Research Article

Construction of miRNA-mRNA-TF Regulatory Network for Diagnosis of Gastric Cancer

Zhenjie Fu,[1] Yuqin Xu,[1] Yan Chen,[1] Hang Lv,[2] Guiping Chen [ID],[2] and Yitao Chen [ID][1]

[1]*College of Life Science, Zhejiang Chinese Medical University, Hangzhou 310053, China*
[2]*Department of Gastrointestinal Surgery, The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China*

Correspondence should be addressed to Guiping Chen; 1941206479@qq.com and Yitao Chen; cytworld@163.com

Gastric cancer (GC), as an epidemic cancer worldwide, has more than 1 million new cases and an estimated 769,000 deaths worldwide in 2020, ranking fifth and fourth in global morbidity and mortality. In mammals, both miRNAs and transcription factors (TFs) play a partial role in gene expression regulation. The mRNA expression profile and miRNA expression profile of GEO database were screened by GEO2R for differentially expressed genes (DEGs) and differentially expressed miRNAs (DEMs). Then, DAVID annotated the functions of DEGs to understand the functions played in biological processes. The prediction of potential target genes of miRNA and key TFs of mRNA was performed by mipathDB V2.0 and CHEA3, respectively, and the gene list comparison was performed to look for overlapping genes coregulated by key TFs and DEMs. Finally, the obtained miRNAs, TF, and overlapping genes were used to construct the miRNA-mRNA-TF regulatory network, which was verified by RT-qPCR. 76 upregulated DEGs, 199 downregulated DEGs, and 3 upregulated miRNAs (miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p) were identified from the expression profiles of mRNA (GSE26899, GSE29998, GSE51575, and GSE13911) and miRNA (GSE93415), respectively. Through database prediction and gene list comparison, it was found that among the 199 downregulated DEGs, 61, 71, and 69 genes were the potential targets of miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p, respectively. 199 downregulated DEGs were used as the gene list for the prediction of key TFs, and the results showed that RFX6 ranked the highest. The potential target overlap genes of miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p were 4 genes (SH3GL2, ATP4B, CTSE, and SORBS2), 7 genes (SLC7A8, RNASE4, ESRRG, PGC, MUC6, Fam3B, and FMO5), and 6 genes (CHGA, PDK4, TMPRSS2, CLIC6, GPX3, and PSCA), respectively. Finally, we constructed a miRNA-mRNA-TF regulatory network based on the above 17 mRNAs, 3 miRNAs, and 1 TF and verified by RT-qPCR and western blot results that the expression of RFX6 was downregulated in GC tissues. These identified miRNAs, mRNAs, and TF have a certain reference value for further exploration of the regulatory mechanism of GC.

1. Introduction

Gastric cancer (GC), as an epidemic cancer worldwide, has more than 1 million new cases and an estimated 769,000 deaths worldwide in 2020, ranking fifth and fourth in global morbidity and mortality [1]. Despite great advances in our understanding of cancer, the causes of its occurrence and progression are not fully understood. In general, the related factors involved in the occurrence and development of GC can be divided into genetic and epigenetic changes, as well as environmental and lifestyle factors [2, 3]. The molecular pathogenesis of GC has always been the focus of scientists. In recent years, miRNA and transcription factors (TFs) have become an important breakthrough direction in the diagnosis, prognosis, and targeted therapy of GC.

As a class of small noncoding RNAs (ncRNAs) composed of 17-25 nucleotides [4], miRNAs regulate mRNA and protein levels by guiding Argonaute (AGO) proteins to the $3'$-UTR region of mRNA leading to mRNA degradation [4–7]. Since the incidental discovery of miRNAs in *Caenorhabditis elegans* [8–12], accumulated evidence has shown that miRNAs play a key regulatory role in biological processes in animals,

especially in recent years, new achievements have been made in the study of miRNAs and human cancers [13, 14]. Previous studies have shown that in the diagnosis of GC, miR-940 showed 81.25% sensitivity and 98.57% specificity. The AUC, sensitivity, and specificity of miR-21, miR-93, miR-106a, and miR-106b combinations were 0.887, 84.80%, and 79.2%, respectively [15, 16]. In addition, the expression of miR-204 (HR = 3.900, 95% CI 1.300-11.800) and miR-15a (RR = 1.950, 95% CI 0.470 9.130) was strongly associated with poor survival in patients with GC [17, 18]. Therefore, it is very promising to complete the personalized management, diagnosis, and prognosis of GC through miRNA combination in the future [19, 20].

TFs are proteins that bind to chromatin to activate or inhibit transcription by helical binding of DNA to specific regulatory sequences in the form of trans-activated or trans-repressive domains. They are expressed in tissues spatially, temporally, and sequentially during development, cell renewal, or differentiation, and any change in their expression may lead to uncontrolled cell integrity or dynamic homeostasis, leading to pathological changes [21], such as diabetes [22] and cancers [23]. TFs are classified into different families, reflecting the homology of their DNA binding domains, thus, reflecting the binding sequence of DNA [24, 25]. However, in-depth studies have found that TFs can play a double-edged sword role, so they may play both anticancer and oncogenic roles, even though TFs of the same family may have two-way effects on the same cancer [26, 27]. Taking the role of SOX family TFs in the pathological process of GC as an example, since the first discovery of SRY protein as a TF involved in mammalian male determination, up to 20 members of SOX family have been identified in mammals [28, 29]. However, although they share a DNA binding domain consisting of 79 amino acids, their effects on GC cells are quite different [30]. For example, SOX1 and SOX6 have anticancer effects; SOX3, SOX4, and SOX5 have procancer effects; and SOX2, SOX7, and SOX10 have both anticancer and procancer effects [31]. Therefore, to accurately understand the role of a TF in a certain cancer is the focus of cancer mechanism research and treatment in recent years.

The aim of this study is to further understand the direct regulatory relationship between mRNA, miRNA, and TFs. Through the analysis of existing mRNA and miRNA expression profile data in the GEO database, the functional annotation, and TF prediction of DEGs, target gene prediction of DEMs was performed with the help of databases such as DAVID, mipathDB V2.0, and CHEA3, respectively, and the regulation network of miRNA-mRNAs and TF was finally constructed based on the results above.

