doi: 10.3978/j.issn.2095-6959.2020.11.001
View this article at: http://dx.doi.org/10.3978/j.issn.2095-6959.2020.11.001

·论著·

心肌肥厚模型：内皮素 1 刺激人诱导的多能干细胞源性心肌细胞 lncRNA 表达谱及 ceRNA 网络的生物信息学分析

尤红俊[1]，赵倩倩[2]，常凤军[1]，韩稳琦[1]，程功[1]，寿锡凌[1]，刘富强[1]，刁佳宇[1]

(1. 陕西省人民医院心血管内科，西安 710068；2. 空军军医大学第一附属医院临床免疫科，西安 710032)

[摘　要]　**目的：**对心肌肥厚模型——内皮素1(endothelin 1，ET1)刺激人诱导的多能干细胞源性心肌细胞(human induced pluripotent stem cell derived cardiomyocytes，hiPSC-CMs)的基因芯片数据进行生物信息学分析，探讨长链非编码RNA(long non-coding RNA，lncRNA)的表达谱及竞争性内源性RNA(competing endogenous RNA，ceRNA)调控网络。**方法：**在GEO(Gene Expression Omnibus)数据库中筛选心肌肥厚相关基因芯片，对下载的原始探针矩阵信息和平台注释文件进行预处理，获取基因表达矩阵信息，利用实用报表提取语言(practical extraction and report language，Perl)赋予基因不同的生物型属性，从而区分出lncRNA与mRNA。利用R语言limma包分别分析lncRNAs和mRNAs的基因表达差异。使用在线下载的数据库预测与每个lncRNA关联的microRNA(miRNA)。在此基础上，利用基于TargetScan，miRDB和miRTarBase的预测软件对miRNA的靶基因(mRNA)进行预测。以Cytoscape3.7.1软件构建lncRNA-miRNA-mRNA ceRNA调控网络。针对ceRNA调控网络中的靶基因，运用在线数据库DAVID6.8进行基因本体论(Gene Ontology，GO)富集分析，并通过在线工具KOBAS3.0进行京都基因和基因组百科全书(Kyoto Encyclopedia of Genes and Genomes，KEGG)通路富集分析。**结果：**与对照组相比，ET1刺激的hiPSC-CMs中19个lncRNA和717个mRNA表达差异具有统计学意义。其中，lncRNA表达上调的有4个，下调的有15个；mRNA表达上调的有312个，下调的有405个。成功构建了ceRNA调控网络，该网络中，显著上调的lncRNA包括AC017002和MIR210HG，分别与9，10个miRNA关联密切，显著下调的有LINC00342，其与9个miRNA关联密切。该网络中靶基因共32个，其中18个上调，14个下调。对靶基因进行GO富集分析，发现被显著富集到细胞增殖、质膜和蛋白结合等13个不同的GO子集中。KEGG富集分析发现：靶基因被富集到MAPK信号通路、细胞凋亡、Wnt信号通路、JAK-STAT信号通路、TNF信号通路等20个信号通路上。**结论：**本研究构建了ceRNA网络，为探讨lncRNA参与心肌肥厚致病机制提供了一个新的视角。LncRNA可能通过ceRNA调控网络，影响心肌肥厚发生发展。在ceRNA网络中有3个lncRNA(表达上调的为AC017002、MIR210HG，表达下调的为LINC00342)可能以竞争性“海绵吸附”miRNA的方式调控下游mRNA的表达水平，使AP1，c-MYC，PIM1表达上调，ASK1(即MAP3K5)，MSK1/2(即RPS6KA5)及SFRP1表达下调，并通过在MAPK信号通路、细胞凋亡、Wnt信号通路、JAK-STAT信号通路、TNF信号通路等途经参与心肌肥厚的病理生理过程。这些基因有可能成为潜在的治疗靶点，研发阻断上述lncRNA或下游靶基因的相关药物可能对治疗心肌肥厚有着深远的临床意义。

[关键词]　长链非编码RNA；竞争性内源性RNA；心肌肥厚；基因表达谱；差异表达基因；GEO数据库

收稿日期 (Date of reception)：2019–11–18

通信作者 (Corresponding author)：刁佳宇，Email: diaojiayu2007@163.com

基金项目 (Foundation item)：陕西省青年科技新星项目 (2017KJXX-70)。This work was supported by the Youth Science and Technology Rising Star Project of Shaanxi Province, China (2017KJXX-70).

