Differential Expression and Interaction network construction of Noncoding RNA and mRNA in Heart Tissue associated with Tetralogy of Fallot

Hongdan Wang[1,2,3], Cunying Cui[4,5], Yanan Li[4,5], Yuanyuan Liu[4,5], Taibing Fan[4,6] Bangtian Peng[4,6] Lin Liu[4.5*]

[1] Institute of Henan Provincial Medical Genetic, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou, China

[2] Key Laboratory Shaanxi Province for Craniofacial Precision Medicine Research, College of Stomatology, Xi'an Jiaotong University, Xi'an, China

[3] College of Forensic Science, Xi'an Jiaotong University Health Science Center, Xi'an Jiaotong University, Xi'an, China

[4] Department of Ultrasound, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Fuwai Central China Cardiovascular Hospital, Zhengzhou, China

[5] Key Laboratory of Congenital Heart Disease, Fuwai Central China Cardiovascular Hospital, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou, China

[6] Pediatric Heart Center, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Fuwai Central China Cardiovascular Hospital, Zhengzhou, China

* Corresponding authour: Lin Liu, liulin_819@126.com

ABSTRACT

Tetralogy of Fallot (TOF) is still the most common and complicated cyanotic congenital heart defect of all congenital heart diseases with a 10% incidence. Surgery repair is often necessary in infancy. The etiology of TOF is complex and genetic and epigenetic mechanisms such as chromosomal abnormalities, gene mutations, nucleic acid modifications, non-coding RNA, and circular RNA(circRNA) play an important role in its occurrence. RNA not only plays an auxiliary role of genetic information carrier, but also plays a more important role in various regulatory functions. There are few studies on the action mechanism of non-coding RNA. Aim to gain more in-depth knowledge of TOF, we collected tissue samples of the right ventricular outflow tract of 5 TOF children with no other intracardiac and extracardiac malformations and 5 normal fetuses. We systematically analyzed the specific long non-coding RNA (lncRNA), microRNA(miRNA), circRNA and messenger RNA(mRNA) profiles of TOF. To our knowledge, there are no reports of genome-wide study of transcriptome in TOF and we first obtained meaningful differentially expressed lncRNAs, miRNAs, circle RNA and mRNAs.

KEYWORDS long non-coding RNAs, microRNA, circular RNA, messenger RNA, Tetralogy of Fallot

37 The total incidence of neonatal defects in China is 5.6‰, and 0.9 million neonates are born with
38 defects every year. The incidence of neonatal defects is increasing annually. The heart is the first
39 major internal organ in the process of embryogenesis and is important for the survival of an embryo.
40 Congenital heart disease (CHD) has ranked first for five consecutive years in neonates. The overall
41 incidence of CHD in the population is 1‰–13‰. In China, the incidence of CHD is 4‰–8‰.
42 Children with severe CHD die in childhood, thereby causing psychological trauma and economic
43 burden to families and society. Tetralogy of Fallot (TOF) remains the most common complex
44 cyanotic congenital heart defect among all congenital heart diseases given its 10% incidence
45 (BEDARD *et al.* 2009). Surgery repair is often necessary in infancy, and the perioperative mortality is
46 less than 5%. However, long-term follow-up has shown that patients who have undergone total repair
47 of TOF are often at high risk for restenosis of the right ventricular outflow tract (RVOT), pulmonary
48 regurgitation, and ventricular dysfunction, and late sudden cardiac death may occur in 0.5% to 6% of
49 patients due to ventricular tachycardia or other unknown causes(MASSIN *et al.* 2011; ZHANG *et al.*
50 2013; LATUS *et al.* 2015; SHIN *et al.* 2016) . The exact etiology of TOF remains unclear, and its
51 pathogenesis is complex. Genetic and epigenetic mechanisms, such as chromosomal abnormality,
52 gene mutation, nucleic acid modification, noncoding RNA (ncRNA), and circular RNA (circRNA),
53 play an important role in its occurrence. At present, chromosomal abnormalities, gene mutations, and
54 other genetic mechanisms have been deeply studied and widely used in the diagnosis and treatment
55 of clinical diseases. However, the action mechanism of ncRNAs on TOF needs to be further studied.

56 Recently, genome-wide transcriptional studies have represented that only almost 1% of the
57 human genome produces biologically meaningful RNA transcripts, while a much larger proportion of
58 the genome is transcribed into ncRNAs(PONTING *et al.* 2009). RNA not only plays an auxiliary role
59 as a carrier of genetic information but also demonstrates various regulatory functions. In comparison
60 with the diversity of RNA transcripts which encode for proteins, the functions of a small number of
61 ncRNAs have been demonstrated. ncRNAs refer to RNA that is not translated into protein. ncRNAs
62 involved in gene regulation include miRNA, long-chain ncRNA (lncRNA), circRNA, PIWI
63 interacting RNA, and small interfering RNA(BAJPAI *et al.* 2010). Among them, lncRNAs,
64 microRNAs (miRNAs), and circRNAs have attracted the most attention.

65 lncRNAs can competitively bind miRNA through competitive endogenous RNA (ceRNA) to
66 relieve the translation inhibition of miRNA on the target gene messenger RNA (mRNA) (SALMENA
67 *et al.* 2011). The expression of lncRNAs are not only cell type and tissue-specific. Some are
68 expressed only at specific stages of eukaryotic development. Sufficient evidence has shown that
69 lncRNAs are key molecules that regulate gene expression, chromatin number maintenance,
70 transcriptional regulation, and post regulation (TANO AND AKIMITSU 2012; LI *et al.* 2013). lncRNAs
71 play significant parts in various organisms' physiological and pathological processes, such as
72 embryonic development, organogenesis, and tumorigenesis (JARIWALA AND SARKAR 2016). In recent
73 years, increasing attention has been paid to research on lncRNAs in the cardiovascular field. Studies
74 have demonstrated that lncRNAs play a crucial part in promoting cardiomyocyte differentiation and
75 maintaining cardiac function (SCHONROCK *et al.* 2012; PHILIPPEN *et al.* 2015; GAO *et al.* 2017).

76 miRNAs are a kind of endogenous noncoding, single-chain, and small molecular RNA composed
77 of approximately 22 nucleotides. They widely exist in eukaryotic cells and are highly conserved.
78 miRNAs are a multifunctional regulator of gene expression in higher eukaryotes. The regulation
79 mode is that the targeted binding of mRNA is mainly conducted, thereby further inhibiting or
80 blocking translation(YATES *et al.* 2013). Heart development is a process of polygenic regulation.
81 Studies have shown that miRNAs can participate in the expression, transcription, splicing, and

82 translation of important genes in the development of the heart and in the regulation of related
83 pathways, which may lead to heart defects (O'BRIEN *et al.* 2012).

84 circRNA is a kind of special RNA molecule. Different from traditional linear RNA, circRNA has
85 a closed ring structure, which is not affected by RNA exonuclease and is difficult to degrade. Its
86 expression is more stable compared with traditional linear RNA. New evidence suggests that
87 circRNAs are widely expressed in mammalian cells and exhibit cell or tissue specificity. circRNAs
88 have been proven to get involved in regulating various biological processes(YU AND KUO 2019).
89 circRNAs can also affect cell physiological processes through various molecular mechanisms. They
90 can be used as mediator of miRNA and RNA binding protein to regulate protein expression and
91 translation. Similar to lncRNAs, circRNAs are rich in miRNA binding sites. They can play the role
92 of miRNA through mechanism competition of ceRNA and relieve miRNA inhibition on their target
93 gene. As such, the target gene's expression level can be increased. circRNAs play a significant
94 regulatory part in diseases through their interaction with miRNAs associated with diseases or other
95 functional elements (HAN *et al.* 2018). However, our understanding of the regulation and function of
96 circRNAs is limited.

