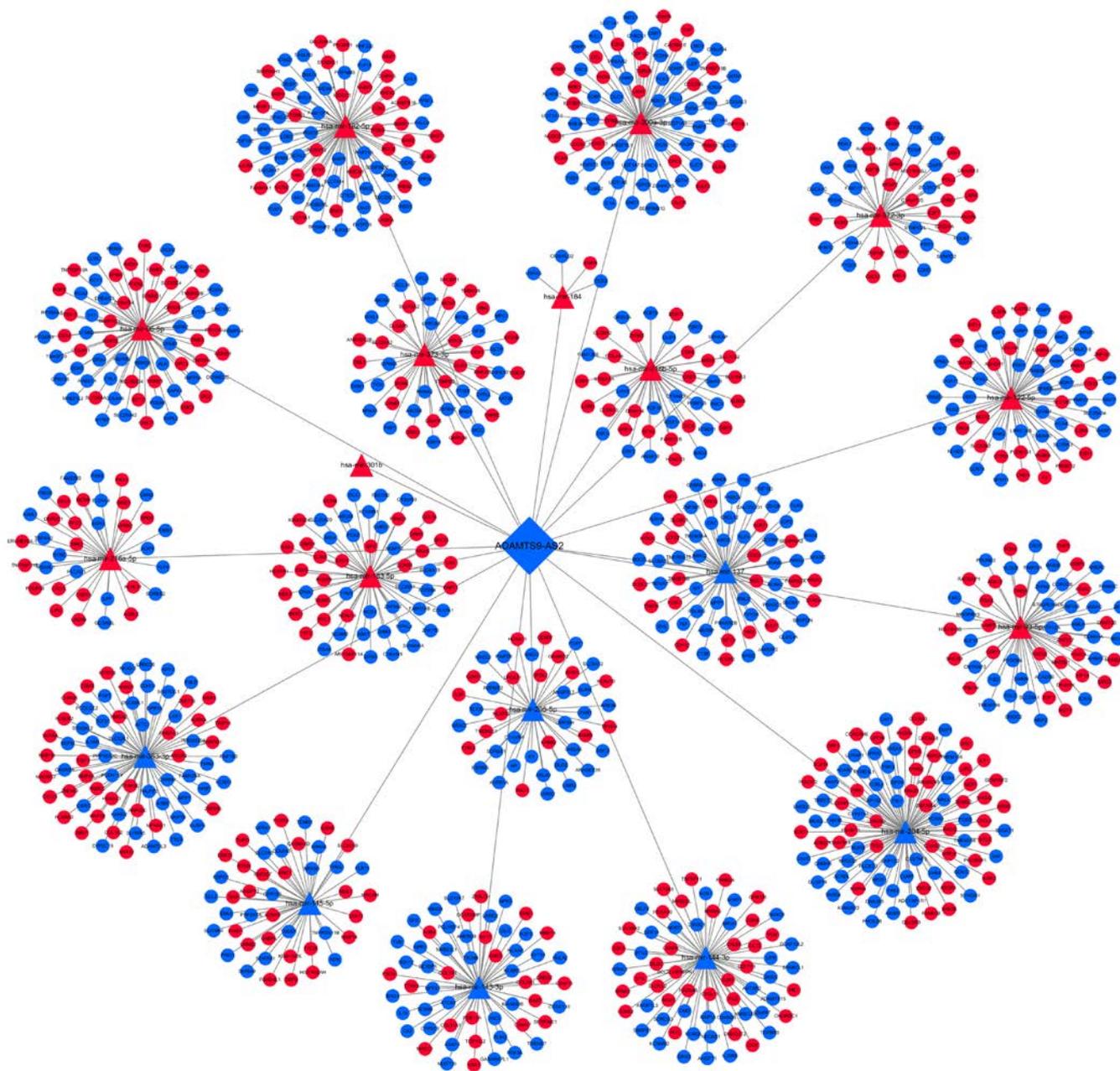
(Figure 7C). The KEGG pathways were mainly enriched in PI3K-Akt signalling pathway. (Figure 7D).

In the lncRNA *ADAMTS9-AS2* regulatory network, univariate Cox survival analysis showed that 179 mRNAs were associated with survival in GC ( $P < 0.05$ ). The survival curves of the top 20 genes significantly related to survival in GC are shown in Figure S2. *AKAP12*, *AR*, *ASF1B*, *CDH6*, *DYNC111*, *GPC3*, *GPR176*, *HIST1H2AH*, *MAP6*, *MCC*, *PCOLCE2*, *PLCL1*, *POLQ*, *RAB9B*, *SCUBE2*, *SERPINE1*, *SH3BGRL*, *TCEAL7*, *TMEM55A*, and *TMTC1* were positively correlated with survival in GC, whereas *ASF1B*, *HIST1H2AH*, and *POLQ* were negatively correlated with survival in GC. Thus, *AKAP12*, *AR*, *CDH6*, *DYNC111*, *GPC3*, *GPR176*, *MAP6*, *MCC*, *PCOLCE2*, *PLCL1*, *RAB9B*, *SCUBE2*, *SERPINE1*, *SH3BGRL*, *TCEAL7*, *TMEM55A*, *TMTC1*, *ASF1B*, *HIST1H2AH*, *POLQ*, and *ADAMTS9-AS2* can be used as prognostic indicators of GC landmarks. Simultaneously, based on the ceRNA hypothesis, *ADAMTS9-AS2* may act as a ceRNA regulating 179 mRNAs through miRNAs, causing abnormal activation of signal pathways such as PI3K-Akt and affecting the prognosis of GC.

### *PVT1*-related regulatory network

*PVT1* has a direct regulatory relationship with 19 miRNAs and indirectly regulates the expression of 769 mRNAs (Figure 8). R software was used to perform GO enrichment and KEGG signal pathway enrichment analysis on the 769 differentially expressed genes in the *PVT1* regulatory network. GO enrichment analysis revealed that the enriched BP were extracellular matrix tissue, DNA replication, extracellular matrix cell cycle G1/S phase transition, and Wnt signalling pathway (Figure 9A). The significantly enriched MF categories were transcriptional factor activity and DNA-binding transcription activator activity (Figure 9B). The significantly enriched CC categories were transmembrane transporter complex, receptor complex, and extracellular matrix (Figure 9C). KEGG pathway analysis of the lncRNA *PVT1* regulatory network mainly showed enrichment in signal pathways such as the PI3K-Akt signalling pathway, cAMP signalling, and cell cycle signalling pathway (Figure 9D).