Bioinformatics analysis of lncRNA expression profiles and competing endogenous RNA network of myocardial hypertrophy model: human induced pluripotent stem cell derived cardiomyocytes stimulated by endothelin 1

YOU Hongjun[1], ZHAO Qianqian[2], CHANG Fengjun[1], HAN Wenqi[1], CHENG Gong[1], SHOU Xiling[1], LIU Fuqiang[1], DIAO Jiayu[1]

(1. Department of Cardiology, Shaanxi Provincial People's Hospital, Xi'an 710068; 2. Department of Clinical Immunology, First Affiliated Hospital of Air Force Military Medical University, Xi'an 710032, China)

Abstract **Objective:** To perform bioinformatics analysis on the gene chip data of myocardial hypertrophy model-human induced pluripotent stem cell derived cardiomyocytes induced by endothelin-1 (hiPSC-CMs), and to explore the expression profile of long non-coding RNAs (lncRNAs) and competing endogenous RNA (ceRNA) regulatory network. **Methods:** The gene chip related to cardiac hypertrophy in the Gene Expression Omnibus (GEO) database was screened, and the downloaded original probe matrix information and platform annotation file were preprocessed to obtain the gene expression matrix information. The different biotype attributes were assigned to the genes by using the Practical Extraction and Report Language (Perl), to distinguish lncRNAs from mRNAs. Gene differential expression analysis of lncRNAs and mRNAs was performed using the R language limma package. The microRNAs associated with each lncRNA were predicted using a database downloaded online. On this basis, target genes (mRNAs) of microRNAs were predicted using prediction softwares based on TargetScan, miRDB and miRTarBase. The lncRNA-miRNA-mRNA ceRNA regulatory network was constructed using Cytoscape 3.7.1 software. The Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of target genes from the ceRNA network were performed by the online database DAVID6.8 and the web servers KOBAS3.0. **Results:** Compared with the control groups, there were 19 lncRNAs and 717 mRNAs expressed differentially in ET1-stimulated hiPSC-CMs with statistical significance, among which 4 lncRNAs up-regulated, 15 lncRNAs down-regulated, 312 mRNAs up-regulated and 405 mRNAs down-regulated. ceRNA regulatory network was constructed successfully, in which lncRNA AC017002 and MIR210HG were up-regulated, LINC00342 was down-regulated. The 3 lncRNAs related closely to 9, 10 and 9 miRNAs, respectively. There were 32 target genes in the network, among which 18 up-regulated and 14 down-regulated. GO enrichment analysis showed that the target genes were significantly enriched to 13 different GO subsets, including cell proliferation, plasma membrane, protein binding and so on. KEGG enrichment analysis found that the target genes were enriched to 20 signaling pathways, including the MAPK signaling pathway, apoptosis, Wnt signaling pathway, JAK-STAT signaling pathway, TNF signaling pathway and so on. **Conclusion:** Our study constructed the ceRNA network, providing a new perspective for exploring the pathogenic mechanism of lncRNAs involved in cardiac hypertrophy. LncRNAs may affect the pathogenesis of cardiac hypertrophy by regulating in the ceRNA network. In the ceRNA network, up-regulated AC017002 and MIR210HG and down-regulated LINC00342 may regulate the expression level of downstream mRNA by means of competitive "sponge adsorption" of miRNA, so that the expression of AP1, c-MYC and PIM1 were up-regulated, and the expression of ASK1 (MAP3K5), MSK1/2 (RPS6KA5) and SFRP1 were down-regulated. These genes were involved in the pathophysiological process of cardiac hypertrophy through MAPK signaling pathway, apoptosis, Wnt signaling pathway, JAK-STST

signaling pathway and TNF signaling pathway. It may become a potential therapeutic target, and the development of drugs that block the above-mentioned lncRNAs and/or downstream target genes may have profound clinical significance in the treatment of cardiac hypertrophy.