2. Material and Methods

2.1. Microarray Data. To identify DEGs and DEMs in the pathogenesis of GC and build a regulatory network of TF-miRNA-mRNA, we searched the Gene Expression Omnibus (GEO) database with keywords such as "gastric cancer," "mRNA," and "miRNA." Only the data sets containing both GC tissue samples and control group information

and without any drug intervention before sequencing could meet the criteria. Finally, 4 mRNA data sets (GSE26899, GSE29998, GSE51575, and GSE13911) and 1 miRNA data set (GSE93415) were selected for the follow-up analysis of this work.

2.2. Identification of DEGs and DEMs. GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/) is a software based on GEO database to perform differential analysis on expression profile chips. After the processing of 5 data sets using GEO2R, DEGs and DEMs were screened out using adjust p value < 0.05, |logFC| >1 and adjust p value < 0.01, |logFC| >1.9 as the criteria, respectively. The p value adjusted by Benjamini Hochberg false discovery rate (FDR) method can significantly reduce the false positive rate [32].

2.3. Biological Significance Analysis of DEGs. In order to further understand the biological function of DEGs, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed via a database for annotation, visualization, and integrated discovery (DAVID, https://david.ncifcrf.gov/home.jsp). The cut-off criterion, p value < 0.05, was set to selected GO terms and KEGG pathway enrichment analysis, and then the analysis results were visualized through OriginPro (2021b_Beta7) software and bioinformatics, an online data visualization tool (http://www.bioinformatics.com.cn).

2.4. Prediction of miRNA Target Genes. miRpathDB v2.0 (https://mpd.bioinf.uni-sb.de/overview.html) is a multifunctional online miRNA research site with a large database of miRNA target genes. In order to comprehensively and accurately excavate the regulatory relationship between miRNA and mRNA in GC tissues, the potential targets of candidate miRNA were comprehensively predicted by mipathDB v2.0, and the possible miRNA-mRNA regulation was found by comparing potential targets with DEGs.

2.5. Prediction of mRNA Transcription Factors. In the process of gene transcription and expression, miRNA and TF, respectively, play a partial regulatory role. Therefore, further digging out key TFs after clarifying the regulatory relationship between miRNA and mRNA is more conducive to in-depth understanding of the pathological mechanism of GC. TF prediction was performed via ChIP-X Enrichment Analysis 3 (ChEA3). ChEA3, whose database contains a collection of gene set libraries generated from multiple sources including TF-gene coexpression from RNAseq studies, TF-target associations from ChIP-seq experiments, and TF-gene cooccurrence computed from crowd-submitted gene lists, is a TFs enrichment analysis tool that ranks TFs associated with user-submitted gene sets.

3. Patients and Samples

To further verify the reliability of the above bioinformatics analysis, 4 pairs of clinical tissue samples from GC patients were obtained from the First Affiliated Hospital of Zhejiang Chinese Medical University. After surgical removal from the patient, all tissue samples were stored in a -80°C cryogenic

refrigerator and kept cryogenic during the samples transfer and nonexperimental disposal period. All biometric tests of tissue samples were performed with the informed consent of the patients. The Ethics Committee of the First Affiliated Hospital of Zhejiang Chinese Medical University approved the research (approval no.: 2017-K-002-01).

3.1. RNA Isolation and Quantitative Reverse Transcription-PCR (RT-qPCR).

Total RNA was extracted from tissue samples by Trizol method. First of all, 0.05 g tissue samples were weighed and put into 1.0 ml Trizol (Carlsbad, California, USA) solution, fully ground, and then stood for 10 min. Then, 200 μl trichloromethane (Xilong Chemical Co., Ltd., Guangdong Province, China) was added, shaken, and stood for 2 min, and the supernatant was extracted by centrifugation at 4°C 12000 rpm for 15 min. Next, add the same amount of isopropyl alcohol (Huadong Medicine Co., Ltd., Zhejiang Province, China) and rest overnight in the -80°C cryogenic refrigerator. The next day, the samples were taken out, and the supernatant was centrifuged at 4°C 12000 rpm for 10 min and washed twice with an appropriate amount of 75% ethanol (Hangzhou Longshan Fine Chemical Co., Ltd., Zhejiang Province, China). Finally, 20-30 μl nuclease-free water (Promega Corporation, Wisconsin, USA) was added for mixing.

Then, Revert Aid First Strand cDNA Synthesis Kit (Waltham, Massachusetts, USA) was used to obtain cDNA by the instructions, and the iTaqTM Universal SYBR@ Green Master Mix kit was used to detect the expression level of mRNA. All primers (RFX6 forward: CATGGCAAGCC GAGGAAGTGTC; RFX6 reverse: GGTATGTGGAGCAG TGTGATGGAG) were synthesized by Sangon Bioengineering (Shanghai) Co., Ltd. GAPDH (GAPDH forward: CGAGCC ACATCGCTCAGACA, GAPDH; GAPDH reverse: CTGG TGAAGACGCCAGTGGA) was used as endogenous control, and all results were calculated and expressed as $2^{-\Delta\Delta Ct}$.

3.2. Western Blot.

Clinical tissue samples were lysed in RIPA lysis buffer, and protein concentrations were determined using a BCA Protein Assay Kit (Thermo Scientific, MA, USA). Tissue lysates were separated with SDS-PAGE gels and transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore, Eschborn, Germany). The blots were blocked in 5% milk for 1 h at room temperature. PVDF membranes were incubated with the primary antibodies overnight in a cold room at 4°C. Subsequently, bound primary antibodies were reacted with corresponding secondary antibodies for 1 h at room temperature and evaluated with by chemiluminescence.

3.3. Statistical Analysis.

Statistical analysis was performed through GraphPad Prism (version 6, San Diego, CA) software. Student's t-tests were utilized for the comparison of two sample groups. Differences were considered as statistically significant when $p < 0.05$.

4. Results

4.1. Differentially Expressed Genes.

Four mRNA expression profiles from GEO database included 210 gastric cancer tissues and 118 normal tissue samples. Four expression profiles (GEO accession nos. GSE26899, GSE29998, GSE51575, and GSE13911) identified 172 upregulated genes and 338 downregulated genes, 1068 upregulated genes and 767 downregulated genes, 1319 upregulated genes and 1868 downregulated genes, and 1001 upregulated genes and 2306 downregulated genes, respectively. Common DEGs were composed of upregulated genes and downregulated genes that overlapped, respectively, among the four data sets. In the end, a total of 275 DEGs were obtained, including 76 upregulated DEGs and 199 downregulated DEGs (Figure 1 and Table 1). In addition, in the miRNA expression profile (GSE93415), three DEMs were screened, which were miR-199a-3p/miR-199b-3p with adjust p value = 3.02E-04 (log FC = 2.15675), miR-125b-5p with adjust p value = 1.58E-04 (log FC = 1.98925), and miR-199a-5p adjust p value = 9.38E-06 (log FC = 1.9549).