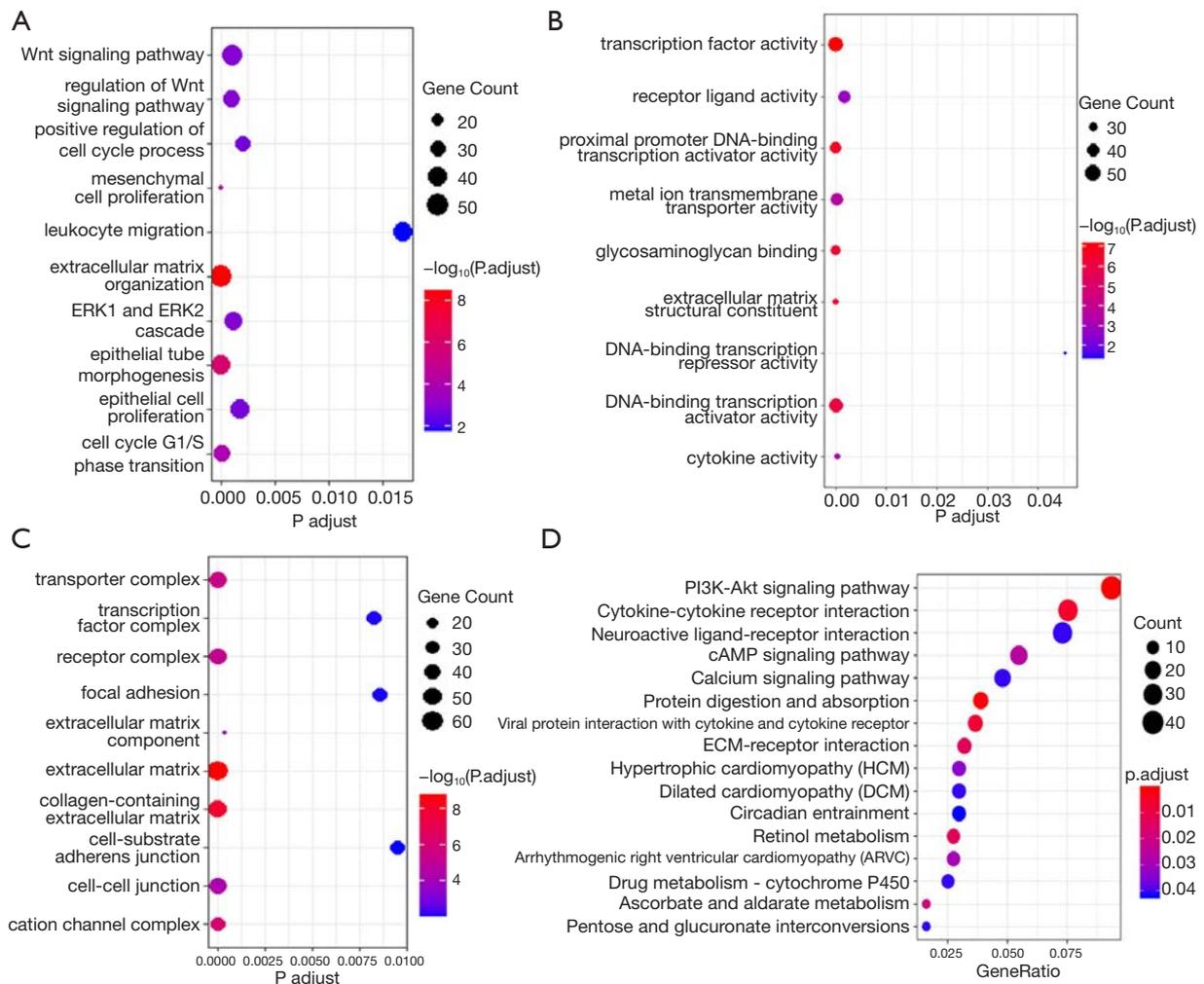
Univariate Cox survival analysis of genes in the *PVT1*



**Figure 6** Regulatory network of the lncRNA ADAMTS9-AS2. Circles represent mRNA, triangles represent miRNA, and diamonds represent lncRNA. Red represents upregulation, and blue represents downregulation.

regulatory network revealed that one miRNA (has-miR-216a) and 144 mRNAs were associated with survival in GC ( $P < 0.05$ ). The survival curves of the top 20 genes significantly related to survival in GC are shown in [Figure S3](#). *AKAP12*, *AR*, *ARMCX1*, *CNRIP1*, *COL4A5*, *FAT3*, *GHR*, *GLP2R*, *JAM3*, *MCC*, *MICB*, *PDGFD*,

*RAB9B*, *SERPINE1*, *SLC24A2*, *SSBP2*, and *TUBB6* were positively correlated with survival in GC, whereas *ASF1B*, *HIST1H2AH*, and *POLQ* were negatively correlated with survival in GC. The above results showed that *AKAP12*, *AR*, *ARMCX1*, *CNRIP1*, *COL4A5*, *FAT3*, *GHR*, *GLP2R*, *JAM3*, *MCC*, *MICB*, *PDGFD*, *RAB9B*, *SERPINE1*, *SLC24A2*,



**Figure 7** Analysis of functional enrichment of target genes in the lncRNA ADAMTS9-AS2 regulatory network using GO and KEGG pathway analysis. (A) Biological process, BP; (B) Molecular function, MF; (C) Cellular component, CC; (D) KEGG pathway analysis. (A), (B), (C), and (D) show the GO analysis results.  $P < 0.05$  was considered statistically significant. GO, Gene Ontology; KEGG, Kyoto Encyclopaedia of Genes and Genomes.

*SSBP2*, *TUBB6*, *ASF1B*, *HIST1H2AH*, *POLQ*, *hsa-mir-216a*, and *PVT1* can be used as prognostic biomarkers for GC.

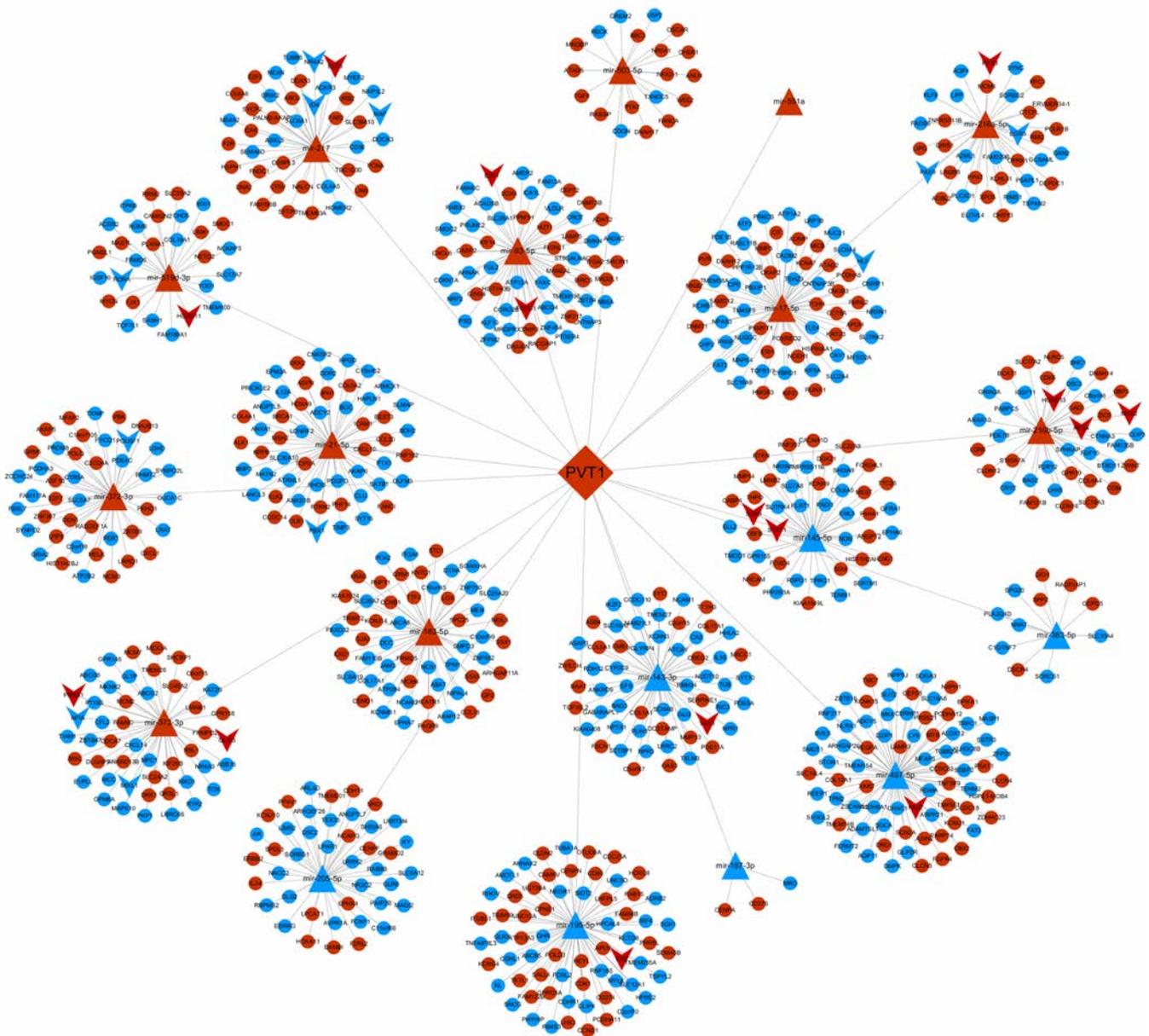
**Expression of the hub lncRNA, ADAMTS9-AS2 and PVT1, in GC samples**