97 TOF is the most common cyanotic congenital heart disease. Most of TOF have nothing to do with
98 chromosome or microduplication/deletion. The pathogenesis of simple TOF has not been elucidated.
99 However, ncRNAs may play a significant regulatory part in heart formation. Little is known
100 regarding the changes in the transcriptome and the possible associations with TOF. Thus, we
101 systematically analyzed the specific lncRNA, miRNA, circRNA, and mRNA profiles of TOF to gain
102 more in-depth knowledge. As we know, no genome-wide study has concentrated on the
103 transcriptome in TOF. This is the first work to obtain meaningful differentially expressed lncRNAs,
104 miRNAs, circle RNA, and mRNAs.

105

Materials and methods

Subjects

108 The case group included five children with TOF (two boys and three girls, aged 6–22 months).
109 Cyanosis was found in all children, and rough systolic murmur was found between the second and
110 fourth intercostals at the left border of the sternum by auscultation. Echocardiography showed
111 ventricular septal defect, aortic straddle, RVOT obstruction, pulmonary artery stenosis, right
112 ventricular enlargement, and right ventricular anterior wall thickening. The children and their parents
113 had no other intracardiac and extracardiac malformations. In the control group, five fetuses, including
114 three males and two females with gestational age of 25–28 weeks, had normal hearts. Given fetal
115 distress, cord twist, and placental abruption, odinopoeia was conducted.

Extraction of RNA

117 The Total RNA including small RNA was collected from 10–20 mg of frozen tissue of the right
118 ventricle by using a TRIzol reagent (Invitrogen, Gaithersburg, MD, USA), which was purified with a
119 mirVana miRNA Isolation Kit (Ambion, Austin, TX, USA) according to the instructions of the
120 manufacturer. RNA's purity and concentration were defined by OD260/280 by using a
121 spectrophotometer (NanoDrop ND-1000). RNA integrity was defined by 1% formaldehyde
122 denaturing gel electrophoresis. RNA integrity was defined by capillary electrophoresis with an RNA

123   6000 Nano Lab-on-a-Chip kit and Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA).
124   Only if the integrity number values of RNA were more than 6, it was analyzed later.

Microarray

126   The specific lncRNA, mRNA, miRNA, and circle RNA profiles of TOF were analyzed by using
127   Agilent human lncRNA+mRNA Array v4.0, Agilent human miRNA Array V21.0, and Agilent
128   human circRNA Array v2.0.

129   lncRNA and mRNA profiling was performed using the Agilent human lncRNA+mRNA Array
130   V4.0 designed with four identical arrays per slide(4×180k), which included the number of probes of
131   human lncRNAs and human mRNAs being approximately 34,000 and 41,000 respectively. lncRNA
132   and mRNA target sequences were selected from diverse databases such as GENCODE/ENSEMBL,
133   Human LincRNA Catalog [1], RefSeq, USCS, NRED (ncRNA Expression Database), LNCipedia, H-
134   InvDB, LncRNAs-a (Enhancer-like), Antisense ncRNA pipeline, etc.. Repeated probes were used to
135   detect each RNA for two times. The control probes were contained in the array.

136   The Agilent miRNA array ($8 \times 60K$) was used to detect miRNA profiling, which included 2549
137   probes of human mature miRNAs and the miRNA target sequences were from the miRBase R21.0.
138   Repeated probes were used to detect each miRNA for times. The control probes were contained in
139   the array.

140   The Agilent human circRNA Array V2.0 ($4 \times 180K$) was used to detect circRNA profiling,
141   which included approximately 170,340 probes of human circRNAs and the circRNA target
142   sequences were from circBase and DeepBase. Long and short probes were used to detect each
143   circRNA at the same time. The control probes were contained in the array.

RNA amplification, labeling, and hybridization

145   *Agilent human lncRNA+mRNA Array v4.0*

146   Cy5 and Cy3-dCTP was used to label the cDNA) by using the Eberwine's linear RNA amplification
147   method and subsequent enzymatic reaction (PATTERSON *et al.* 2006). By using the CbcScript reverse
148   transcriptase with cDNA synthesis system according to the manufacturer's protocol with T7 Oligo
149   (dT) and T7 Oligo (dN), double-stranded cDNAs were synthesized using 1 µg of total RNA.

150   DNA polymerase and RNase H were used for double-stranded cDNA (dsDNA) synthesis PCR
151   NucleoSpin Extract II Kit eluted with 30 µL elution buffer was used for the purifying of the dsDNA
152   products. The products were concentrate to 16 µL and conducted with vitro transcription reactions in
153   40 µL using the T7 Enzyme Mix at 37 °C for 14 h. the RNA Clean-up Kit was used for purifying the
154   amplified cRNA.

155   Klenow enzyme labeling was performed next by using CbcScript II reverse transcriptase. In
156   short, the amplified RNA (2 µg) was mixed with random nanomer (4 µg), denatured for 5 min at 65
157   °C, and immediately cooled on the ice. Then, 4× first-strand buffer (5 µL), o0.1M DTT (2 µL), and
158   CbcScript II reverse transcriptase (1.5 µL) were mixed and added. The mixtures were incubated for
159   10 min at 25 °C ad then for 90 min at 37 °C. The purification method of cDNA products was as
160   follows. cDNA was added with random nanomer (4 µg), 95 °C for 3 min, and cooled on ice
161   immediately for 5 min. Then, Klenow buffer (5 µL), dNTP, and Cy5-dCTP or Cy3-dCTPwere mixed
162   and added. Then, added Klenow enzyme for1.2 µL, and incubated for 90 min at 37 °C. Labeled

4

163　cDNA was purifiedas follows. The hybridization solution of the labeled controls samples and test
164　samples were dissolved in 80 μL including 3× SSC, 0.2% SDS, 5× Denhardt's solution, and 25%
165　formamide. DNA in the hybridization solution was denatured at for 3 min 95 °C before loading on
166　the microarray. Arrays were hybridized overnight at 42 °C in an Agilent Hybridization Oven at the
167　rotation speed of 20 rpm. At last, the microarray was washed using the consecutive solutions (0.2%
168　SDS, 2× SSC and 0.2× SSC).

169　*Agilent human miRNA Array V21.0*

170　The experiments were carried out in accordance with the protocol of the manufacturer. In short, the
171　miRNA labeling reagent was used to label the miRNAs. pCp-Cy3 was used to dephosphorylate and
172　ligate 200 ng total RNA. Next, we purify and hybridize the labeled RNA. Last, the Agilent
173　microarray scanner was used to scan the images and Agilent feature extraction software version
174　10.10 was used to grid and analyze them.

175　*Agilent human circRNA Array v2.0*

176　Total RNAs were first digested by using Ribonuclease R to ensure the specific and efficient labeling
177　of cirRNA. Then, Cy3-dCTP was used to label cDNA in accordance with Eberwine's linear RNA
178　amplification method. The succeeding enzymatic reaction was as described above. The next steps of
179　RNA amplification, labeling, and hybridization are described in the "Agilent human lncRNA+mRNA
180　Array v4.0" section.

181　**Microarray imaging and sample analysis**

182　The Agilent GeneSpring software V13.0 was used to summary, normalize and control in quality the
183　data of the lncRNA+mRNA array, miRNA array, and circRNA array.