Keywords lncRNA; ceRNA; cardiac hypertrophy; gene expression profile; differentially expressed genes; GEO database

心肌肥厚是心力衰竭等心血管疾病的独立危险因素[1]。基因表达调控机制参与心肌肥厚的发生发展。非编码RNA(non-coding RNAs，ncRNAs)是一类不编码蛋白质但调控蛋白质表达的RNA分子，如lncRNA，microRNA (miRNA)、环状RNA (circular RNA，circRNA)和内源性小干扰RNA(small interfere RNA，siRNA)等。作为一种新型的、长度超过200个核苷酸的非编码RNA，lncRNA在剂量补偿效应、细胞周期的调节、细胞命运决定、细胞分化的调节、染色质重塑的调节、RNA加工、维护核结构完整性和表观遗传学的调节中均发挥着重要作用，已成为遗传研究的热点[2-3]。LncRNA在调节信使RNA(mRNA)和各种心血管疾病的发病机制中发挥重要的作用[4-5]，其可能通过与mRNA竞争性“海绵吸附”miRNA，抑制miRNA活性，从而微调miRNA介导的靶基因(mRNA)的表达，进而影响疾病的发生发展。LncRNA参与心肌生长发育的基因调控，参与心肌肥厚及心力衰竭的发病过程。研究lncRNA在基因调控网络中的作用将提高对心肌肥厚甚至心力衰竭的诊治理念，并揭示疾病初始失调及其潜在机制。然而，全面研究心肌肥厚基因调控水平却受到人类心肌细胞可用性的限制。既往研究[6]发现：内皮素1(endothelin 1，ET-1)刺激的人诱导多能干细胞源性心肌细胞(human induced pluripotent stem cell-derived cardiomyocytes，hiPSC-CMs)可作为人类心肌肥厚的研究模型。随着生物信息技术的发展，基因组学、转录组学、蛋白质组学等大数据已经产生，生物信息技术与计算机科学相结合进行分析，可研究多个生物分子之间的关系。生物信息学分析已成为肿瘤学、免疫性疾病、炎症性疾病、心血管疾病等多学科领域的有效研究工具，受到很多数科研人员的青睐。本研究从GEO(https://www.ncbi.nlm.nih.gov/geo/)数据库中筛选出ET-1刺激的hiPSC-CMs的RNA表达谱芯片数据，通过生物信息学方法，分析肥厚心肌差异表达基因(differentially expressed genes，DEGs)，并构建lncRNA-miRNA-mRNA ceRNA调控网络，对了解心肌肥厚的病理生理过程具有重要的参考价值，为寻找新的治疗靶点提供实验理论依据。

1 资料与方法

1.1 数据下载

以“cardiac hypertrophy”为搜索词，在美国国立生物中心的GEO数据库中搜索已公布的心肌肥厚相关基因芯片，获取一个数据集GSE60291。GSE60291来源于美国威斯康星州医学院。采用Affymetrix Human Genome U133 Plus 2.0 Array平台，以hiPSC-CMs为研究对象，ET1刺激的hiPSC-CMs为实验组(3组)，无ET1刺激的hiPSC-CMs为对照组(3组)。

1.2 数据处理

对下载的原始探针矩阵信息和平台注释文件进行预处理，获取基因表达矩阵信息。实用报表提取语言(Perl)被用来赋予基因属性，以区分出lncRNA和mRNA。Perl是一种高级、通用、动态的编程语言。

1.3 LncRNA 和 mRNA 差异表达分析

采用统计软件R 3.5.0和R语言limma包分析基因表达差异，该软件为基于Bioconductor的专门用于处理表达谱芯片数据。Limma软件包中的经验贝叶斯检验用于检测ET1刺激的hiPSC-CMs与对照组之间的DEGs。DEGs的筛选条件为差异倍数log fold change>1或<−1，P<0.05。并运用R语言绘制热图，以可视化DEGs。

1.4 预测 miRNA 和靶基因

使用预先下载的数据库和Perl预测每个lncRNA相关的miRNA，然后利用TargetScan，miRDB和miRTarBase等预测软件预测miRNA的靶基因(mRNA)；利用Cytoscape3.7.1软件构建lncRNA-microRNA-mRNA ceRNA可视化调控网络。

1.5 GO 和 KEGG 富集分析

对ceRNA网络中靶基因实施基因本体论

(Gene Ontology，GO)和京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes，KEGG)通路富集分析。利用在线数据库DAVID 6.8(https://david.ncifcrf.gov/)对基因进行GO富集分析，在生物学过程(biological processes，BP)、细胞组分(cellular component，CC)和分子功能(molecular function，MF)方面对基因进行功能注释。利用在线工具KOBAS(http://kobas.cbi.pku.edu.cn/anno_iden.php)进行KEGG通路富集分析，探讨DEGs显著富集的信号通路。P<0.05为差异有统计学意义。