4.2. GO Terms and KEGG Pathway Enrichment.

In order to further understand the function and mechanism of these DEGs, DAVID was employed to analyze GO terms including biological process (BP), cell component (CC), and molecular function (MF), as well as KEGG pathway, and the results were illustrated. In the BP category, the upregulated gene enrichment GO terms were mainly cell adhesion (GO: 0007155) with $p = 5.09E − 09$ and extracellular matrix organization (GO: 0030198) with $p = 1.33E − 12$, while the downregulated gene enrichment GO terms were mainly oxidation-reduction process (GO: 0055114) with $p = 5.22E − 07$, digestion (GO: 0007586) with $p = 3.21E − 12$, and xenobiotic metabolic process (GO: 0006805) with $p = 1.70E − 07$. Besides, in the CC category, the upregulated gene enrichment GO terms were mainly extracellular region (GO: 0005576) with $p = 5.56E − 11$, extracellular space (GO: 0005615) with $p = 5.09E − 08$, extracellular exosome (GO: 0070062) with $p = 0.0032$, and extracellular matrix (GO: 0031012) with $p = 2.91E − 08$, while the downregulated gene enrichment GO terms were mainly extracellular exosome (GO: 0070062) with $p = 9.58E − 08$, extracellular space (GO: 0005615) with $p = 2.36E − 09$, extracellular region (GO: 0005576) with $p = 0.003517504$, and integral component of plasma membrane (GO: 0005887) with $p = 0.015674589$. Moreover, in the MF category, the upregulated gene enrichment GO terms were mainly protein binding (GO: 0005515) with $p = 0.009744135$, while the downregulated gene enrichment GO terms were mainly zinc ion binding (GO: 0008270) with $p = 0.003724062$ (Figures 2(a) and 2(c)). As for the KEGG pathway, the most significant results were 31 downregulated genes involved in metabolic pathways (Figures 2(b) and 2(d)).

4.3. miRNA-mRNA-TF Regulatory Network Construction.

In order to construct miRNA-mRNA-TF regulatory network, mipathDB V2.0 was used to predict the list of potential target genes of miRNAs and compare the list of lower DEGs, and the results showed that the potential genes of miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p were 61, 71, and 69 in the list of downregulated genes, respectively (Figure 3(a)). The results of the CHEA3 prediction show that RFX6 ranks highest among the TFs involved in

(a)   (b)
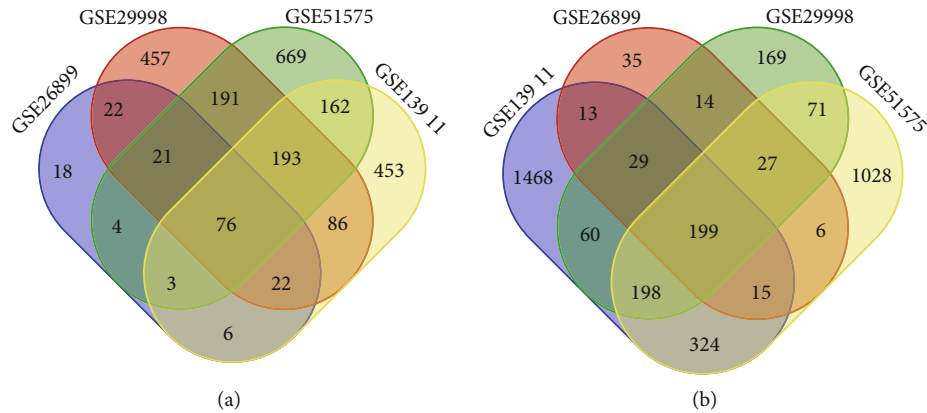
FIGURE 1: Screening of differentially expressed genes (DEGs). After the data sets of GSE26899, GSE29998, GSE51575, and GSE13911 were processed by GEO2R, the common upregulated DEGs (a) and downregulated DEGs (b) were screened out according to the criteria of adjust p value < 0.05 and logFC > 1 and adjust p value < 0.05 and logFC < -1, respectively.

TABLE 1: Screening of differentially expressed genes (DEGs).

Genes	
Downregulated	IGF2BP3 COL1A1 CCNE1 FNDC1 S100A3 COL5A2 THY1 MARVELD3 SPP1 OLR1 ITGA2 TIMP1 BGN COL8A1 CDCA5 CKS2 TEAD4 MMP11 CDH3 MMP3 CRABP2 KRT80 FSCN1 LRP8 LIPG BMP1 PLA2G7 MFAP2 TFAP2A COL4A1 TGIF1 SULF1 IL32 LAMC2 RARRES1 CLDN1 IGSF6 HAVCR2 APOC1 GINS2 COL12A1 ESM1 PLAU SERPINB5 TGM2 FOXC1 FZD2 CST1 CLDN7 NDUFA4L2 HOXA10 ETV4 IER5L CAMK2N1 TNFRSF12A ISG15 ANGPT2 TMEM158 IFITM3 SERPINH1 FAP GJB2 THBS2 S100A10 SERPINE1 MSLN LY6E UPP1 COL6A3 COL10A1 F12 COL5A1 PMEPA1 HOXC6 CTHRC1 FJX1
Upregulated	EPB41L4B COL4A5 ISL1 GHR MAOA ZNF385B PRDM16 KIAA1324 GKN1 GHRL CELA3A APOBEC2 SULT1C2 FCGBP GPRC5C KCNJ13 AKR7A3 IRX3 DGKD SORBS2 KLF4 MT1G CYP3A5 GC ALDH6A1 C5 VSIG1 PDGFD FAM3B KIT NTN4 TNFRSF17 CA9 SH3BGRL2 CBS ADH1C NPY IGFBP2 PIK3C2G SLC9A2 GCNT2 SST PLCXD3 ECHDC3 XYLT2 TPD52L1 ZBTB7C GGT6 PDK4 CXCL17 LYPD6B NR3C2 PDIA2 SLC22A23 TCEAL2 MT1E ANXA10 CA4 ADHFE1 ADAM28 MUC6 OSBPL7 KRT20 HRASLS2 CTSE GIF ABCC5 RNASE1 MT1F SOSTDC1 ACACB RORC PNPLA7 VSIG2 FAM46C CHRM3 TMED6 SH3GL2 SCNN1G CKMT2 PSCA KCNJ16 REG3A TFF1 RPRM CYFIP2 FAM3D CKM FBXL13 RNASE4 ST6GALNAC1 DERL3 PTGER3 COBLL1 DNASE1 CXCL14 LDHD PROM2 FBP2 AKR1B10 SOX21 ARL14 AADAC GSTA4 ATP4A POU2AF1 UGT2B15 CKB CD36 LRRC17 RAP1GAP IGFALS TSC22D3 PIGR CDH2 MAL AQP4 CHGA RAB26 MT1M ESRRG APLP1 VILL TMEM171 CPA2 ABCA8 RAB27B FA2H ADH7 PDILT AMPD1 SULT2A1 GAMT TFF2 CCKBR HPGD RBPMS2 ARRDC4 SCGB2A1 SLC7A8 ADA TMPRSS2 DUOX1 LIPF DPT GUCA2B CYP4X1 CAPN13 TCN1 EPN3 SLC2A12 SORBS1 ENPP5 LIFR NQO1 CD79A CYB5R1 CYP2C18 FGA ALDH3A1 FUT9 PTGR1 AZGP1 ALDH1A1 FMO5 IRX2 CA2 LTF GATA5 ATP4B BCAS1 SIDT2 SLC26A9 ADH1A CHIA METTL7A FAM189A2 SSTR1 PGC MYRIP FABP4 KCNE2 FAM107A GPX3 PPP1R3C MT1H SIGLEC11 PTGS1 ALDOB KLK11 DNER MAMDC2 C1orf116 CLDN18 HDC RASSF6 SELENBP1 CLIC6 TRIM50