184　　The genes which were differentially expressed were selected by using the Threshold values of ≥ 2
185　and ≤ -2-fold change and a corrected P value of Benjamini–Hochberg, less than 0.05, which had
186　significant differences. Log2 was used to process the data and the CLUSTER 3.0 software was used
187　to adjust the data. The hierarchical clustering was used to analyze the data and cluster heat maps were
188　generated. At last, the Java TreeView (Stanford University School of Medicine, Stanford, CA, USA)
189　was used to conduct the tree visualization. On the basis of the normalized data, the Pearson
190　correlation coefficient matrix of the correlation between samples was made. The box plot was drawn,
191　showing the gene expression of different samples before and after data normalization. Principal
192　component analysis (PCA) was used to reflect the similarity of samples, and the expression of the
193　samples was shown in the 3D space through the dimension reduction of data.

194　**Comparative analysis of case and control groups**

195　T-test was used to analyze the difference between groups. The default screening criteria for
196　significant difference in this study were as follows: when the gene was upregulated, the number of
197　detected samples required to account for more than 60% of the total number of this group. When the
198　gene was downregulated, the requirement was the same. The difference multiple FC(ABS)≥ 2.0 and
199　$P \leq 0.05$ were also required. The larger the difference between the multiple of FC (ABS), the greater
200　the difference between the two groups. The lower the *P* value, the higher the reliability of differential
201　genes. Cluster analysis was performed, and a scatter map and volcano map were drawn to illustrate
202　the differences between groups.

203 ***Construction of the coding–noncoding gene co-expression (CNC) network***

204 *Construction of the lncRNA–mRNA network*

205 The correlation analysis between the lncRNA and mRNA differentially expressed was used for
206 constructing the network of CNC. For construction of the network, , Pearson correlation was
207 calculated, and the critical correlation lncRNA and mRNA pairs were chosen. When Pearson
208 correlation coefficients, which are not less than 0.99, the lncRNAs and mRNAs pairswere selected to
209 draw the network by using software Cytoscape. In the network construction analysis, a degree was
210 the simplest and most important measure of a gene centrality which was defined as the number of
211 connections between one node **to** the other(BARABASI AND OLTVAI 2004).

212 The lncRNA prediction for Cis-acting and trans-acting was performed in this study. The
213 prediction of Cis-acting lncRNA was conducted by their correlation to the expressed protein-coding
214 genes. When the Pearson's correlation coefficient less than 0.99, Cis-acting lncRNA could be
215 selected. The lncRNA resides at the genomic loci where a protein-coding gene and an lncRNA gene
216 are within 10 kb of each other along the genome(JIA *et al.* 2010).

217 Trans-prediction was performed by using blat tools (http://hgdownload.cse.ucsc.edu/admin/exe/)
218 with the setting of the default parameter to compare the lncRNA with its co-expression mRNAs' 3′
219 UTR.,

220 *Construction of the miRNA–mRNA network*

221 miRNA target gene prediction was analyzed by using miRWalk2.0, which gave the prediction results
222 of 12 miRNA target gene prediction programs. The programs included miRWalk, DIANA-
223 microTv4.0, miRanda-rel2010, mirBridge, miRDB4.0, miRmap, miRNAMap, PicTar2, PITA,
224 RNA22v2, RNAhybrid2.1, and Targetscan 6.2.

225 *Construction of the circRNA–miRNA network*

226 CircRNAs play crucial parts in the function and transcriptional control of miRNA through acting as
227 ceRNAs or positive regulators on their parent coding genes. The network of circRNA–miRNA was
228 performed based on the miRanda-3.3 software. The open-source bioinformatics software Cytoscape
229 was used for constructing the network.

230 ***Gene annotation and enrichment analysis***

231 Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment,
232 and disease analyses were conducted for the target gene mRNA of significantly different lncRNAs,
233 miRNAs, and circRNAs. In the KEGG pathway enrichment and disease analyses, the first 20
234 critically enriched GO terms were selected, and the first 20 significantly enriched terms from
235 biological process (BP), molecular function (MF), and cellular component (CC) were selected in GO.
236 The histogram drawn on the basis of the P value could directly reflect significantly enriched terms. A
237 bubble chart was also provided to show the enrichment analysis results. Finally, candidate signal
238 pathways related to TOF were used to further determine candidate lncRNAs, mRNAs, miRNAs, and
239 circRNAs.

240 ***Quantitative real-time polymerase chain reaction (qRT-PCR)***

241 The RNA was extracted by using the TRIzol method, which was reverse, transcribed to cDNA by
242 using a TaKaRa reverse transcription kit. Reaction amplification was performed by using ABI
243 StepOne Real-Time PCR system (ABI, Vernon, USA). The reaction conditions were as follows: 50
244 °C, 2 min; 95 °C, 10 min; 95 °C, 20 s; 60 °C, 30 s; and 72 °C, 30 s for a total of 40 cycles; 95°C, 15
245 s, 60°C, 1 min, 95°C, 15 s.

246

Data Availability Statement

248 The datasets analyzed for this study can be found in a shared Microsoft OneDrive network disk
249 (https://1drv.ms/f/s!Akmkg4ZxWXm5jbJis0B97A2RBPPr-g)

250

251 **Results**

252 *lncRNA/mRNA*

253 According to the screening criteria mentioned previously (FC≥Ccordin0.05), altogether of 3228
254 lncRNAs which are differentially expressed and 4295 mRNAs which are differentially expressed
255 were found. Among them, 1683 were upregulated lncRNAs, and 1545 were downregulated lncRNAs;
256 by contrast, 2817 mRNAs were upregulated, and 1478 mRNAs were downregulated (Attached Lists
257 1 and 2). The clustering analysis result (Fig.1) suggests that in-group samples have a good correlation
258 and high consistency, whereas the between-group samples are significantly different and can be
259 clearly distinguished. Thus, the selection feasibility of tissue samples is high. Co-expression analysis
260 was conducted on the selected lncRNAs that are differentially expressed and mRNAs. A co-
261 expression network map was constructed (Fig. 2), and 44 lncRNAs and 497 mRNAs co-expressed
262 were identified. The target genes of the co-expressed lncRNA and mRNA were predicted, and the co-
263 expression network map was constructed (Fig. 3). Finally, 66 lncRNAs and 62 mRNAs that may
264 target each other were identified.

265 GO, KEGG pathway enrichment, and disease analyses were conducted for the candidate genes
266 that target each other. In GO analysis, 20 significantly enriched terms were selected from BP, MF,
267 and CC. A histogram was drawn on the basis of the P value, as shown in Fig. 4. In KEGG pathway
268 enrichment and disease analyses, a histogram was drawn for the top 20 terms, as shown in Fig. 5. A
269 total of five lncRNAs and mRNAs that interacted with one another were selected finally (Table 1).

270 *miRNA*

271 A total of 118 differentially expressed miRNAs, including 94 upregulated and 26 downregulated
272 miRNAs, were screened (Attached List 3). A Circos plot was drawn on the basis of the differences
273 (Fig. 6). This plot shows the degree of difference among genes on the basis of their location. Figure 6
274 shows 37 upregulated miRNAs and 10 downregulated miRNAs with a high degree of difference. The
275 number 13, 15, and Y chromosomes are not distributed. Clustering analysis and principal component
276 analysis suggest that in-group samples have a good correlation and a high consistency. The between-
277 group samples are significantly different and can be clearly grouped, with strong data reliability. GO,
278 KEGG pathway enrichment, and disease analyses were conducted for different genes (Figs. 7 and 8).
279 Target gene prediction was conducted for differential miRNA. miRNAs that interacted with another
280 and conformed to the expression trend of the target mRNA were selected (Table 2).