2 结果

2.1 基因差异表达分析

与对照组相比，ET1刺激的hiPSC-CMs中19个lncRNA和717个mRNA差异表达具有统计学意义。其中，lncRNA表达上调的有4个，下调的有15个；mRNA表达上调的有312个，下调的有405个。差异表达的lncRNA和mRNA的数目见表1。LncRNA和mRNA差异表达分析如火山图(图1，表2，3)。对差异表达的lncRNA及部分mRNA绘制热图见图2，3。

表1 差异表达基因数目(logFC>1或<−1)

Table 1 Number of DEGs (logFC>1或<−1)

Regulation	LncRNA expression	mRNA expression
Up-regulation	4	312
Down-regulation	15	405
Total Sum	19	717

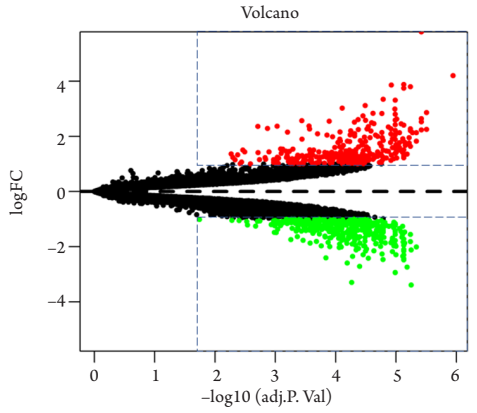

图1 基因差异表达谱火山图(虚线框内所示为差异表达基因)

Figure 1 Gene differential expression profile volcano map (DEGs were showed in dotted box)

表2 LncRNA差异表达分析

Table 2 LncRNA differential expression analysis

LncRNA	LogFC	AveExpr	*t*	*P*	*B*
ZBED5-AS1	−1.666	6.512	−16.254	1.58E-05	8.514
HAND2-AS1	−1.212	8.085	−16.128	1.64E-05	8.450
C1QTNF1-AS1	−1.155	5.561	−14.693	2.33E-05	7.675
ASB16-AS1	−1.054	4.455	−14.220	2.77E-05	7.402
RARA-AS1	1.086	7.538	12.024	5.31E-05	5.991
AC017002	1.469	4.775	11.397	6.76E-05	5.541
GATA6-AS1	−1.221	9.675	−10.822	8.70E-05	5.106
LINC01351	−1.507	6.212	−10.680	9.29E-05	4.995
LINC00648	−2.406	5.344	−9.767	<0.001	4.247
A2M-AS1	−1.382	5.087	−8.897	<0.001	3.472
PXN-AS1	−1.420	5.627	−8.889	<0.001	3.465
PAXIP1-AS1	−1.124	5.769	−8.689	<0.001	3.277
FGF13-AS1	−1.196	5.250	−8.652	<0.001	3.242
LINC00342	−1.060	5.733	−8.445	<0.001	3.043
LINC00702	2.565	8.613	8.178	<0.001	2.780
CAHM	−1.001	4.713	−7.587	0.001	2.173
UNC5B-AS1	−1.038	9.967	−7.389	0.001	1.961
MIR210HG	1.406	6.101	5.949	0.002	0.279
LOH12CR2	−1.212	4.403	−5.493	0.003	-0.311

“logFC”是两组表达值间以2为底对数化的变化倍数，“AveExpr”为平均表达量，“B”为贝叶斯检验后得到的标准差的对数化值。

“LogFC” is the change multiple of logarithmic base 2 between the expression values of two groups, “AveExpr” is the average expression, “B” is the logarithmic value of the standard deviation obtained after bayes test.

2.2 靶基因预测结果

预测与每个lncRNA关联的miRNA，用于预测靶基因，部分展示见表4。

2.3 LncRNA–miRNA-mRNA 网络构建

基于2.2的结果构建lncRNA-miRNA-mRNA ceRNA网络(图4)。该图中菱形、三角形、椭圆形分别代表lncRNA，miRNA，mRNA。上调的AC017002与9个miRNA关联密切，包括hsa-miR-