downregulated DEGs (Table 2). Moreover, in the list of downregulated DEGs, the potential target genes of miR-199a-3p overlap with the RFX6 downstream genes (RDGs) in 4 genes, namely, SH3GL2, ATP4B, CTSE, and SORBS2. The potential target genes of miR-125b-5p overlapped with RDGs in 7 genes, which were SLC7A8, RNASE4, ESRRG, PGC, MUC6, FAM3B, and FMO5. The potential target genes of miR-199a-5p overlapped with RDGs in 6 genes, which were CHGA, PDK4, TMPRSS2, CLIC6, GPX3, and PSCA (Figure 3(b)). Finally, the miRNA-mRNA-TF regulatory network was constructed based on the above screened 17 mRNAs, 3 miRNAs, and 1 TF by Cytoscape 3.8.2 software (Figure 4(a)).

4.4. Validation of Significant TF in GC Tissues. In order to further verify the accuracy of the above bioinformatics analysis results, the mRNA and protein levels of the significant regulatory factor RFX6 were verified in 4 pairs of GC tissues and normal tissues by RT-qPCR and western blot. RFX6 activates transcription by forming a heterodimer with RFX3 and binding to the X-box in the target gene promoter, so the expression of RFX6 is positively correlated with downstream genes. The validation results showed that the expression of RFX6 mRNA and protein was significantly downregulated, which suggested that the expression of RFX6 was positively correlated with the expression of downstream genes (Figure 4(b)). Therefore, the above bioinformatics
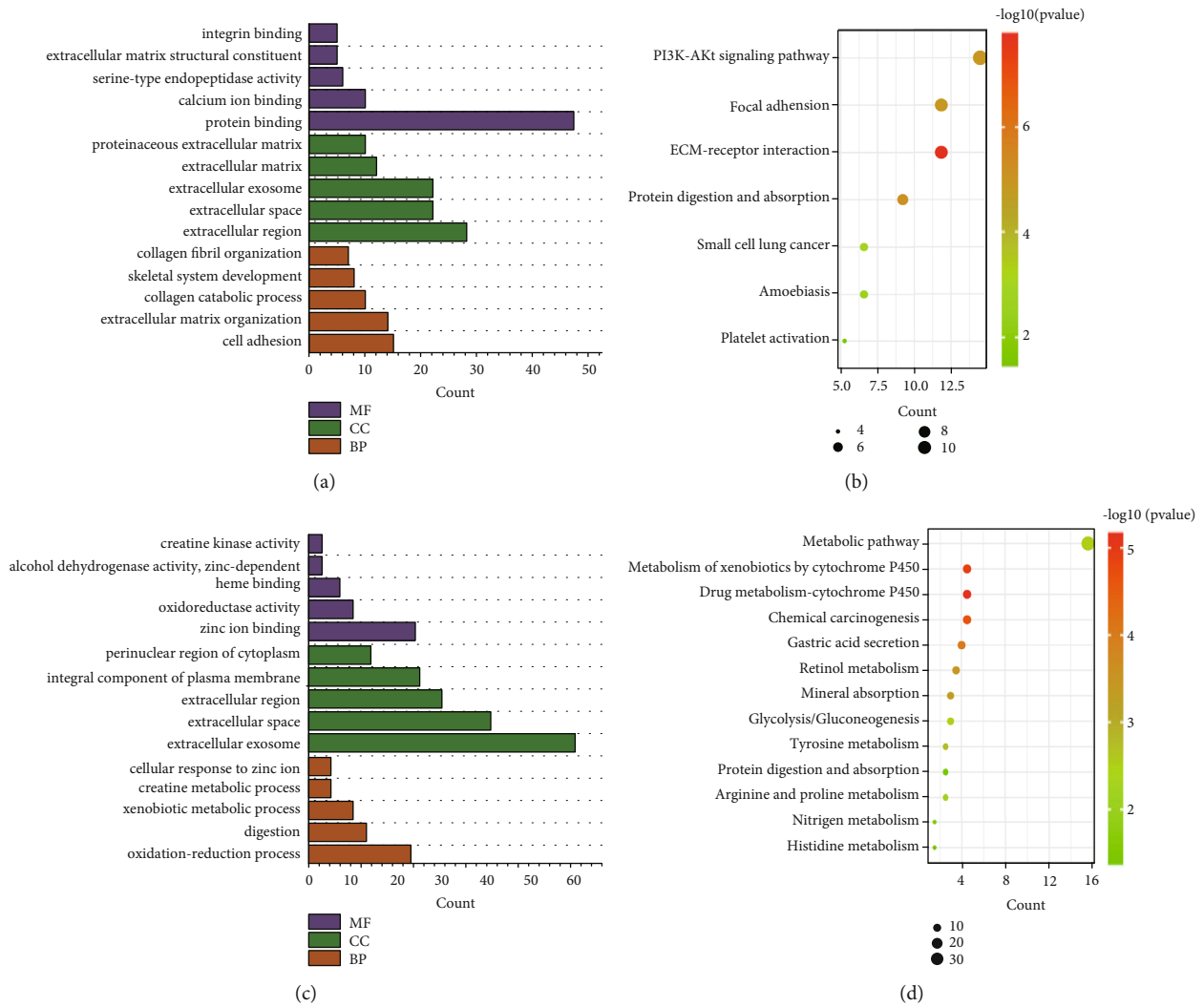
(a)