Clinicopathological characteristics of 54 patients with GC are shown in Table S2. Through analysis of *ADAMTS9-AS2* and *PVT1* expression data from patients with GC in TCGA, we found that *ADAMTS9-AS2* expressions were downregulated whereas *PVT1* expressions were upregulated in GC samples (n=343) compared with those in adjacent

normal tissue samples (n=30) ( $P < 0.0001$ , Figure 10A and B). Upon validation by qRT-PCR analysis, as shown in Figures 10C and D, the expression of *ADAMTS9-AS2* was downregulated, whereas *PVT1* was upregulated in GC tissues (n=54) compared with those in the matched non-cancerous gastric tissues (n=54).

**Discussion**

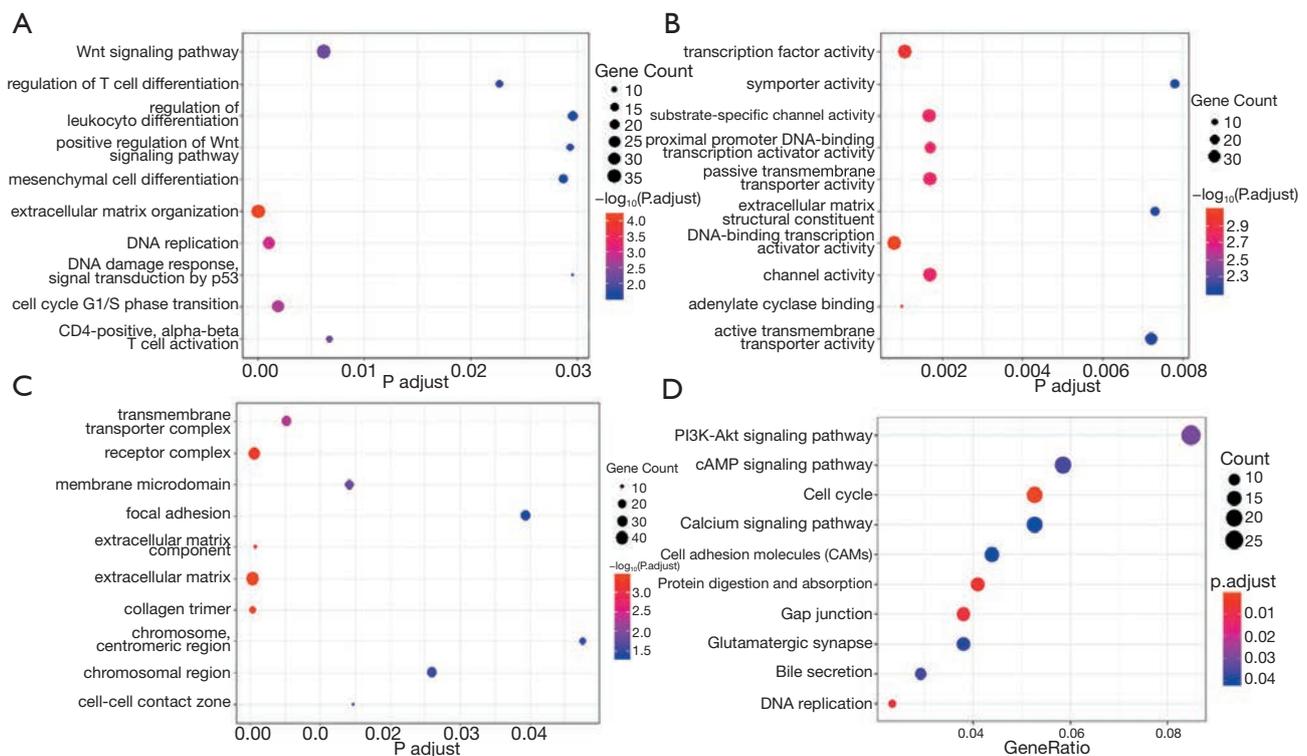
In this study, we used TCGA in combination with an online database to obtain transcriptome data for GC and differential miRNA genes (2,684 differentially expressed



**Figure 8** Regulatory network of the lncRNA PVT1. Circles represent mRNA, triangles represent miRNA, diamonds represent lncRNA, and “V” shapes represent transcription factors. Red represents upregulation, and blue represents downregulation.

lncRNAs, 240 differentially expressed miRNAs, and 4,285 differentially expressed mRNAs) and construct an lncRNA-miRNA-mRNA ceRNA regulatory network with a total of 1,626 network nodes and 2,704 interaction relationships. KEGG signal pathway enrichment analysis for the ceRNA network target genes showed that PI3K-Akt, stimulation of neural ligand and receptor interactions, and cytokine and cytokine receptor interactions were significantly

enriched. In addition, the lncRNAs *ADAMTS9-AS2* and *PVT1* were found to be key node genes in the ceRNA network. Functional enrichment and survival analyses were performed on the subnetworks regulated by the lncRNAs *ADAMTS9-AS2* and *PVT1*. Validation by qRT-PCR in 54 pairs of GC and normal tissues adjacent to the cancer showed that *ADAMTS9-AS2* was downregulated, whereas *PVT1* was upregulated in GC tissues, consistent with the