27a-3p，hsa-miR-425-5p，hsa-miR-135a-5p，hsa-miR-142-3p，hsa-miR-761，hsa-miR-3619-5p，hsa-miR-23b-3p，hsa-miR-338-3p，hsa-miR-449c-5p；上调的MIR210HG与10个miRNA关联密切，包括hsa-miR-135a-5p，hsa-miR-206，hsa-miR-761，hsa-miR-3619-5p，hsa-miR-216b-5p，hsa-miR-24-3p，hsa-miR-27a-3p，hsa-miR-449c-5p，hsa-miR-125b-5p，hsa-miR-129-5p；下调的LINC00342与9个miRNA关联密切，包括hsa-miR-139-5p，hsa-miR-142-3p，hsa-miR-193a-3p，hsa-miR-761，hsa-miR-3619-5p，hsa-miR-217，hsa-miR-27a-3p，hsa-miR-338-3p，hsa-miR-129-5p(图4)。该网络中靶基因共32个，其中18个上调，包括*TPM3*，*PIM1*，*AMOTL2*，*INPP5F*，*FSCN1*，*KANK4*，*LITAF*，*SLC7A1*，*KCNJ2*，*GBA2*，*ENPEP*，*BDNF*，*TNFRSF21*，*FOS*，*LIMS1*，*SPRY4*，*MYC*，*PLAU*；14个下调，包括*SEMA6D*，*ARL15*，*RPS6KA5*，*NR3C2*，*SFRP1*，*TGFBR3*，*BMF*，*EZH2*，*MAP3K5*，*SLC16A1*，*MXI1*，*GPAM*，*ZWINT*，*BHLHB9*。

表3 mRNA(前20)差异表达分析

Table 3 mRNA differential expression analysis (top 20)

mRNA	logFC	AveExpr	*t*	*P*	*B*
SERPINE1	4.200	7.151	39.741	1.13E-06	15.005
FJX1	2.255	6.458	31.737	3.13E-06	13.607
MAP2K1	2.862	10.059	30.913	3.13E-06	13.431
CCL4	5.791	6.929	28.842	3.80E-06	12.954
SPRY4	2.232	6.845	27.881	3.80E-06	12.714
LPL	2.639	7.797	27.794	3.80E-06	12.692
ENPEP	2.527	4.204	27.296	3.80E-06	12.563
ALDH1B1	1.827	6.659	25.730	4.60E-06	12.133
EML1	−2.010	6.723	−25.679	4.60E-06	12.119
NANOS1	2.134	9.218	25.568	4.60E-06	12.087
RDH10	2.132	7.505	24.837	5.33E-06	11.872
DRD1	−3.391	5.712	−24.197	5.63E-06	11.676
HELLS	−2.120	4.735	−24.187	5.63E-06	11.673
UBE2J1	1.634	7.302	23.791	5.76E-06	11.548
SOX9	3.802	7.528	23.641	5.76E-06	11.500
IRS2	2.275	8.519	23.354	5.76E-06	11.408
GINS1	−1.558	8.633	−23.080	5.76E-06	11.317
PDYN	2.209	5.625	23.071	5.76E-06	11.314
NUDT6	−1.816	6.512	−23.049	5.76E-06	11.307
TRIM55	1.754	10.989	22.704	6.08E-06	11.192

“logFC”为两组表达值间以2为底对数化的变化倍数，“AveExpr”为平均表达量，“B”为贝叶斯检验后得到的标准差的对数化值。

“LogFC” is the change multiple of logarithmic base 2 between the expression values of two groups, “AveExpr” is the average expression, “B” is the logarithmic value of the standard deviation obtained after bayes test.

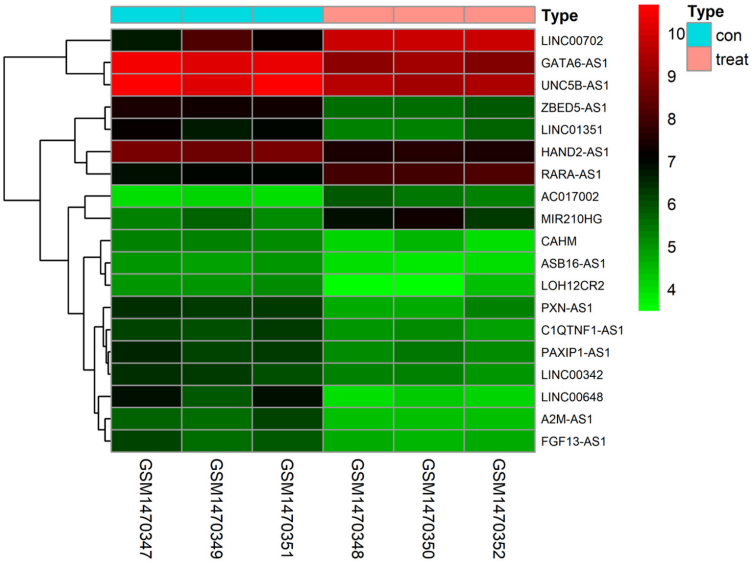

图2 LncRNA表达谱热图(con：对照组；treat：实验组)