7

281 *circRNA*

282 A total of 18,016 circRNAs which are differentially expressed, including 9404 upregulated circRNAs
283 and 8612 downregulated circRNAs, were screened. On the basis of differences, a Circos plot was
284 drawn (Fig. 9). This plot is based on the location of genes. It can show the degree of difference
285 between differential genes. The plot was distributed in all chromosomes. Clustering analysis and
286 PCA were conducted for cases and control samples. The results show that in-group samples have a
287 good correlation and high consistency; between-group samples are significantly different and can be
288 clearly grouped, and data have strong reliability. GO, KEGG pathway enrichment, and disease
289 analyses were conducted for differential circRNAs (Figs. 10 and 11). The most prominent function of
290 circRNAs is its role as miRNA sponge to regulate target gene expression by inhibiting miRNA
291 activity. A circRNA–miRNA network was constructed (Fig. 12). The figure shows the interaction
292 network between the first eight circRNAs and multiple target miRNAs. In these eight circRNAs,
293 except for hsa_circ_0056861_CBC1, the expression of the remaining circRNAs was downregulated
294 compared with that in the control group. Three differentially expressed circRNAs, namely, hsa-
295 circRNA301, hsa_circ_0113626, and hsa_circ_0129079, may be related with the target mRNA and
296 were selected. The target mRNA of hsa-circRNA301 and hsa_circ_0113626 is CPT2, and that of
297 hsa_circ_0129079 is GHR.

298 **qRT-PCR**

299 qRT-PCR was conducted for the five pairs of candidate lncRNA/mRNAs in the case and control
300 groups. The test result shows that the relative quantitative analysis results of four lncRNAs and four
301 mRNAs are consistent with the chip results (Fig. 13). The four differential candidate lncRNAs are
302 ASO1873, ENST00000452466.1, NR_073058.1, and uc002qim.1, except for ENST00000517747.1.
303 The four differential mRNAs include QRSL1 (NM_018292), CPT2 (NM_000098), GHR
304 (NM_000163), and ZAK (NM_133646), except for MRPS18C (NM_016067).

305

306 **Discussion**

307 TOF is the most common cyanotic congenital heart disease, accounting for approximately 0.2‰ of
308 live births. Its pathological features include pulmonary artery stenosis, aortic straddle, ventricular
309 septal defect, and right ventricular hypertrophy. The prognosis of the disease is closely related to the
310 cause of the disease (whether there is genetic change). Some TOF cases are accompanied by the
311 change in detectable genetic materials, including common genetic syndrome and changes in CNVs,
312 SNPs, and Small InDels (LIU *et al.* 2016; LIU *et al.* 2017; WANG *et al.* 2017). The common genetic
313 syndrome includes trisomy 21 syndrome, trisomy 18 syndrome, DiGeorge syndrome (more than 80%
314 of cases have 22q11 microdeletion), and jaw–heart–face syndrome(HOU *et al.* 2019; SOUTO FILHO *et*
315 *al.* 2019). Patients with TOF accompanied by genetic material changes have different degrees of
316 growth retardation, mental retardation, immune dysfunction, and malformation. After the follow-up,
317 the prognosis of TOF patients with gene mutation is poorer than that of patients with TOF alone. At
318 present, chromosome abnormalities, gene mutation, and other genetic mechanisms have been widely
319 used in the diagnosis and treatment of TOF. For most patients with TOF, the condition involves
320 changes in the heart and is not accompanied by detectable chromosomal abnormalities.

321 Cardiac development is a process of polygenic regulation. The genetic and epigenetic
322 mechanisms such as nucleic acid modification and ncRNA also play an important role in its

8

323 occurrence(SERRA-JUHE *et al.* 2015; GRUNERT *et al.* 2016; MUNTEAN *et al.* 2017). However, the role
324 of genetic and epigenetic modifications in congenital heart disease should be studied in depth. The
325 role of ncRNA in heart development has attracted increasing attention. Evidence shows that ncRNA
326 can take part in the regulation of heart development-related pathways and lead to the occurrence of
327 congenital heart disease(WANG *et al.* 2016; TURTON *et al.* 2019; WU *et al.* 2019). The studies on
328 miRNA in the heart were carried out earlier and there were more studies on it. miRNA-1275,
329 miRNA-27b, miRNA-421, miRNA-1201, and miRNA-122 may be associated with TOF; miRNA-
330 27b may inhibit Mef2c, which affects myocardium development; miRNA-145 may regulate cell
331 apoptosis and mitochondrial function by affecting the expression of FXN gene, thus leading to the
332 occurrence of congenital heart disease. Studies on LncRNA have proved that lncRNAs exist in all
333 stages of embryonic heart development. A very small amount of lncRNAs may be directly related to
334 congenital heart disease. However, whole genome-based studies on lncRNAs, especially on intron or
335 intergenic lncRNAs, are lacking. Given the diversity in the categories and functions of lncRNAs, the
336 reference significance of study results on different lncRNAs is not high. In addition, given the lack of
337 conservation among species, the expression level is generally low. There is no uniform name for
338 lncRNA characteristics in the world, so there is difficulty in understanding its real meaning and
339 function from the name. Given the absence of an lncRNA database, studying lncRNAs is difficult.
340 Moreover, the role of lncRNAs in heart development should be elucidated. With progress in this
341 field, many circRNAs have been identified to be distributed in many organisms; an evolutionarily
342 conserved endogenous ncRNA was found to play a crucial part in the regulation of gene expression
343 and participate in the occurrence and development of many diseases and embryonic
344 development(ZHAO *et al.* 2016; LIU *et al.* 2018). Studies on the role of circRNAs in the regulation of
345 congenital heart disease are few. Thus, the role of circRNAs in congenital heart disease should be
346 investigated. Thus far, the whole genome transcription of TOF has not been reported yet. In the
347 present work, RNA was detected in the RVOT among patients with TOF. The differential lncRNA,
348 miRNA, circRNA, and mRNA profiles of TOF are first reported comprehensively to have a basis for
349 further exploring the etiology and individualized treatment of TOF.

350 In this work, 3228 lncRNAs which are differentially expressed, 4295 mRNAs which are
351 differentially expressed, 118 miRNAs which are differentially expressed, and 18,016 circRNAs
352 which are differentially expressed were found. Clustering analysis and principal component analysis
353 were used to evaluate the reliability of data. The results show that in-group samples have a good
354 correlation and consistency; between-group samples are significantly different and can be clearly
355 grouped with strong data reliability. Next, target gene prediction, GO, and KEGG pathway
356 enrichment analyses were conducted for differential locus. Five pairs of lncRNA–mRNAs were
357 selected and verified via qPCR technology. Finally, four pairs of lncRNA–mRNA, namely,
358 ASO1873-QRSL1, ENST00000452466.1-CPT2, NR_073058.1–GHR, and uc002qim.1-ZAK, were
359 identified. Given the diversity in the categories and functions of lncRNA, the reference significance
360 among different lncRNA study results is low. Candidate lncRNAs were identified on the basis of
361 mRNA. The related literature shows that ZAK in four candidate genes is most likely related to TOF.
362 ZAK is a mitogen-activated protein kinase kinase kinase (MAPKKK) which activates the stress-
363 activated protein kinase/c-jun N-terminal kinase pathway and NF-κB. Twenty-seven different tissue
364 samples from 95 individuals were tested via RNA-seq. The expression of ZAK in the heart is the
365 highest. This gene, as a member of the MAPKKK family of signal transduction molecules, encodes a
366 protein with an N-terminal kinase catalytic domain, which are followed by a leucine zipper motif and
367 a sterile-alpha motif. This magnesium-binding protein, which forms homodimers, is located in the
368 cytoplasm. The protein mediates gamma radiation signaling leading to cell cycle arrest, and the
369 activity of which plays a role in cell cycle checkpoint regulation in cells. The protein also has pro-
370 apoptotic activity(JANDHYALA *et al.* 2008). ZAK can directly or indirectly affect the growth and