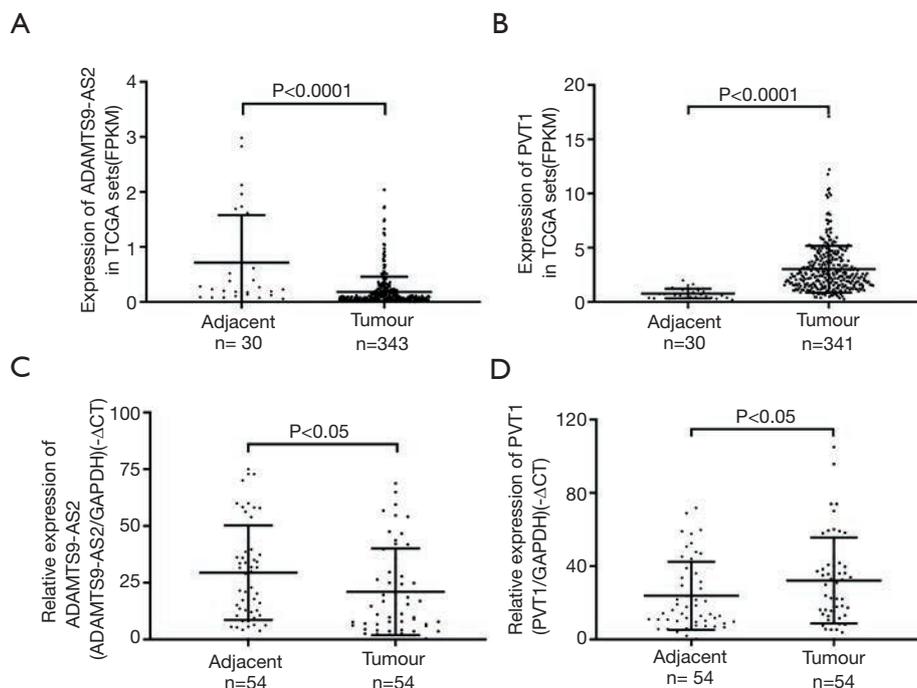
**Figure 9** Analysis of functional enrichment of target genes in the lncRNA PVT1 regulatory network using GO and KEGG pathway analysis. (A) Biological process (BP); (B) molecular function (MF); (C) cellular component (CC); (D) KEGG pathway analysis. A, B, C, and D show the GO analysis results.  $P < 0.05$  was considered statistically significant. GO, Gene Ontology; KEGG, Kyoto Encyclopaedia of Genes and Genomes.

differential analysis results from TCGA data.

In the lncRNA *ADAMTS9-AS2* regulatory network, *ADAMTS9-AS2* was found to regulate the expression of 974 mRNAs through 19 miRNAs. KEGG enrichment analysis showed that the target genes in the lncRNA *ADAMTS9-AS2* regulatory network are significantly enriched in the PI3K-Akt signalling pathway. Recent studies have shown that *ADAMTS9-AS2* is abnormally expressed in a variety of cancers, including lung cancer, breast cancer, ovarian cancer, and GC (20-23). The results of this study indicate that *ADAMTS9-AS2* is downregulated in tumours and cell lines. *In vivo* and *in vitro* experiments have confirmed that *ADAMTS9-AS2* functions as a tumour suppressor gene. However, in salivary cystadenomas, *ADAMTS9-AS2* acts as an oncogene (24). This suggests that the role of *ADAMTS9-AS2* in tumours is complex because different target genes are regulated in different tumours and may play different or even opposite roles. Downregulation of *ADAMTS9-AS2* expression in ovarian cancer and GC is associated with poor prognosis (22,23). In breast cancer, downregulation

of *ADAMTS9-AS2* expression can decrease *PTEN* expression and tamoxifen resistance through miRNA-130a-5p (21). The PI3K-Akt signalling pathway has been confirmed to be related with GC growth, differentiation, invasion, and metastasis in previous studies (25,26). Cell-culture experiments have confirmed that PI3K inhibitors can reverse the *ADAMTS9-AS2*-mediated inhibition of cell viability, colony forming ability, migration, invasion, and increased apoptosis in SGC-7901 cells, suggesting that PI3K-Akt signalling pathway plays an important role in GC tumorigenesis and development (23). The lncRNA-miRNA-mRNA regulatory network constructed by Wang Jie identified the lncRNA *ADAMTS9-AS2* as a hub gene. KEGG functional enrichment analysis showed that *ADAMTS9-AS2* regulatory genes were significantly enriched in the PI3K-Akt signalling pathway (17), consistent with the results of this study.

In the *PVT1* regulatory network, *PVT1* regulates the expression of 769 mRNAs through 19 miRNAs. The human *PVT1* gene is located on chromosome 8 q24.21,



**Figure 10** Using TCGA data, we observed that the expression of ADAMTS9-AS2 was downregulated (A), whereas that of PVT1 was upregulated (B) in gastric cancer tissues (n=343) compared with those in the matched non-cancerous gastric tissues (n=30). Our qRT-PCR (C and D) results are consistent with gene expression in gastric cancer tissues and adjacent non-cancerous gastric tissues according to TCGA. TCGA, The Cancer Genome Atlas.

with a total of 21 exons and 25 annotated transcripts (27). When studying the role of *PVT1* in cancer development, the heterogeneity of these different transcripts needs to be considered because the biological characteristics of different transcripts may have different effects on tumorigenesis, development, and patient survival. Recently, two studies on GC have confirmed that *PVT1* expression is upregulated in GC (28,29). Upregulation of *PVT1* is related to the interaction of transcription factors such as FOXM1 and STAT3, which promote *PVT1* expression in GC (28). Three studies have also confirmed that *PVT1* can promote the growth, colony formation, invasion, metastasis and tumour angiogenesis of GC cells (28-31). *PVT1* can inhibit the degradation of STAT3, which binds to the *PVT1* promoter region to promote its expression via a positive feedback mechanism. Upregulated *PVT1* eventually increases VEGFA expression to promote tumour angiogenesis. Upregulated *PVT1* and VEGFA are both associated with poor prognosis in GC (31). Furthermore, based on the ceRNA hypothesis, *PVT1* may act as a ceRNA to regulate 144 mRNAs through miRNAs, causing abnormal activation of signalling pathways such

as PI3K-Akt and affecting GC prognosis. However, this needs additional confirmation through *in vivo* and *in vitro* experimental studies.

However, there are some limitations to our study. First, the research on non-coding RNA is expanding, we could not cover all lncRNAs and miRNAs. Second, analysis of the complex relationship between lncRNAs and mRNAs was not included in this study. Finally, we just verified expressions of two key genes in 54 patients with GC but did not verify their clinical relevance, impact on survival, and molecular mechanism in GC. Our follow-up research will investigate the molecular mechanisms of the lncRNAs *ADAMTS9-AS2* and *PVT1*.

GC arises from a highly heterogeneous malignant tumour with a complex pathogenesis that involves abnormalities in gene mutations, DNA copy number, DNA methylation, and variable splicing. The study of single DNA, RNA, and transcription factor or protein regulatory relationships to analyse GC pathogenesis has certain limitations. In this study, an lncRNA-miRNA-mRNA regulatory network was constructed from a systematic perspective of differentially expressed genes in GC, and the network was analysed for

its application in the prediction of GC prognosis. The two key nodes in the network, lncRNA *ADAMTS9-AS2* and *PVT1*, along with their regulatory subnetworks, may act as molecular targets for GC treatment.

## Conclusions

The GC lncRNA-miRNA-mRNA regulatory network constructed in this study will help understand GC pathogenesis from a systematic perspective and simultaneously provide new information for early diagnosis, prognosis, and selection of potential therapeutic targets in GC.

## Acknowledgments

We thank the TCGA and GEO database builders and participants for providing open access to gene expression and clinical phenotype data.

*Funding:* This work was supported by the Gansu Provincial Health Industry Scientific Research Management (2019-19), the First Hospital of Lanzhou University Foundation (ldyyyn2019-53), and the National Natural Science Foundation of China (81372145).