Figure 2 LncRNA expression profile heatmap (con: hiPSC-CMs stimulated without ET1; treat: hiPSC-CMs stimulated with ET1)

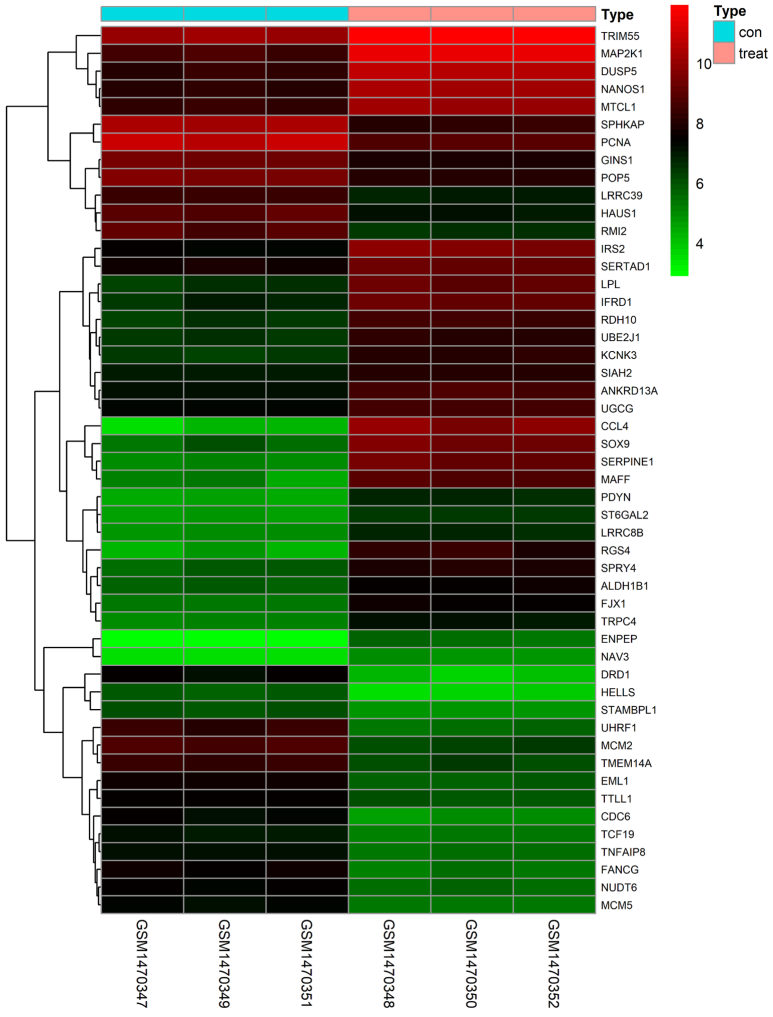

图3 mRNA(前50)表达谱热图(con：对照组；treat：实验组)

Figure 3 mRNA expression profile heatmap (con: hiPSC-CMs stimulated without ET1; treat: hiPSC-CMs stimulated with ET1)

表4 LncRNA关联的miRNA部分展示及其下游靶基因(mRNA，前20)预测

Table 4 Prediction of miRNAs and downstream target genes (mRNA) associated with lncRNAs (Top 20)

miRNA	Target gene	miRDB	miRTarBase	TargetScan	Sum
hsa-miR-216b-5p	TPM3	1	1	1	3
hsa-miR-216b-5p	*MCM4*	1	1	1	3
hsa-miR-761	*PIM1*	1	1	1	3
hsa-miR-761	*AMER1*	1	1	1	3
hsa-miR-24-3p	*AMOTL2*	1	1	1	3
hsa-miR-129-5p	*EI24*	1	1	1	3
hsa-miR-27a-3p	*NPEPPS*	1	1	1	3
hsa-miR-125b-5p	*CSNK2A1*	1	1	1	3
hsa-miR-142-3p	*CLTA*	1	1	1	3
hsa-miR-23b-3p	*ZMYM2*	1	1	1	3
hsa-miR-27a-3p	*NR2F2*	1	1	1	3
hsa-miR-425-5p	*LCOR*	1	1	1	3
hsa-miR-142-3p	*RICTOR*	1	1	1	3
hsa-miR-142-3p	*AKT1S1*	1	1	1	3
hsa-miR-129-5p	*SORBS2*	1	1	1	3
hsa-miR-10a-5p	*MTF2*	1	1	1	3
hsa-miR-142-3p	*KAT7*	1	1	1	3
hsa-miR-10a-5p	*MED1*	1	1	1	3
hsa-miR-129-5p	*DUSP10*	1	1	1	3
hsa-miR-125b-5p	*ZNF385A*	1	1	1	3