371 apoptosis of cardiomyocytes. In 2004, Huang et al. reported that ZAK's expression of a wild-type
372 form induces the characteristic hypertrophic growth features, which include increased cell size,
373 elevated atrial natriuretic factor (ANF) expression, and increased actin fiber organization(HUANG *et*
374 *al.* 2004a). In 2004, Huang et al. reported that ZAK's expression of a dominant-negative form
375 inhibited the characteristic TGF-β-induced features of cardiac hypertrophy, containing increased cell
376 size, elevated expression of ANF, and increased organization of actin fibers(HUANG *et al.* 2004b). In
377 2016, Fu et al. reported that ZAKα signaling activation is crucial for cardiac hypertrophy, and their
378 findings revealed the inherent regulatory role of ZAKβ in antagonizing ZAKα and subsequently
379 downregulating the cardiac hypertrophy and apoptosis induced by ZAKα (FU *et al.* 2016). After
380 selecting ZAK, lncRNA uc002qim.1 was selected and may affect ZAK expression on the basis of the
381 target gene prediction result. The selected miRNA was consistent and interacted with target mRNA
382 in the expression trend. Thus, hsa-miR-130b-3p, hsa-miR-372-3p, and hsa-miR-758-3p were
383 miRNAs that may get involved in the lncRNA–mRNA interaction. The miRNA's seed sequences
384 should be compared with lncRNA to further understand their interaction relationship. Through
385 comparison of seed sequences, lncRNA uc002qim.1 has three potential binding loci with hsa-miR-
386 130b-3p, two potential binding loci with hsa-miR-372-3p, and one potential binding locus with hsa-
387 miR-758-3p. Therefore, hsa-miR-130b-3p may likely play a regulatory role in lncRNA uc002qim.1-
388 mRNA ZAK interaction. The relationship between hsa-miR-130b-3p and congenital heart disease has
389 not been established before. However, hsa-miR-130b-3p has been widely studied. Some studies have
390 shown that hsa-miR-130b-3p can regulate the differentiation of keratinocytes and play a significant
391 part in the skin disease formation(LI *et al.* 2017). miR-130b-3p also acts a pivotal part in the invasion
392 and differentiation of tumor cells. It can inhibit the breast cancer cells' invasion and migration by
393 targeting Notch ligand Delta like-1. miR-130b-3p can negatively regulate PTEN through PI3K and
394 integrin β1 signaling pathway, thereby playing a part in bladder cancer's occurrence and
395 development(YU *et al.* 2015; SHUI *et al.* 2017). Thus, the interaction among uc002qim.1, ZAK, and
396 miR-130b-3p should be studied.

397 In this work, whole genome chip technology screening found that circRNA has the largest data
398 volume. A total of 18,016 circRNAs, which are differentially expressed, including 9404 upregulated
399 and 8612 downregulated circRNAs, were screened. Target miRNA interaction analysis was
400 conducted for the first eight circRNAs, and the interactive network was constructed for further
401 studies. We could not observe the interaction of circRNA with uc002qim.1, ZAK, and miR-130b-3p.
402 cirRNA may have an effect on heart development through other unknown pathways.

403 In summary, the patterns of expression of lncRNAs, mRNAs, miRNAs, and circRNAs in the
404 RVOT of patients with TOF were studied using Agilent human lncRNA+mRNA Array v4.0, Agilent
405 human miRNA Array V21.0, and Agilent human circRNA Array v2.0. In this paper, the full
406 transcriptome data of RVOT myocardial tissue among patients with TOF were systematically
407 reported for the first time, and differentiation, GO, KEGG pathway enrichment, and target gene
408 prediction analyses were conducted to obtain these data. Finally, the optimal RNA sequencing
409 information was provided. As such, the function of these ncRNAs, their molecular regulatory
410 mechanisms, and their pathological mechanism in the occurrence and development of simple TOF
411 can be analyzed. Given sample size limitations, this study should be further strengthened. The
412 function and regulatory mechanism of ncRNAs should be further studied.

413

414 **Conflict of Interest**

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

417

Author Contributions

419 In this study, HDW and LL conceived the concept of the work and designed the study. CYC, YNL,
420 YYL, collected the data and performed the statistical analyses. TBF and BTP collected the sample.
421 HDW and LL revised and finalized the manuscript. All authors read and approved the final
422 manuscript.

423

430

References

432 Bajpai, R., D. A. Chen, A. Rada-Iglesias, J. Zhang, Y. Xiong *et al.*, 2010 CHD7 cooperates with
433   PBAF to control multipotent neural crest formation. Nature 463**:** 958-962.

434 Barabasi, A. L., and Z. N. Oltvai, 2004 Network biology: understanding the cell's functional
435   organization. Nat Rev Genet 5**:** 101-113.

436 Bedard, E., K. P. McCarthy, K. Dimopoulos, G. Giannakoulas, M. A. Gatzoulis *et al.*, 2009
437   Structural abnormalities of the pulmonary trunk in tetralogy of Fallot and potential clinical
438   implications: a morphological study. J Am Coll Cardiol 54**:** 1883-1890.

439 Fu, C. Y., W. W. Kuo, T. J. Ho, S. Y. Wen, L. C. Lin *et al.*, 2016 ZAKbeta antagonizes and
440   ameliorates the cardiac hypertrophic and apoptotic effects induced by ZAKalpha. Cell
441   Biochem Funct 34**:** 606-612.

442 Gao, J., W. Xu, J. Wang, K. Wang and P. Li, 2017 The Role and Molecular Mechanism of Non-
443   Coding RNAs in Pathological Cardiac Remodeling. Int J Mol Sci 18.

444 Grunert, M., C. Dorn, H. Cui, I. Dunkel, K. Schulz *et al.*, 2016 Comparative DNA methylation and
445   gene expression analysis identifies novel genes for structural congenital heart diseases.
446   Cardiovasc Res 112**:** 464-477.

447 Han, B., J. Chao and H. Yao, 2018 Circular RNA and its mechanisms in disease: From the bench to
448   the clinic. Pharmacol Ther 187**:** 31-44.

11

449 Hou, H. T., H. X. Chen, X. L. Wang, C. Yuan, Q. Yang *et al.*, 2019 Genetic characterisation of
450   22q11.2 variations and prevalence in patients with congenital heart disease. Arch Dis Child.

451 Huang, C. Y., P. J. Chueh, C. T. Tseng, K. Y. Liu, H. Y. Tsai *et al.*, 2004a ZAK re-programs atrial
452   natriuretic factor expression and induces hypertrophic growth in H9c2 cardiomyoblast cells.
453   Biochem Biophys Res Commun 324**:** 973-980.

454 Huang, C. Y., W. W. Kuo, P. J. Chueh, C. T. Tseng, M. Y. Chou *et al.*, 2004b Transforming growth
455   factor-beta induces the expression of ANF and hypertrophic growth in cultured
456   cardiomyoblast cells through ZAK. Biochem Biophys Res Commun 324**:** 424-431.

457 Jandhyala, D. M., A. Ahluwalia, T. Obrig and C. M. Thorpe, 2008 ZAK: a MAP3Kinase that
458   transduces Shiga toxin- and ricin-induced proinflammatory cytokine expression. Cell
459   Microbiol 10**:** 1468-1477.

460 Jariwala, N., and D. Sarkar, 2016 Emerging role of lncRNA in cancer: a potential avenue in
461   molecular medicine. Ann Transl Med 4**:** 286.