## Footnote

*Reporting Checklist:* The authors have completed the MDAR checklist. Available at <http://dx.doi.org/10.21037/tcr-20-2168>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/tcr-20-2168>). The authors have no conflicts of interest to declare

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of Lanzhou University Second Hospital (2020A-031) and informed consent was taken from all the patients.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with

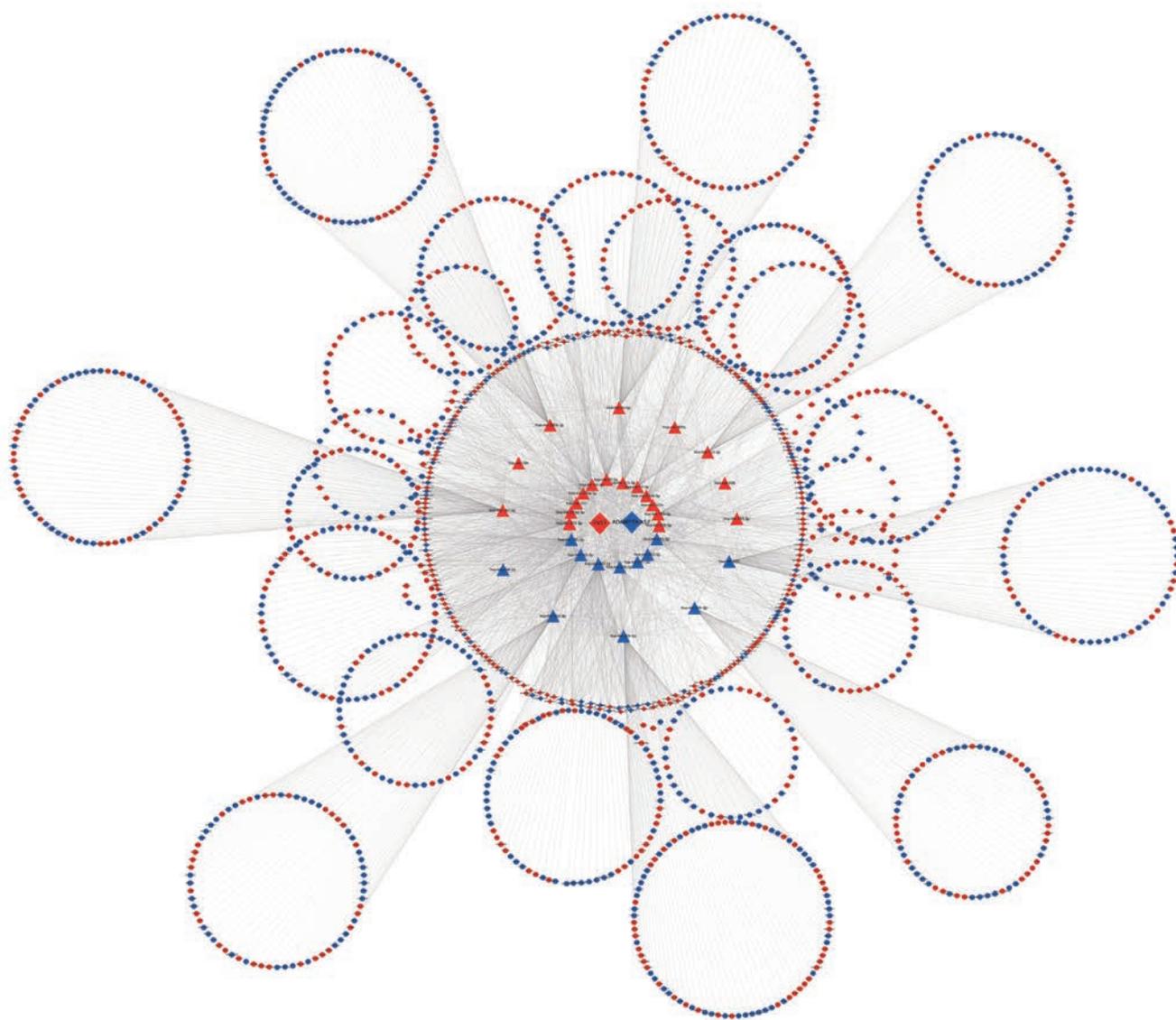
the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

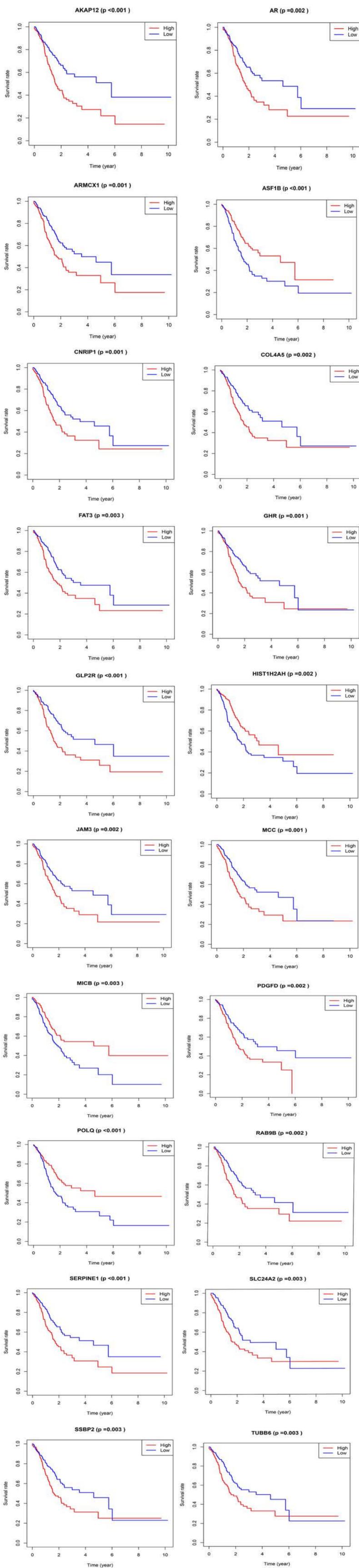
1. Siegel RL, Miller KD, Jemal A. Cancer statistics. *CA Cancer J Clin* 2017;67:7-30.
2. Chapelle N, Bouvier AM, Manfredi S. Early gastric cancer: trends in incidence, management, and survival in a well-defined French population. *Ann Surg Oncol* 2016;23:3677-83.
3. Esteller M. Non-coding RNAs in human disease. *Nat Rev Genet* 2011;12:861-74.
4. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer* 2015;15:321-33.
5. Calin GA, Ferracin M, Cimmino A. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 2005;353:1793-801.
6. Matsumura T, Sugimachi K, Iinuma H, et al. Exosomal microRNA in serum is a novel biomarker of recurrence in human colorectal cancer. *Br J Cancer* 2015;113:275-81.
7. Inamura K, Ishikawa Y. MicroRNA in lung cancer: novel biomarkers and potential yools for treatment. *J Clin Med* 2016;5:36.
8. Zhu XP, Mou KJ, Xu QF. Microarray analysis of the, aberrant microRNA expression pattern in gliomas of different grades. *Oncol Rep* 2015;34:318-24.
9. Kjersem JB, Ikdahl T, Lingjaerde OC. Plasma microRNAs predicting clinical outcome in metastatic colorectal cancer patients receiving first-line oxaliplatin-based treatment. *Mol Oncol* 2014;8:59-67.
10. Toiyama Y, Hur K, Tanaka K, et al. Serum mi R-200c is a novel prognostic and metastasis-predictive biomarker in patientswith colorectal cancer. *Ann Surg* 2014;259:735-43.
11. Sun Y, Wang M, Lin G, et al. Serum MicroRNA-155 as a Potential Biomarker to Track Disease in Breast Cancer. *PLoS One* 2012;7:e47003.
12. Wang H, Tan G, Dong L, et al. Circulating miR-125b as a marker predicting chemoresistance in breast cancer. *PLoS One* 2012;7:e34210.
13. Zhou X, Ye F, Yin C, et al. The interaction between miR-141and lncRNA-H19 in regulating cell proliferation and migration in gastric cancer. *Cell Physiol Biochem* 2015;36:1440-52.
14. Dan J, Wang J, Wang Y, et al. LncRNA-MEG3 inhibits