“miRNA”为lncRNA关联的microRNA，“Target gene”为相应miRNA的靶基因，其预测基于TargetScan，miRDB和miRTarBase三个数据库，“Sum”代表总得分，若得分为3表示同时在3个数据库中均被证明为相应的miRNA的靶基因。

“miRNA” is the microRNA associated with lncRNA. “Target gene” is the target gene of the corresponding miRNA, and its prediction is based on 3 databases of TargetScan, miRDB and miRTarBase. “Sum” represents the total score. If the score is 3, it means that the mRNA is the target gene of corresponding miRNA proved in the three databases simultaneously.

2.4 GO 富集分析

对ceRNA网络中的靶基因进行GO富集分析，发现靶基因被富集到13个不同的GO子集中，包括BP，CC和MF 3个方面，其中各自最显著富集的子集分别为细胞增殖、质膜和蛋白结合(表5)，差异具有统计学意义(P<0.05)。采用R语言ggplot2包进行可视化呈现(图5)。

2.5 KEGG 富集分析

对ceRNA网络中的靶基因进行KEGG通路富集分析，发现靶基因被富集到20个信号通路上(表6)，差异具有统计学意义(P<0.05)。对符合要求的以上KEGG通路，采用R语言进行部分可视化呈现(图6)。这些显著富集的通路包括：MAPK信号通路(图7)、细胞凋亡(图8)、Wnt信号通路(图9)、JAK-STAT信号通路(图10)、TNF信号通路(图11)、心肌细胞中肾上腺素能信号、癌症中的miRNAs、神经营养因子信号通路、甲状腺癌、膀胱癌、急性髓系白血病、结直肠癌、昼夜节律、胆碱能突触、乙型肝炎、缩宫素信号通路、癌症中转录错误调节、cAMP信号通路、癌症中蛋白聚糖等。