462 Jia, H., M. Osak, G. K. Bogu, L. W. Stanton, R. Johnson *et al.*, 2010 Genome-wide computational
463   identification and manual annotation of human long noncoding RNA genes. RNA 16**:** 1478-
464   1487.

465 Latus, H., P. Hachmann, K. Gummel, M. Khalil, C. Yerebakan *et al.*, 2015 Impact of residual right
466   ventricular outflow tract obstruction on biventricular strain and synchrony in patients after
467   repair of tetralogy of Fallot: a cardiac magnetic resonance feature tracking study. Eur J
468   Cardiothorac Surg 48**:** 83-90.

469 Li, C. X., H. G. Li, L. T. Huang, Y. W. Kong, F. Y. Chen *et al.*, 2017 H19 lncRNA regulates
470   keratinocyte differentiation by targeting miR-130b-3p. Cell Death Dis 8**:** e3174.

471 Li, X., Z. Wu, X. Fu and W. Han, 2013 Long Noncoding RNAs: Insights from Biological Features
472   and Functions to Diseases. Med Res Rev 33**:** 517-553.

473 Liu, H., Y. Hu, B. Zhuang, J. Yin, X. H. Chen *et al.*, 2018 Differential Expression of CircRNAs in
474   Embryonic Heart Tissue Associated with Ventricular Septal Defect. Int J Med Sci 15**:** 703-
475   712.

476 Liu, L., H. D. Wang, C. Y. Cui, Y. Y. Qin, T. B. Fan *et al.*, 2017 Whole exome sequencing identifies
477   novel mutation in eight Chinese children with isolated tetralogy of Fallot. Oncotarget 8**:**
478   106976-106988.

479 Liu, L., H. D. Wang, C. Y. Cui, D. Wu, T. Li *et al.*, 2016 Application of array-comparative genomic
480   hybridization in tetralogy of Fallot. Medicine (Baltimore) 95**:** e5552.

481 Massin, M. M., S. G. Malekzadeh-Milani, S. Schifflers, H. Dessy and T. Verbeet, 2011 Long-term
482   electrocardiographic follow-up after repair of tetralogy of Fallot. Ann Noninvasive
483   Electrocardiol 16**:** 336-343.

484 Muntean, I., R. Toganel and T. Benedek, 2017 Genetics of Congenital Heart Disease: Past and
485   Present. Biochem Genet 55**:** 105-123.

486 O'Brien, J. E., Jr., N. Kibiryeva, X. G. Zhou, J. A. Marshall, G. K. Lofland *et al.*, 2012 Noncoding
487   RNA expression in myocardium from infants with tetralogy of Fallot. Circ Cardiovasc Genet
488   5**:** 279-286.

489 Patterson, T. A., E. K. Lobenhofer, S. B. Fulmer-Smentek, P. J. Collins, T. M. Chu *et al.*, 2006
490   Performance comparison of one-color and two-color platforms within the MicroArray Quality
491   Control (MAQC) project. Nat Biotechnol 24**:** 1140-1150.

492 Philippen, L. E., E. Dirkx, P. A. da Costa-Martins and L. J. De Windt, 2015 Non-coding RNA in
493   control of gene regulatory programs in cardiac development and disease. J Mol Cell Cardiol
494   89**:** 51-58.

495 Ponting, C. P., P. L. Oliver and W. Reik, 2009 Evolution and functions of long noncoding RNAs.
496   Cell 136**:** 629-641.

497 Salmena, L., L. Poliseno, Y. Tay, L. Kats and P. P. Pandolfi, 2011 A ceRNA hypothesis: the Rosetta
498   Stone of a hidden RNA language? Cell 146**:** 353-358.

499 Schonrock, N., R. P. Harvey and J. S. Mattick, 2012 Long noncoding RNAs in cardiac development
500   and pathophysiology. Circ Res 111**:** 1349-1362.

501 Serra-Juhe, C., I. Cusco, A. Homs, R. Flores, N. Toran *et al.*, 2015 DNA methylation abnormalities
502   in congenital heart disease. Epigenetics 10**:** 167-177.

503 Shin, Y. R., J. W. Jung, N. K. Kim, J. Y. Choi, Y. J. Kim *et al.*, 2016 Factors associated with
504   progression of right ventricular enlargement and dysfunction after repair of tetralogy of Fallot
505   based on serial cardiac magnetic resonance imaging. Eur J Cardiothorac Surg 50**:** 464-469.

506 Shui, Y., X. Yu, R. Duan, Q. Bao, J. Wu *et al.*, 2017 miR-130b-3p inhibits cell invasion and
507   migration by targeting the Notch ligand Delta-like 1 in breast carcinoma. Gene 609**:** 80-87.

508 Souto Filho, J. T. D., H. A. A. Ribeiro, I. P. B. Fassbender, J. Ribeiro, W. D. S. Ferreira Junior *et al.*,
509   2019 Bernard-Soulier syndrome associated with 22q11.2 deletion and clinical features of
510   DiGeorge/velocardiofacial syndrome. Blood Coagul Fibrinolysis 30**:** 423-425.

511 Tano, K., and N. Akimitsu, 2012 Long non-coding RNAs in cancer progression. Front Genet 3**:** 219.

512 Turton, N., R. Swan, T. Mahenthiralingam, D. Pitts and I. M. Dykes, 2019 The Functions of Long
513   Non-Coding RNA during Embryonic Cardiovascular Development and Its Potential for
514   Diagnosis and Treatment of Congenital Heart Disease. J Cardiovasc Dev Dis 6.

515 Wang, H. D., L. Liu, D. Wu, T. Li, C. Y. Cui *et al.*, 2017 Clinical and molecular cytogenetic analyses
516   of four families with 1q21.1 microdeletion or microduplication. J Gene Med 19.

517 Wang, L., D. Tian, J. Hu, H. Xing, M. Sun *et al.*, 2016 MiRNA-145 Regulates the Development of
518   Congenital Heart Disease Through Targeting FXN. Pediatr Cardiol 37**:** 629-636.

519 Wu, J., J. Li, H. Liu, J. Yin, M. Zhang *et al.*, 2019 Circulating plasma circular RNAs as novel
520   diagnostic biomarkers for congenital heart disease in children. J Clin Lab Anal 33**:** e22998.

521 Yates, L. A., C. J. Norbury and R. J. Gilbert, 2013 The long and short of microRNA. Cell 153**:** 516-
522   519.

523 Yu, C. Y., and H. C. Kuo, 2019 The emerging roles and functions of circular RNAs and their
524   generation. J Biomed Sci 26**:** 29.

525 Yu, T., R. Cao, S. Li, M. Fu, L. Ren *et al.*, 2015 MiR-130b plays an oncogenic role by repressing
526   PTEN expression in esophageal squamous cell carcinoma cells. BMC Cancer 15**:** 29.

527 Zhang, J., J. J. Chang, F. Xu, X. J. Ma, Y. Wu *et al.*, 2013 MicroRNA deregulation in right
528   ventricular outflow tract myocardium in nonsyndromic tetralogy of fallot. Can J Cardiol 29**:**
529   1695-1703.

530   Zhao, Y., P. N. Alexandrov, V. Jaber and W. J. Lukiw, 2016 Deficiency in the Ubiquitin Conjugating
531   Enzyme UBE2A in Alzheimer's Disease (AD) is Linked to Deficits in a Natural Circular
532   miRNA-7 Sponge (circRNA; ciRS-7). Genes (Basel) 7.