- proliferation and metastasis by regulating miRNA-21 in gastric cancer. *Biomed Pharmacother* 2018;99:931-8.
15. Wang X, Kan J, Han J, et al. LncRNA SNHG16 functions as an oncogene by sponging miR-135a and promotes JAK2/STAT3 signal pathway in gastric cancer. *J Cancer* 2019;10:1013-22.
  16. Yuan H, Chen Z, Bai S, et al. Molecular mechanisms of lncRNA SMARCC2/miR-551b-3p/TMPRSS4 axis in gastric cancer. *Cancer Letters* 2018;418:84-96.
  17. Wang J, Ding Y, Wu Y, et al. Identification of the complex regulatory relationships related to gastric cancer from lncRNA-miRNA-mRNA network. *J Cell Biochem* 2020;121:876-87.
  18. Wang P, Li J, Zhao W, et al. A Novel LncRNA-miRNA-mRNA Triple Network Identifies LncRNA RP11-363E7.4 as An Important Regulator of miRNA and Gene Expression in Gastric Cancer. *Cell Physiol Biochem* 2018;47:1025-41.
  19. Mao Y, Liu R, Zhou H, et al. Transcriptome analysis of miRNA-lncRNA-mRNA interactions in the malignant transformation process of gastric cancer initiation. *Cancer Gene Ther* 2017;24:267-75.
  20. Liu C, Yang Z, Deng Z, et al. Upregulated lncRNA ADAMTS9-AS2 suppresses progression of lung cancer through inhibition of miR-223-3p and promotion of TGFBR3. *IUBMB Life* 2018;70:536-46.
  21. Shi YF, Lu H, Wang HB. Downregulated lncRNA ADAMTS9-AS2 in Breast Cancer Enhances Tamoxifen Resistance by Activating microRNA-130a-5p. *Eur Rev Med Pharmacol Sci* 2019;23:1563-73.
  22. Wang A, Jin C, Li H, et al. LncRNA ADAMTS9-AS2 regulates ovarian cancer progression by targeting miR-182-5p/FOXF2 signaling pathway. *Int J Biol Macromol*, 2018;120:1705-13.
  23. Cao B, Liu C, Yang G. Down-regulation of lncRNA ADAMTS9-AS2 contributes to gastric cancer development via activation of PI3K/Akt pathway. *Biomed Pharmacother* 2018;107:185-93.
  24. Xie S, Yu X, Li Y, et al. Upregulation of lncRNA ADAMTS9-AS2 promotes salivary adenoid cystic carcinoma metastasis via PI3K/Akt and MEK/Erk signaling. *Mol Ther* 2018;26:2766-78.
  25. Liu Y, Chen S, Xue R, et al. Mefloquine effectively targets gastric cancer cells through phosphatase-dependent inhibition of PI3K/Akt/mTOR signaling pathway. *Biochem Biophys Res Commun* 2016;470:350-5.
  26. Zhang XB, Song L, Wen HJ, et al. Upregulation of microRNA-31 targeting integrin  $\alpha 5$  suppresses tumor cell invasion and metastasis by indirectly regulating PI3K/AKT pathway in human gastric cancer SGC7901 cells. *Tumour Biol* 2016;37:8317-25.
  27. Martínez-Barriocanal Á, Arango D, Dopeso H. PVT1 long non-coding RNA in gastrointestinal cancer. *Front Oncol* 2020;10:38.
  28. Xu MD, Wang Y, Weng W, et al. A positive feedback loop of lncRNA-PVT1 and FOXM1 facilitates gastric cancer growth and invasion. *Clin Cancer Res* 2017;23:2071-80.
  29. Huang T, Liu H, Chen J, et al. The long noncoding RNA PVT1 functions as a competing endogenous RNA by sponging miR-186 in gastric cancer. *Biomed Pharmacother* 2017;88:302-8.
  30. Zhao S, Fan N, Chen X, et al. Long noncoding RNA PVT1-214 enhances gastric cancer progression by upregulating TrkC expression in competitively sponging way. *Eur Rev Med Pharmacol Sci* 2019;23:4173-84.
  31. Zhao J, Du P, Cui P, et al. LncRNA PVT1 promotes angiogenesis via activating the STAT3/VEGFA axis in gastric cancer. *Oncogene* 2018;37:4094-109.

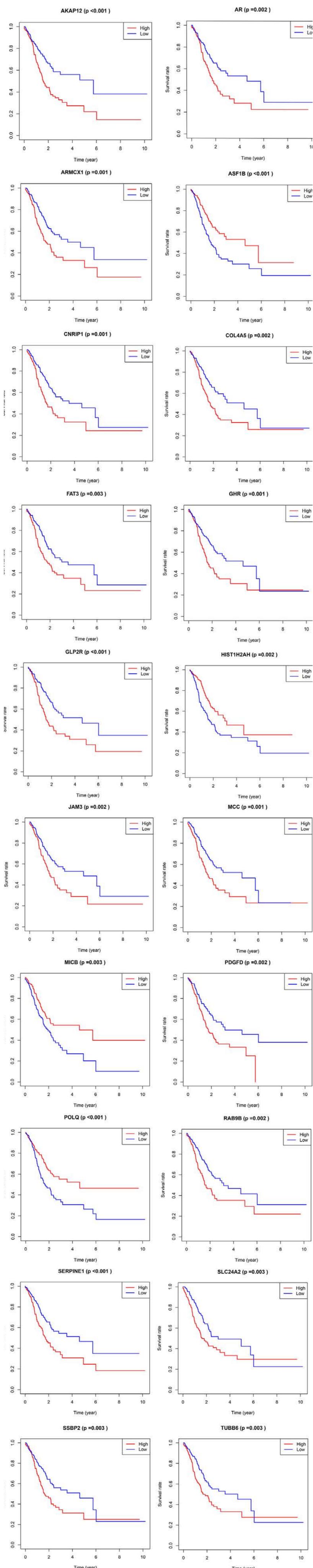
**Cite this article as:** Wang S, Li XC, Zhu JR, Ma ZJ, Ran JT, Zhou YN. Construction of a novel ceRNA network and identification of lncRNA *ADAMTS9-AS2* and *PVT1* as hub regulators of miRNA and coding gene expression in gastric cancer. *Transl Cancer Res* 2021;10(2):938-952. doi: 10.21037/tcr-20-2168



**Figure S1** lncRNA-miRNA-mRNA regulatory network of gastric cancer. LncRNAs are represented by diamonds, miRNAs are represented by triangles, and mRNA are represented by circles; the connections represent the regulatory relationship between the two RNAs; red represents the upregulation, and blue represents downregulation.



**Figure S2** Survival analysis of the lncRNA ADAMTS9-AS2 regulatory network target genes. Vertical axis: overall survival. Horizontal axis: total survival time.  $P < 0.05$  was considered statistically significant.



**Figure S3** Survival analysis of lncRNA PVT1 regulatory network target genes. Vertical axis: overall survival. Horizontal axis: total survival time. P<0.05 was considered statistically significant.