(b)

(c)

(d)

Figure 2: Annotation of differentially expressed genes (DEGs). The screening results were divided into two groups, upregulated genes and downregulated genes, for enrichment analysis by a database for annotation, visualization, and integrated discovery (DAVID). Gene Ontology (GO) terms included biological processes (BP), cellular component (CC), and molecular function (MF), and the top five most significant items in each group and Kyoto Encyclopedia of Genes and Genomes (KEGG) results were illustrated. (a) and (b) The upregulated gene group. (c) and (d) The downregulated gene group.
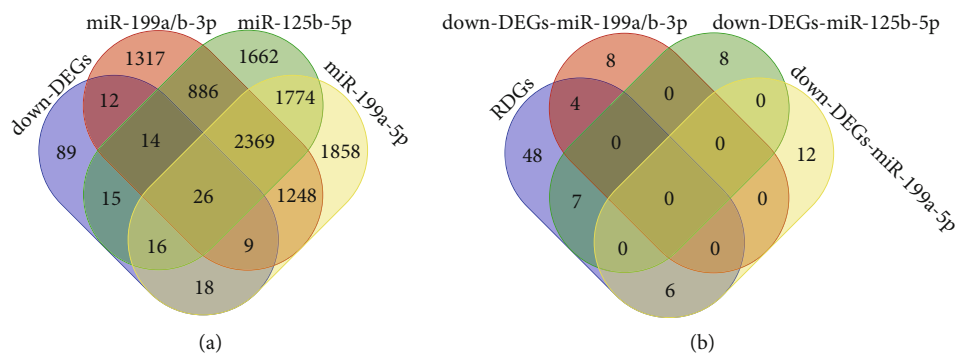

(a)

(b)

Figure 3: Screening of miRNA-mRNA-TF regulatory network nodes. (a) Among the 199 common downregulated DEGs (down-DEGs) screened from the mRNA expression profile, the overlapping genes with the potential targets of differentially expressed miRNAs (DEMs) were identified. (b) The DEMs targets overlap with the RFX6 downstream genes (RDGs).

Table 2: The top 10 transcription factors of prediction results.

Rank	Transcription factors	Mean rank
1	RFX6	12
2	FOXA3	14
3	TFCP2L1	15.33
4	GATA5	18.67
5	MYRFL	26
6	ISX	26
7	FOXA2	26.5
8	ISL1	30.67
9	CDX1	31.33
10	FOXA1	34

This prediction is the result obtained by integrating multiple databases. Different databases have different rankings, so the higher the mean rank is, the more accurate it will be.

analysis results are proved to be credible from the perspective of experiment.

5. Discussion

According to the data released by the Global Cancer Observatory (GCO), in 2020, there were 1,9292,789 new cancer patients in the world, among which 1,089,103 were GC patients, accounting for 5.6%, ranking the fifth. The research on the mechanism of GC and the development of new drugs have always been the shining research direction of medical and scientific researchers.

In this study, DEGs and DEMs were screened out from the existing mRNA and miRNA expression profiles, and the potential targets of DEMs and TF of DEGs were predicted through the database, so as to construct miRNA-mRNA-TF regulatory network and further understand the pathogenesis of GC. First of all, in the 4 mRNA expression profiles containing 210 GC tissues and 118 normal tissues, 275 DEGs were screened using adjust p value < 0.05, |logFC| >1 as the criterion, among which 76 genes were significantly upregulated in GC tissues compared with normal tissues, while the remaining 199 genes were significantly downregulated in GC tissues compared with normal tissues. Second, to further understand the mechanism of DEGs in biological processes, GO terms and KEGG pathway enrichment analysis were conducted for the upregulated genes and downregulated genes, respectively, and the results showed that the downregulated genes with a higher proportion of DEGs had more obvious enrichment. Then, using adjust p value < 0.01, |logFC| >1.9 as the criterion, three miRNAs, including miR-199a-3p/miR-199b-3p (adjust p value = 3.02E-04, log FC = 2.15675), miR-125b-5p (adjust p value = 1.58E-04, log FC = 1.98925), and miR-199a-5p (adjust p value = 9.38E-06, log FC = 1.9549), were screened out from a miRNA expression profile containing 20 GC tissues and 20 normal tissues, which were overexpressed in GC tissues relative to normal tissues. After screening the DEMs, to explore the regulatory relationship between miRNA and mRNA and construct miRNA-mRNA regula-

tory network, the potential targets of miRNAs were predicted by the mipathDB V2.0 data platform, and some of the 199 low-expressed genes that might be regulated by miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p were obtained by using the Venn diagram. Finally, the prediction of TFs, 199 significantly low-expressed genes were used to predict key TFs that played an important role in the progression of GC cases. The prediction results showed that RFX6 ranked the highest in the average of multiple databases, suggesting that RFX6 may play a key regulatory role in the pathogenesis of GC. RT-qPCR results confirmed that the expression of RFX6 was significantly changed in clinical GC tissues.