533

534

535

536
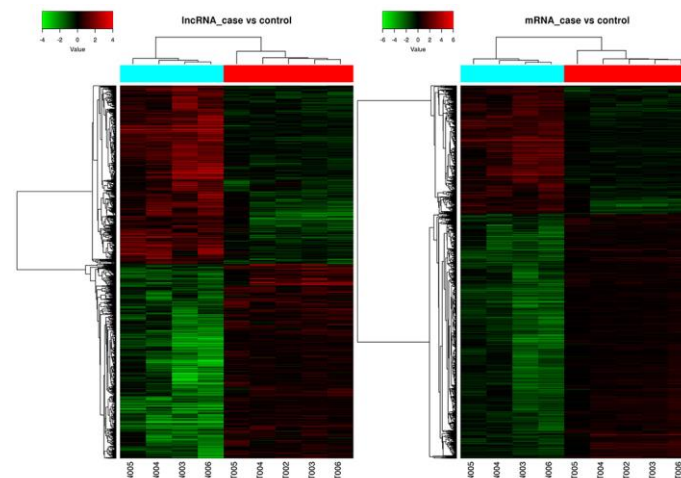
538   **Figure 1**. The sample clustering analysis results of lncRNAs and mRNAs

539

540

14
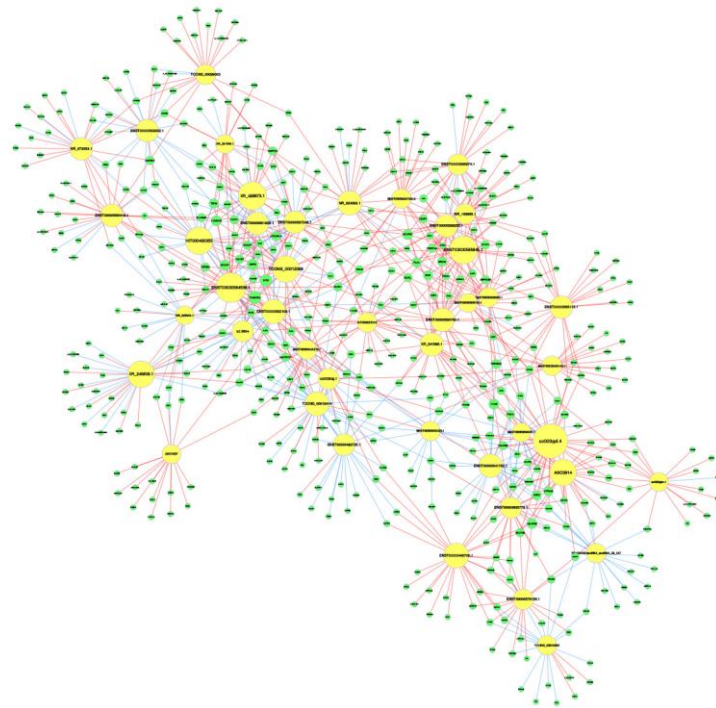
Figure 2. The co-expression network map of lncRNAs and mRNAs
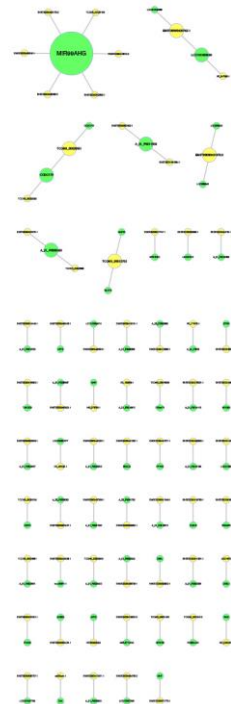
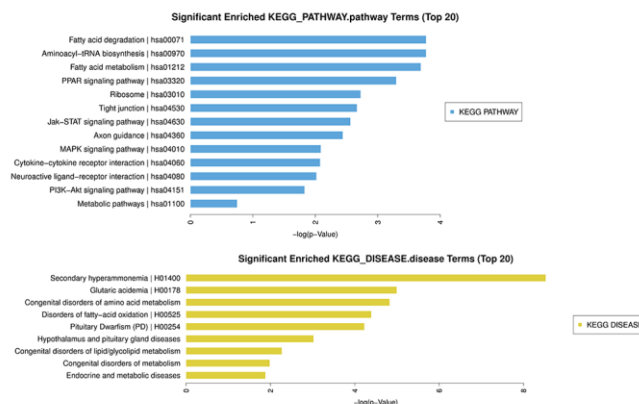**Figure 3**. The target genes of the co-expressed lncRNA and mRNA

548

549

550
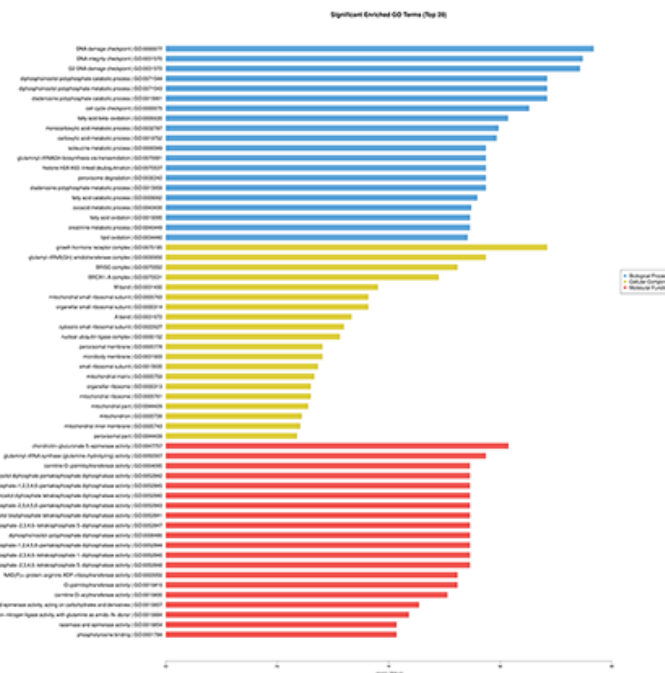
Figure 4. A histogram of 20 significantly enriched terms selected from BP, MF, and CC on the basis of the GO analysis
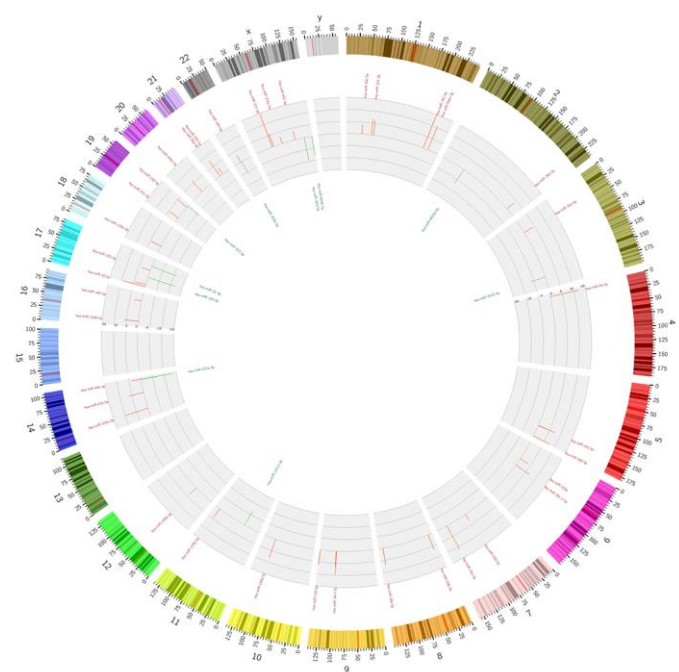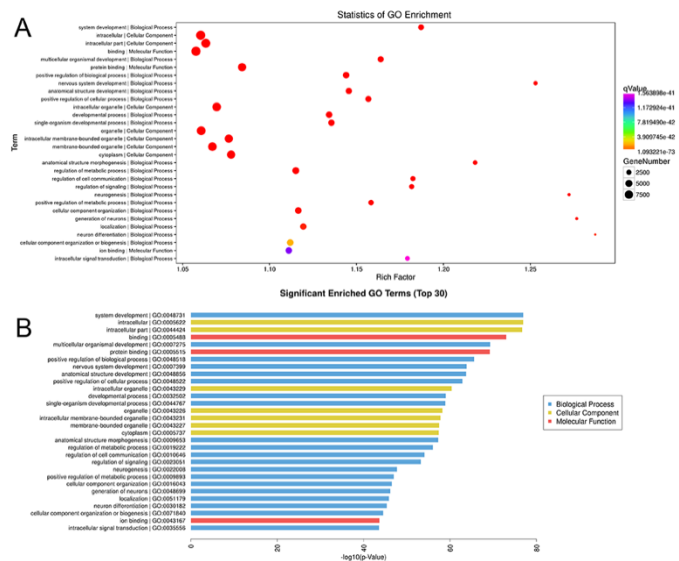
553