**Table S1** gene list of lncRNA-miRNA-mRNA regulatory network in TCGA gastric cancer dataset (lncRNA top100, mRNA top100, according to logFC)

Gene category	Gene symbol	logFC	P value	FDR
LncRNA	LINC00392	7.10	8.74E-07	8.03E-06
	ERVMER61-1	6.16	1.64E-07	1.87E-06
	HOTAIR	5.51	1.32E-19	1.62E-17
	AL139002.1	5.50	3.60E-07	3.71E-06
	LINC00355	5.16	2.18E-09	4.13E-08
	AL513123.1	4.87	1.26E-12	4.85E-11
	DSCR4-IT1	4.77	1.12E-05	7.48E-05
	HULC	4.69	1.47E-07	1.70E-06
	C10orf126	4.56	7.92E-07	7.37E-06
	AP002478.1	4.42	4.03E-15	2.47E-13
	POU6F2-AS2	4.40	1.46E-06	1.25E-05
	LINC00454	4.32	2.36E-05	1.41E-04
	AC061975.6	4.18	7.07E-05	3.60E-04
	LINC00326	3.88	0.000139	6.48E-04
	C7orf69	3.84	6.19E-13	2.56E-11
	LINC00200	3.73	0.000354	1.45E-03
	LINC00393	3.72	4.83E-05	2.59E-04
	NOVA1-AS1	3.70	8.78E-06	6.05E-05
	NKX2-1-AS1	3.47	0.002207	6.75E-03
	AC026320.1	3.46	0.000811	2.93E-03
	LINC00460	3.26	2.84E-09	5.25E-08
	LINC00052	3.24	4.28E-05	2.35E-04
	AP000525.1	3.16	7.25E-12	2.38E-10
	BOK-AS1	3.13	8.79E-05	0.000435
	DLX6-AS1	3.10	1.29E-07	1.52E-06
	UCA1	3.06	5.39E-07	5.25E-06
	LINC00534	3.04	1.19E-05	7.89E-05
	HOTTIP	3.01	6.66E-08	8.57E-07
	AC087269.1	3.01	8.26E-10	1.72E-08
	LINC00221	3.00	0.001243	0.004188
	C8orf31	2.97	4.92E-12	1.68E-10
	FREM2-AS1	2.96	0.000553	0.00211
	MYB-AS1	2.95	1.47E-09	2.91E-08
	LINC00524	2.80	4.51E-05	0.000246
	IGF2-AS	2.79	2.45E-05	0.000146
	AP002992.1	2.77	2.43E-05	0.000145
	NAALADL2-AS2	2.75	1.19E-05	7.89E-05
	C15orf54	2.73	6.81E-09	1.13E-07

TCGA, The Cancer Genome Atlas.