miRNAs affect gene expression by regulating the translation and degradation of mRNA, and the promotion or inhibition effect of this mechanism on cancer cells has been reported in recent years [33]. It was found by retrieval that the significantly upregulated expression of miR-199a-3p, miR-125-5p, and miR-199a-5p in GC cells and tissues analyzed in this study has been reported many times. In 2003, a computational prediction of miR-199a's identity was made based on the conservatism of miR-199a among humans, mouse, and puffer fish [34]. The expression of miRNA in zebrafish was verified, and the ends of the miRNA were localized by cloning. The two miRNA sequences were named miR-199a and miR-199a∗ (from the 3 'arm), respectively. It has been reported that both of these two miRNA expression forms are expressed in humans and are named as miR-199a-5p and miR-199a-3p, respectively [35]. Functional studies have shown that miR-199a-3p can significantly promote the proliferation and inhibit apoptosis of GC cells in vivo and in vitro. It has been found in clinical studies that the overexpression of miR-199a-3p is associated with tumor size, Lauren stage, depth of invasion, lymph node metastasis, distant metastasis, TNM stage, and prognosis [36]. Moreover, in stage I, II, and III tumors, the high expression of miR-199a-3p is associated with a significantly reduced 5-year survival rate. In addition, luciferase report assay demonstrated that miR-199a-3p directly targeted the expression level of ZHX1 regulatory protein. Therefore, researchers hypothesize that the oncogenic activity of miR-199a-3p may be related to the direct targeting and inhibition of ZHX1 [37]. Similar to miR-199a-3p, miR-199a-5p can promote the migration and invasion of GC cells, and its expression level is related to tumor diameter, lymph node metastasis, and TNM stage. However, the difference was that the expression level of miR-199a-5p was not associated with Lauren staging and distant metastasis [38]. Different from miR-199a-3p and miR-199a-5p, the role of miR-125b-5p in the pathogenesis of GC is not fully understood, but the identification results at the cell level and tissue level indicate that miR-125b-5p is significantly upregulated during the development of GC [39]. In addition, in a study on drug resistance, it was found that STAT3 may be a potential target of miR-125b-5p, and the signaling cascades involved in the regulation of chemotherapy drug resistance in GC [40].

In addition to these identified miRNAs, other mRNAs in the miRNA-mRNA-TF regulatory network have also been individually reported, in addition to CTSE, SORBS2,
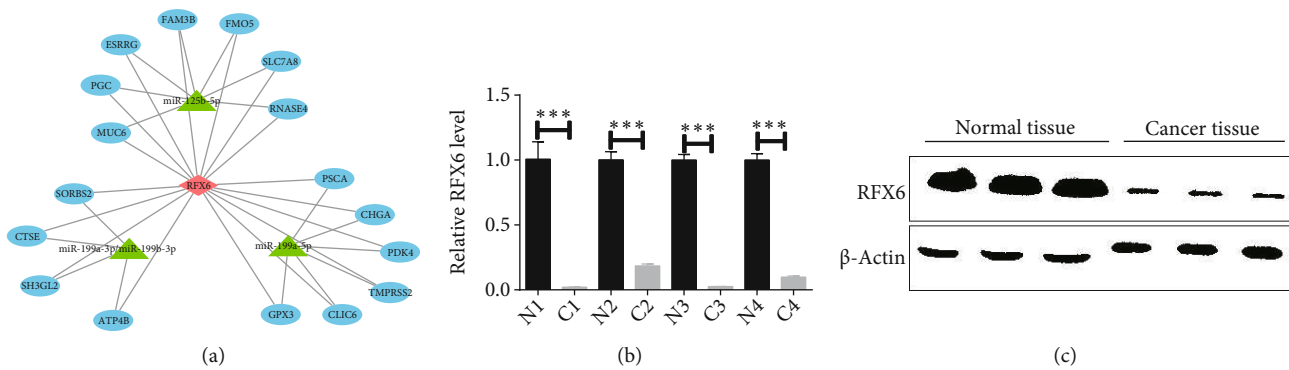
Figure 4: Construction of miRNA-mRNA-TF regulatory network and validation of RFX6 expression level. (a) Three miRNAs (miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p) and transcription factor (TF) RFX6 coregulate gene expression in the miRNA-mRNA-TF regulatory network, and miRNAs negatively regulate their expression levels by binding to the mRNA 3′-UTR region of 17 genes. RFX6 activates transcription by forming a heterodimer with RFX3 and binding to the X-box in the target gene promoter. (b) and (c) Compared with normal gastric tissue, RFX6 was significantly downregulated in GC tissue.

TMPRS, and RNASE4. For example, as a negative regulator of Wnt signaling pathway, ESRRG plays a tumor suppressor role in GC [41]. Moreover, the β subunit of the gastric H$^+$, K$^+$-ATPase encoded by ATPase H$^+$/K$^+$ Transporting Beta Subunit (ATP4B) are decreased in GC cells and tissues due to the interaction between DNA methylation and histone acetylation [42, 43]. Previous studies have found that if ATP4B expression level is restored, it can inhibit the proliferation, activity, migration, invasion, tumorigenicity, and induction of apoptosis of GC cells, and its tumor-suppressive effect may be played through the regulation of mitochondrial metabolism and apoptosis pathway [44].

Besides, pepsinogen C (PGC) belongs to the aspartic protease family, which is expressed in situ in gastric mucosa, expressed in serum, and expressed ectopic. It is secreted by the gastric host cells and has an activation effect on pepsin C and can digest polypeptides and amino acids [45]. In the process from superficial gastritis to atrophic gastritis and finally to the formation of GC, the positive level of PGC continues to decline, which indicates that the in situ expression of PGC may have a negative correlation with the formation of GC [46]. And the results showed that the expression of MUC6 was decreased in GC tissues, but surprisingly, there was no significant difference between the expression of MUC6 in GC tissues and the gender, tumor site, lymphatic infiltration, clinical stage, metastasis, Lauren intestinal type, diffuse type, and mixed type GC tissues [47].

As for the only TF, RFX6, a member of the RFX (regulatory factor X-box binding) family of winged-helix transcription factors [48], which is downstream of Neurog3 and upstream of many other islet transcription factors in the islet development factor hierarchy, is required for differentiation of four of the five islet cell types [49]. In prostate cancer, on the one hand, clinical data showed that the upregulated expression of RFX6 was associated with the risk of tumor progression, metastasis, and biochemical recurrence [50]. In addition, RFX6 has also been reported to be differentially expressed in melanoma, liver cancer, GC, and other cancer tissues, but the further molecular mechanism remains to be explored.