554

555

**Figure 5**. A histogram of top 20 terms by KEGG pathway enrichment and disease analyses

557

558

16

Figure 6. A Circos plot shows 37 upregulated miRNAs and 10 downregulated miRNAs with a high degree of difference on the basis of their location

Figure 7. Statistics of GO Enrichment and significant enrichment Go terms of TOP 30 differentially expressed miRNA

17

569

570

571
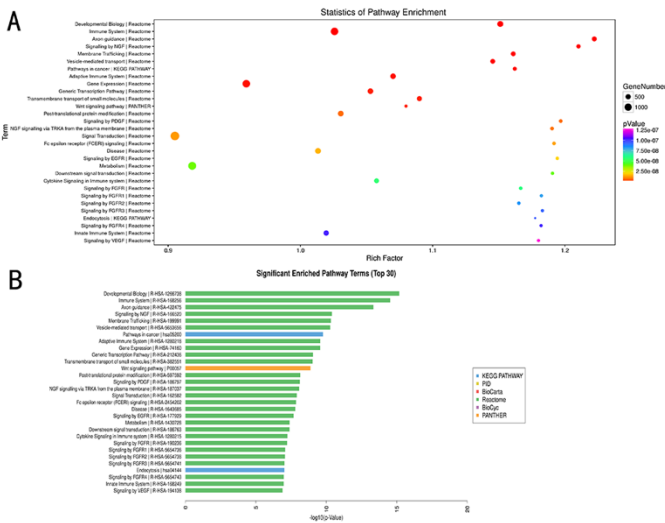
Figure 8. Statistics of KEGG Pathway Enrichment and significant enrichment KEGG Pathway terms of TOP 30 differentially expressed miRNA

574

575

576

577

578

18

579

580 **Figure 9**. A Circos plot shows differentially expressed circRNAs on the basis of their location

581
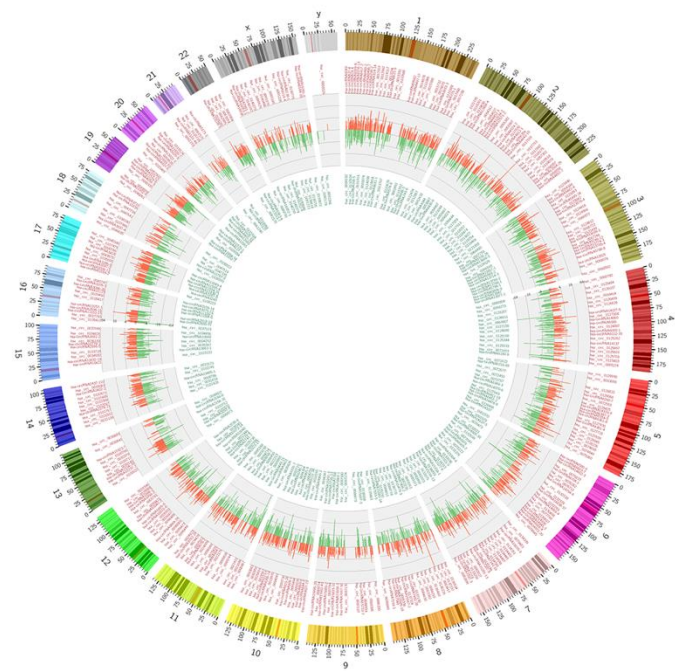
582

583

584 **Figure 10**. Statistics of GO Enrichment and significant enrichment Go terms of TOP 30 differentially
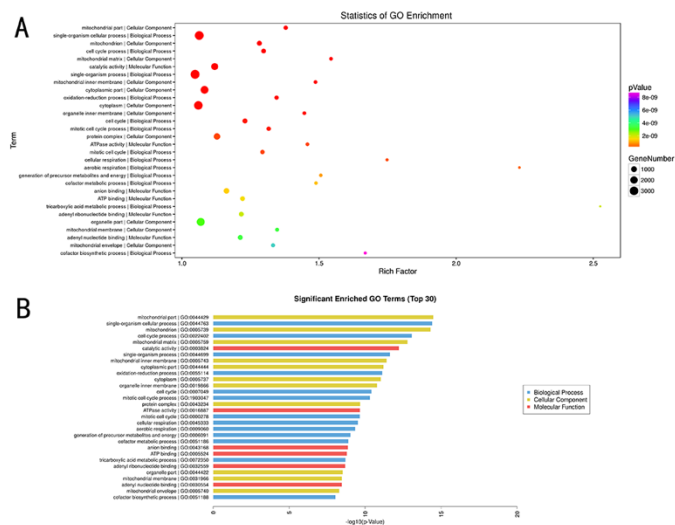585 expressed circRNAs

586

Figure 11. Statistics of KEGG Pathway Enrichment and significant enrichment KEGG Pathway terms of TOP 30 differentially expressed circRNAs

Figure 12. The circRNA–miRNA network shows the interaction network between the first eight circRNAs and multiple target miRNAs

Figure 13. The qRT-PCR results of four lncRNAs and four mRNAs which are consistent with the chip results

20

599

600 **Table 1.** Five selected lncRNAs and their mRNA that may be the target genes.

LncRNA ID	LncRNA Gene name	target mRNA Gene name	mRNA RefSeq Accession	Case *vs* Control（lncRNA）	Case *vs* Control（mRNA）
ASO1873	ASO1873	QRSL1	NM_018292	up	up
ENST00000452466.1	ENSG00000236723.1	CPT2	NM_000098	up	up
NR_073058.1	C7orf55	GHR	NM_000163	up	up
ENST00000517747.1	ENSG00000245281.2	MRPS18C	NM_016067	up	up
uc002qim.1	AL833150	ZAK	NM_133646	up	up

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

21

617   **Table 2.** Seven selected miRNAs and their mRNA that may be the target genes.

Selected miRNA	miRNA Refseq accession	target mRNA Gene name	mRNA Refseq Accession	miRNA case vs control
hsa-miR-130b-3p	MIMAT0000691	QRSL1	NM_018292	down
hsa-miR-372-3p	MIMAT0000724	QRSL1	NM_018292	down
hsa-miR-130b-3p	MIMAT0000691	ZAK	NM_133646	down
hsa-miR-372-3p	MIMAT0000724	ZAK	NM_133646	down
hsa-miR-758-3p	MIMAT0003879	ZAK	NM_133646	down
hsa-miR-432-5p	MIMAT0002814	GHR	NM_001242400	down
hsa-miR-503-5p	MIMAT0002874	GHR	NM_001242400	down

618

619

22