**Table S2** List of genes of gastric cancer lncRNA-miRNA-mRNA regulatory network

Gene category	Gene symbol	logFC	P value	FDR
	AL356133.2	2.70	0.000226	0.00098
	FNDC1-IT1	2.68	4.53E-07	4.54E-06
	H19	2.67	2.88E-06	2.26E-05
	PLCH1-AS2	2.63	4.17E-07	4.22E-06
	LINC00114	2.50	0.000236	0.001017
	LINC00469	2.49	0.001099	0.003779
	LINC00523	2.47	0.002585	0.007707
	DLEU7-AS1	2.46	6.81E-17	5.97E-15
	AC009121.1	2.45	2.59E-12	9.35E-11
	AL357060.1	2.44	2.36E-06	1.90E-05
	AC090398.1	2.43	0.000161	0.000731
	AC074035.1	2.42	0.000302	0.00126
	CACNA1C-IT2	-1.45	0.0008	0.002896
	MIR22HG	-1.47	8.97E-24	2.04E-21
	LINC00278	-1.49	0.000931	0.003281
	HCG23	-1.50	1.87E-07	2.10E-06
	AC107959.1	-1.53	6.36E-22	1.15E-19
	ATP1B3-AS1	-1.54	1.57E-15	1.07E-13
	ALDH1L1-AS2	-1.54	3.48E-05	0.000197
	TTY14	-1.54	0.000322	0.00133
	BX255923.1	-1.57	6.11E-07	5.85E-06
	AC025594.2	-1.58	6.42E-06	4.60E-05
	HCG21	-1.58	2.78E-05	0.000163
	AGAP11	-1.59	7.47E-09	1.22E-07
	AL139240.1	-1.59	0.00014	0.00065
	AC078993.1	-1.63	9.75E-06	6.63E-05
	LINC00092	-1.67	1.10E-14	6.32E-13
	SNHG14	-1.69	8.62E-14	4.21E-12
	GFOD1-AS1	-1.70	2.36E-07	2.57E-06
	LINC00461	-1.70	3.88E-09	6.87E-08
	IQCJ-SCHIP1-AS	-1.71	1.10E-07	1.32E-06
	AC007834.1	-1.71	8.89E-06	6.12E-05
	LINC00163	-1.81	5.17E-07	5.09E-06
	AL158206.1	-1.82	2.03E-15	1.33E-13
	FAM181A-AS1	-1.84	2.68E-05	0.000158
	TDRG1	-1.85	0.000304	0.001266
	MME-AS1	-1.88	1.37E-05	8.91E-05
	PDZRN3-AS1	-2.00	1.68E-09	3.28E-08
	AC002511.1	-2.02	5.24E-13	2.21E-11
	OSTN-AS1	-2.14	5.47E-13	2.28E-11
	IL20RB-AS1	-2.18	7.34E-12	2.40E-10
	AL357153.1	-2.23	2.21E-11	6.48E-10
	ARHGGEF26-AS1	-2.25	1.68E-16	1.33E-14
	CHL1-AS2	-2.26	6.13E-12	2.05E-10
	PART1	-2.35	4.37E-15	2.67E-13
	RBMS3-AS3	-2.35	1.23E-20	1.90E-18
	ADAMTS9-AS2	-2.37	1.08E-25	3.29E-23
	FRMD6-AS2	-2.38	1.59E-13	7.35E-12
	LINC00365	-2.45	6.33E-18	6.32E-16
	SOX21-AS1	-2.51	6.46E-11	1.72E-09
	ADAMTS9-AS1	-2.69	3.19E-37	4.87E-34
	AC092422.1	-2.71	1.20E-22	2.46E-20
	LINC00332	-2.71	4.46E-10	9.92E-09
	C20orf166-AS1	-2.71	1.27E-16	1.04E-14
	MIR205HG	-2.98	2.84E-06	2.24E-05
	LRRC3-AS1	-3.08	8.80E-30	4.95E-27
	LINC00330	-3.12	5.43E-18	5.47E-16
	AC011374.1	-3.52	2.21E-23	4.93E-21
	AC110491.1	-3.56	2.00E-21	3.33E-19
	AC024597.1	-3.64	1.13E-18	1.27E-16
	HCG22	-4.44	5.56E-30	3.50E-27
miRNA	hsa-mir-372	7.52	1.03E-08	4.61E-08
	hsa-mir-373	7.16	3.38E-07	1.14E-06
	hsa-mir-519d	4.14	0.000109	0.000229
	hsa-mir-184	3.92	4.98E-08	2.02E-07
	hsa-mir-122	3.44	0.000327	0.000643
	hsa-mir-935	2.86	1.46E-13	1.49E-12
	hsa-mir-508	2.70	2.31E-09	1.15E-08
	hsa-mir-183	1.94	1.31E-16	1.78E-15
	hsa-mir-551a	1.93	1.04E-06	3.28E-06
	hsa-mir-21	1.91	1.54E-48	4.70E-46
	hsa-mir-217	1.83	5.40E-08	2.17E-07
	hsa-mir-301b	1.80	2.50E-10	1.44E-09
	hsa-mir-96	1.76	7.99E-15	9.40E-14
	hsa-mir-216b	1.76	0.001982	0.003342
	hsa-mir-2115	1.62	8.38E-05	0.00018
	hsa-mir-182	1.56	2.78E-13	2.66E-12
	hsa-mir-200a	1.27	1.54E-09	7.81E-09
	hsa-mir-210	1.24	3.73E-05	8.71E-05
	hsa-mir-503	1.22	2.35E-09	1.16E-08
	hsa-mir-17	1.18	5.21E-14	5.69E-13
	hsa-mir-93	1.16	4.55E-13	4.16E-12
	hsa-mir-497	-1.00	2.06E-12	1.71E-11
	hsa-mir-363	-1.07	2.35E-07	8.18E-07
	hsa-mir-100	-1.20	4.87E-10	2.71E-09
	hsa-mir-144	-1.26	2.73E-10	1.56E-09
	hsa-mir-187	-1.33	5.30E-05	0.000117
	hsa-mir-195	-1.53	1.39E-24	4.25E-23
	hsa-mir-143	-1.86	1.62E-23	4.73E-22
	hsa-mir-204	-1.89	3.51E-12	2.75E-11
	hsa-mir-137	-2.01	4.61E-09	2.17E-08
	hsa-mir-145	-2.42	2.67E-42	5.46E-40
	hsa-mir-383	-2.52	2.22E-18	3.58E-17
	hsa-mir-205	-2.60	1.24E-06	3.81E-06
mRNA	MAGEC2	7.68	6.25E-06	2.44E-05
	C6orf15	7.40	1.63E-08	9.81E-08
	IBSP	6.94	2.09E-21	8.41E-20
	BPIFA1	6.65	5.40E-06	2.13E-05
	CST4	6.44	1.30E-16	2.69E-15
	BAAT	6.36	5.52E-16	1.05E-14
	HTR2C	6.05	2.73E-09	1.88E-08
	SSX1	5.94	1.08E-05	4.03E-05
	COL11A1	5.92	2.02E-21	8.15E-20
	DSCR4	5.77	1.15E-06	5.07E-06
	LIN28B	5.72	4.43E-06	1.78E-05
	FAM9C	5.53	5.20E-08	2.90E-07
	TKTL1	5.38	4.19E-09	2.78E-08
	HOXC13	5.33	1.54E-10	1.29E-09
	LEFTY1	5.31	8.79E-10	6.53E-09
	S100A7A	5.24	1.65E-05	5.91E-05
	SALL4	5.18	1.29E-21	5.45E-20
	CALB1	5.09	1.83E-08	1.09E-07
	CA6	4.97	3.76E-07	1.81E-06
	KRT23	4.87	2.62E-10	2.12E-09
	LY6G6D	4.87	0.000171	0.000496
	AQP2	4.87	1.87E-08	1.12E-07
	TRIM71	4.86	7.37E-09	4.72E-08
	ASB4	4.80	2.08E-07	1.04E-06
	CTCF	4.79	1.29E-07	6.74E-07
	HOXD11	4.73	4.54E-06	1.82E-05
	ZNF716	4.65	2.71E-08	1.58E-07
	C14orf105	4.63	1.91E-11	1.85E-10
	HOXD13	4.61	5.86E-05	0.000187
	MMP3	4.56	8.53E-13	1.01E-11
	GAD1	4.45	2.53E-17	5.78E-16
	EPYC	4.31	6.83E-09	4.40E-08
	HOXC8	4.29	2.94E-21	1.16E-19
	MMP13	4.26	8.31E-13	9.83E-12
	HMG2A	4.21	3.09E-16	6.12E-15
	GNGT1	4.20	4.05E-08	2.30E-07
	NEUROD4	4.20	0.000314	0.000862
	HOXA9	4.06	5.69E-12	5.99E-11
	NKX2-1	4.03	0.000319	0.000877
	EN2	4.02	4.58E-08	2.58E-07
	LHX1	3.98	3.65E-07	1.76E-06
	STATH	3.90	0.001169	0.002818
	CDH12	3.88	2.16E-05	7.56E-05
	ZIC5	3.85	9.91E-09	6.20E-08
	DCSTAMP	3.79	4.12E-18	1.03E-16
	F2	3.79	1.08E-07	5.70E-07
	UGT2B4	3.66	0.00012	0.000361
	FGB	3.64	0.000132	0.000391
	HOXA10	3.64	3.27E-15	5.47E-14
	NNAT	-3.04	6.96E-52	3.62E-49
	TMEM100	-3.05	5.83E-30	6.54E-28
	PAX9	-3.05	1.04E-21	4.49E-20
	EMP1	-3.06	5.70E-52	3.05E-49
	PSD	-3.07	7.05E-32	9.99E-30
	PMP2	-3.11	9.94E-23	5.08E-21
	HPCAL4	-3.12	2.89E-36	6.22E-34
	TMEM236	-3.12	7.76E-27	6.35E-25
	DSC3	-3.12	1.13E-14	1.75E-13
	PKP1	-3.14	1.61E-12	1.84E-11
	PRIMA1	-3.16	1.24E-24	7.94E-23
	ADCYAP1R1	-3.21	5.64E-29	5.74E-27
	KCNA1	-3.21	4.23E-14	6.08E-13
	ABCG8	-3.23	1.03E-21	4.46E-20
	PHOX2B	-3.23	6.09E-16	1.15E-14
	PCK1	-3.23	1.59E-21	6.61E-20
	SLC6A19	-3.23	5.20E-18	1.28E-16
	USH1G	-3.26	8.58E-17	1.83E-15
	PADI1	-3.32	6.97E-15	1.11E-13
	SLC5A7	-3.33	3.17E-21	1.15E-19
	ESRRG	-3.34	1.02E-29	1.10E-27
	UGT2B17	-3.38	7.42E-13	8.84E-12
	SERTM1	-3.40	1.30E-16	2.70E-15
	PSAPL1	-3.49	2.19E-20	7.73E-19
	ZNF365	-3.51	9.02E-42	2.81E-39
	UGT1A1	-3.59	1.15E-24	7.38E-23
	ALOX12	-3.74	1.94E-73	2.79E-70
	ASB11	-3.77	1.90E-31	2.51E-29
	CIDEA	-3.81	7.06E-41	2.07E-38
	SLC6A4	-3.82	2.50E-51	1.27E-48
	SERPINB13	-3.85	3.79E-10	2.97E-09
	RNF222	-3.97	9.76E-31	1.19E-28
	UGT1A7	-4.02	1.19E-29	1.28E-27
	CRISP3	-4.12	2.97E-16	5.90E-15
	C11orf86	-4.19	3.21E-36	6.76E-34
	UGT1A4	-4.23	2.23E-42	7.08E-40
	UGT1A5	-4.24	4.47E-35	8.64E-33
	A2ML1	-4.39	7.66E-21	2.85E-19
	CA7	-4.57	1.75E-75	2.97E-72
	SYNPO2L	-4.59	4.23E-88	1.32E-84
	ENDOU	-4.59	3.07E-38	7.32E-36
	MUC21	-4.72	3.64E-16	7.12E-15
	AQP4	-4.72	1.27E-35	2.60E-33
	MYOC	-4.75	2.93E-38	7.13E-36
	TMPRSS15	-4.78	2.19E-25	1.48E-23
	GKN2	-4.83	1.45E-24	9.20E-23
	TMPRSS11B	-4.85	5.41E-14	7.64E-13
	SLC28A1	-5.00	3.69E-77	6.90E-74
	G6PC	-5.58	1.34E-58	9.31E-56

Note: LogFC: log2foldChange value. FDR, false discovery rate.