6. Conclusion

In conclusion, in this study, 76 upregulated DEGs, 199 downregulated DEGs, and 3 upregulated DEMs (miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p) were identified from the expression profiles of mRNA (GSE26899, GSE29998, GSE51575, and GSE13911) and 1 miRNA (GSE93415), respectively. Through database prediction and gene list comparison, it was found that among the 199 downregulated DEGs, 61, 71, and 69 genes were the potential targets of miR-199a-3p/miR-199b-3p, miR-125b-5p, and miR-199a-5p, respectively. 199 downregulated DEGs were used as the gene list for the prediction of key TFs, and the results showed that RFX6 ranked the highest. The potential target overlap genes of miR-125b-5p and miR-199a-5p were 4 (SH3GL2, ATP4B, CTSE, and SORBS2), 7 (SLC7A8, RNASE4, ESRRG, PGC, MUC6, Fam3B, and FMO5), and 6 (CHGA, PDK4, TMPRSS2, CLIC6, GPX3, and PSCA), respectively. Finally, we constructed a miRNA-mRNA-TF regulatory network based on the above 17 mRNAs, 3 miRNAs, and 1 TF and verified by RT-qPCR and western blot results that the expression of RFX6 was downregulated in GC tissues. These identified miRNAs, mRNAs, and TF have a certain reference value for further exploration of the regulatory mechanism of GC. It is worth noting that, given the complexity of miRNA and TF crosstalk, these regulatory relationships need to be treated with caution.

Data Availability

The datasets of GSE26899, GSE29998, GSE51575, GSE13911, and GSE93415 were downloaded from the GEO database.

Conflicts of Interest

The authors have declared that no competing interest exists.

Authors' Contributions

Zhenjie Fu and Yuqin Xu jointly completed the experiment and the writing and modification of the paper. Yan Chen and Hang Lv provided assistance for the experiment process. Guiping Chen and Yitao Chen provided guidance for the collection of experimental samples and the overall design. Zhenjie Fu and Yuqin Xu have contributed equally to this work and share first authorship. Guiping Chen and Yitao Chen authors have contributed equally to this work and share corresponding authorship.

Acknowledgments

References

[1] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] M. E. Camacho-Moll, J. Macdonald, L. H. J. Looijenga et al., "The oncogene Gankyrin is expressed in testicular cancer and contributes to cisplatin sensitivity in embryonal carcinoma cells," *BMC Cancer*, vol. 19, no. 1, p. 1124, 2019.

[3] L. Zhang, Q. Zhou, Q. Qiu et al., "CircPLEKHM3 acts as a tumor suppressor through regulation of the miR-9/BRCA1/DNAJB6/KLF4/AKT1 axis in ovarian cancer," *Molecular Cancer*, vol. 18, no. 1, p. 144, 2019.

[4] J. Wang, J. Chen, and S. Sen, "MicroRNA as biomarkers and diagnostics," *Journal of Cellular Physiology*, vol. 231, no. 1, pp. 25–30, 2016.

[5] A. Etheridge, I. Lee, L. Hood, D. Galas, and K. Wang, "Extracellular microRNA: a new source of biomarkers," *Mutation Research*, vol. 717, no. 1-2, pp. 85–90, 2011.

[6] J. R. Chevillet, I. Lee, H. A. Briggs, Y. He, and K. Wang, "Issues and prospects of microRNA-based biomarkers in blood and other body fluids," *Molecules*, vol. 19, no. 5, pp. 6080–6105, 2014.

[7] K. Kai, R. L. Dittmar, and S. Sen, "Secretory microRNAs as biomarkers of cancer," *Seminars in Cell & Developmental Biology*, vol. 78, pp. 22–36, 2018.

[8] R. C. Lee, R. L. Feinbaum, and V. Ambros, "The C. elegans heterochronic gene _lin-4_ encodes small RNAs with antisense complementarity to *lin-14*," *Cell*, vol. 75, no. 5, pp. 843–854, 1993.

[9] B. Wightman, I. Ha, and G. Ruvkun, "Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in C. elegans," *Cell*, vol. 75, no. 5, pp. 855–862, 1993.

[10] R. C. Lee and V. Ambros, "An extensive class of small RNAs in Caenorhabditis elegans," *Science*, vol. 294, no. 5543, pp. 862–864, 2001.

[11] N. C. Lau, L. P. Lim, E. G. Weinstein, and D. P. Bartel, "An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans," *Science*, vol. 294, no. 5543, pp. 858–862, 2001.

[12] P. Hydbring and G. Badalian-Very, "Clinical applications of microRNAs," *F1000Research*, vol. 2, p. 136, 2013.

[13] I. Alvarez-Garcia and E. A. Miska, "MicroRNA functions in animal development and human disease," *Development*, vol. 132, no. 21, pp. 4653–4662, 2005.

[14] A. Esquela-Kerscher and F. J. Slack, "Oncomirs – microRNAs with a role in cancer," *Nature Reviews. Cancer*, vol. 6, no. 4, pp. 259–269, 2006.

[15] X. Liu, A. Kwong, A. Sihoe, and K. M. Chu, "Plasma miR-940 may serve as a novel biomarker for gastric cancer," *Tumour Biology*, vol. 37, no. 3, pp. 3589–3597, 2016.

[16] G. Zhao, T. Jiang, Y. Liu et al., "Droplet digital PCR-based circulating microRNA detection serve as a promising diagnostic method for gastric cancer," *BMC Cancer*, vol. 18, no. 1, p. 676, 2018.

[17] L. Cui, Y. Lou, X. Zhang et al., "Detection of circulating tumor cells in peripheral blood from patients with gastric cancer using piRNAs as markers," *Clinical Biochemistry*, vol. 44, no. 13, pp. 1050–1057, 2011.

[18] Y. Naito, K. Yasuno, H. Tagawa et al., "MicroRNA-145 is a potential prognostic factor of scirrhous type gastric cancer," *Oncology Reports*, vol. 32, no. 4, pp. 1720–1726, 2014.

[19] M. V. Iorio and C. M. Croce, "MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review," *EMBO Molecular Medicine*, vol. 4, no. 3, pp. 143–159, 2012.

[20] C. E. Condrat, D. C. Thompson, M. G. Barbu et al., "miRNAs as biomarkers in disease: latest findings regarding their role in diagnosis and prognosis.," *Cells*, vol. 9, no. 2, p. 276, 2020.

[21] M. Lambert, S. Jambon, S. Depauw, and M. H. David-Cordonnier, "Targeting transcription factors for cancer treatment," *Molecules*, vol. 23, no. 6, p. 1479, 2018.

[22] S. Polvani, M. Tarocchi, S. Tempesti, L. Bencini, and A. Galli, "Peroxisome proliferator activated receptors at the crossroad of obesity, diabetes, and pancreatic cancer," *World Journal of Gastroenterology*, vol. 22, no. 8, pp. 2441–2459, 2016.

[23] B. R. B. Pires, R. Silva, G. M. Ferreira, and E. Abdelhay, "NF-kappaB: two sides of the same coin," *Genes*, vol. 9, no. 1, p. 24, 2018.

[24] G. Badis, M. F. Berger, A. A. Philippakis et al., "Diversity and complexity in DNA recognition by transcription factors," *Science*, vol. 324, no. 5935, pp. 1720–1723, 2009.

[25] D. A. Bernstein, "Identification of small molecules that disrupt SSB-protein interactions using a high-throughput screen," *Methods in Molecular Biology*, vol. 922, pp. 183–191, 2012.

[26] L. Song, D. Liu, J. He et al., "SOX1 inhibits breast cancer cell growth and invasion through suppressing the Wnt/β-catenin signaling pathway," *APMIS*, vol. 124, no. 7, pp. 547–555, 2016.

[27] A. Rad, S. Esmaeili Dizghandi, M. R. Abbaszadegan, N. Taghechian, M. Najafi, and M. M. Forghanifard, "SOX1 is correlated to stemness state regulator SALL4 through progression and invasiveness of esophageal squamous cell carcinoma," *Gene*, vol. 594, no. 2, pp. 171–175, 2016.

[28] A. H. Sinclair, P. Berta, M. S. Palmer et al., "A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif," *Nature*, vol. 346, no. 6281, pp. 240–244, 1990.

[29] G. E. Schepers, R. D. Teasdale, and P. Koopman, "Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families," *Developmental Cell*, vol. 3, no. 2, pp. 167–170, 2002.

[30] P. A. Seymour, "Sox9: a master regulator of the pancreatic program," *The Review of Diabetic Studies*, vol. 11, no. 1, pp. 51–83, 2014.

[31] A. J. Abadi, A. Zarrabi, F. Hashemi et al., "The role of SOX family transcription factors in gastric cancer," *International Journal of Biological Macromolecules*, vol. 180, pp. 608–624, 2021.

[32] A. Reiner, D. Yekutieli, and Y. Benjamini, "Identifying differentially expressed genes using false discovery rate controlling procedures," *Bioinformatics*, vol. 19, no. 3, pp. 368–375, 2003.

[33] G. Hutvágner and P. D. Zamore, "A microRNA in a multiple-turnover RNAi enzyme complex," *Science*, vol. 297, no. 5589, pp. 2056–2060, 2002.

[34] L. P. Lim, M. E. Glasner, S. Yekta, C. B. Burge, and D. P. Bartel, "Vertebrate microRNA genes," *Science*, vol. 299, no. 5612, p. 1540, 2003.

[35] P. Landgraf, M. Rusu, R. Sheridan et al., "A mammalian microRNA expression atlas based on small RNA library sequencing," *Cell*, vol. 129, no. 7, pp. 1401–1414, 2007.

[36] L. Li, Y. P. Mou, Y. Y. Wang, H. J. Wang, and X. Z. Mou, "miR-199a-3p targets ETNK1 to promote invasion and migration in gastric cancer cells and is associated with poor prognosis," *Pathology-Research and Practice*, vol. 215, article 152511, 2019.

[37] Z. Wang, X. Ma, Q. Cai, X. Wang, B. Yu, and Q. Cai, "miR-199a-3p promotes gastric cancer progression by targeting ZHX1," *FEBS Letters*, vol. 588, no. 23, pp. 4504–4512, 2014.

[38] X. J. He, Y. Y. Ma, S. Yu et al., "Up-regulated miR-199a-5p in gastric cancer functions as an oncogene and targets klotho," *BMC Cancer*, vol. 14, no. 1, p. 218, 2014.

[39] X. Chen, X. Li, X. Peng et al., "Use of a four-miRNA panel as a biomarker for the diagnosis of stomach adenocarcinoma," *Disease Markers*, vol. 2020, Article ID 8880937, 12 pages, 2020.

[40] P. Deng, M. Sun, W. Y. Zhao et al., "Circular RNA circVAPA promotes chemotherapy drug resistance in gastric cancer progression by regulating miR-125b-5p/STAT3 axis," *World Journal of Gastroenterology*, vol. 27, no. 6, pp. 487–500, 2021.

[41] M. H. Kang, H. Choi, M. Oshima et al., "Estrogen-related receptor gamma functions as a tumor suppressor in gastric cancer," *Nature Communications*, vol. 9, no. 1, p. 1920, 2018.

[42] S. Lin, B. Lin, X. Wang et al., "Silencing of ATP4B of ATPase H(+)/K(+) transporting beta subunit by intragenic epigenetic alteration in human gastric cancer cells," *Oncology Research*, vol. 25, no. 3, pp. 317–329, 2017.

[43] U. M. Raja, G. Gopal, and T. Rajkumar, "Intragenic DNA methylation concomitant with repression of ATP4B and ATP4A gene expression in gastric cancer is a potential serum biomarker," *Asian Pacific Journal of Cancer Prevention*, vol. 13, no. 11, pp. 5563–5568, 2012.

[44] Y. Pan, X. Wang, Y. He et al., "Tumor suppressor ATP4B serve as a promising biomarker for worsening of gastric atrophy and poor differentiation," *Gastric Cancer*, vol. 24, no. 2, pp. 314–326, 2021.

[45] M. I. Hassan, A. Toor, and F. Ahmad, "Progastriscin: structure, function, and its role in tumor progression," *Journal of Molecular Cell Biology*, vol. 2, no. 3, pp. 118–127, 2010.

[46] S. Shen, J. Jiang, and Y. Yuan, "Pepsinogen C expression, regulation and its relationship with cancer," *Cancer Cell International*, vol. 17, no. 1, p. 57, 2017.

[47] A. Z. Khattab, W. A. Nasif, and M. Lotfy, "MUC2 and MUC6 apomucins expression in human gastric neoplasm: an immunohistochemical analysis," *Medical Oncology*, vol. 28, Suppl 1, pp. 207–213, 2011.

[48] S. Aftab, L. Semenec, J. S. C. Chu, and N. Chen, "Identification and characterization of novel human tissue-specific RFX transcription factors," *BMC Evolutionary Biology*, vol. 8, no. 1, pp. 1–11, 2008.

[49] S. B. Smith, H. Q. Qu, N. Taleb et al., "Rfx6 directs islet formation and insulin production in mice and humans," *Nature*, vol. 463, no. 7282, pp. 775–780, 2010.

[50] Q. Huang, T. Whitington, P. Gao et al., "A prostate cancer susceptibility allele at 6q22 increases *RFX6* expression by modulating HOXB13 chromatin binding," *Nature Genetics*, vol. 46, no. 2, pp. 126–135, 2014.