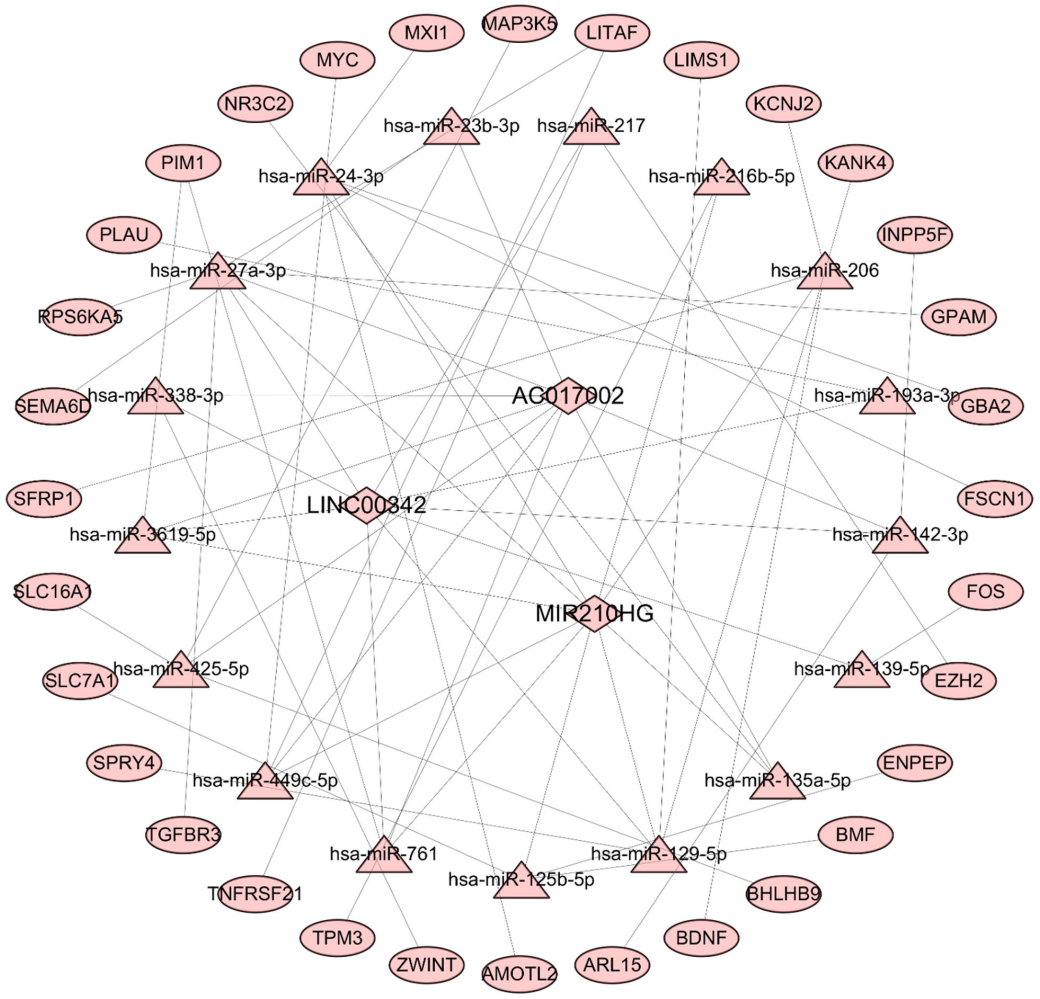

图4 LncRNA-miRNA-mRNA ceRNA网络构建

Figure 4 lncRNA-miRNA-mRNA ceRNA network construction

菱形：lncRNA；三角形：miRNA；椭圆：mRNA

Diamond: lncRNA; triangle: miRNA; ellipse: mRNA

表5 LncRNA-miRNA-mRNA ceRNA网络中靶基因GO富集分析

Table 5 GO enrichment analysis of target genes in lncRNA-miRNA-mRNA ceRNA network

Types	GO subgroup	No. of enrich genes	*P*	Genes
BP	GO:0071356 (cellular response to tumor necrosis factor)	3	0.016	*TNFRSF21*, *MAP3K5*, *SFRP1*
BP	GO:0071391 (cellular response to estrogen stimulus)	2	0.023	*SFRP1*, *MYC*
BP	GO:0014898 (cardiac muscle hypertrophy in response to stress)	2	0.026	*EZH2*, *INPP5F*
BP	GO:0008283 (cell proliferation)	4	0.027	*FSCN1*, *PIM1*, *ENPEP*, *MYC*
BP	GO:0014911 (positive regulation of smooth muscle cell migration)	2	0.035	*SEMA6D*, *MYC*
BP	GO:0051894 (positive regulation of focal adhesion assembly)	2	0.037	*LIMS1*, *SFRP1*
BP	GO:0016477 (cell migration)	3	0.038	*FSCN1*, *TGFBR3*, *ENPEP*

续表5

Types	GO subgroup	No. of enrich genes	*P*	Genes
BP	GO:0042127 (regulation of cell proliferation)	3	0.043	*TNFRSF21*, *EZH2*, *PLAU*
BP	GO:0006306 (DNA methylation)	2	0.044	*FOS*, *EZH2*
CC	GO:0005886 (plasma membrane)	14	0.013	*TNFRSF21*, *LIMS1*, *LITAF*, *PIM1*, *ENPEP*, *KCNJ2*, *GBA2*, *SLC16A1*, *SFRP1*, *SEMA6D*, *SLC7A1*, *GPAM*, *BMF*, *PLAU*
CC	GO:0005887 (integral component of plasma membrane)	7	0.030	*TNFRSF21*, *SLC16A1*, *SEMA6D*, *SLC7A1*, *TGFBR3*, *ENPEP*, *KCNJ2*
CC	GO:0005790 (smooth endoplasmic reticulum)	2	0.043	*GBA2*, *KCNJ2*
MF	GO:0005515 (protein binding)	22	0.036	*TNFRSF21*, *LIMS1*, *LITAF*, *FSCN1*, *EZH2*, *NR3C2*, *PIM1*, *MXI1*, *AMOTL2*, *SPRY4*, *TPM3*, *RPS6KA5*, *FOS*, *MAP3K5*, *SFRP1*, *ZWINT*, *SLC7A1*, *INPP5F*, *TGFBR3*, *BMF*, *MYC*, *PLAU*

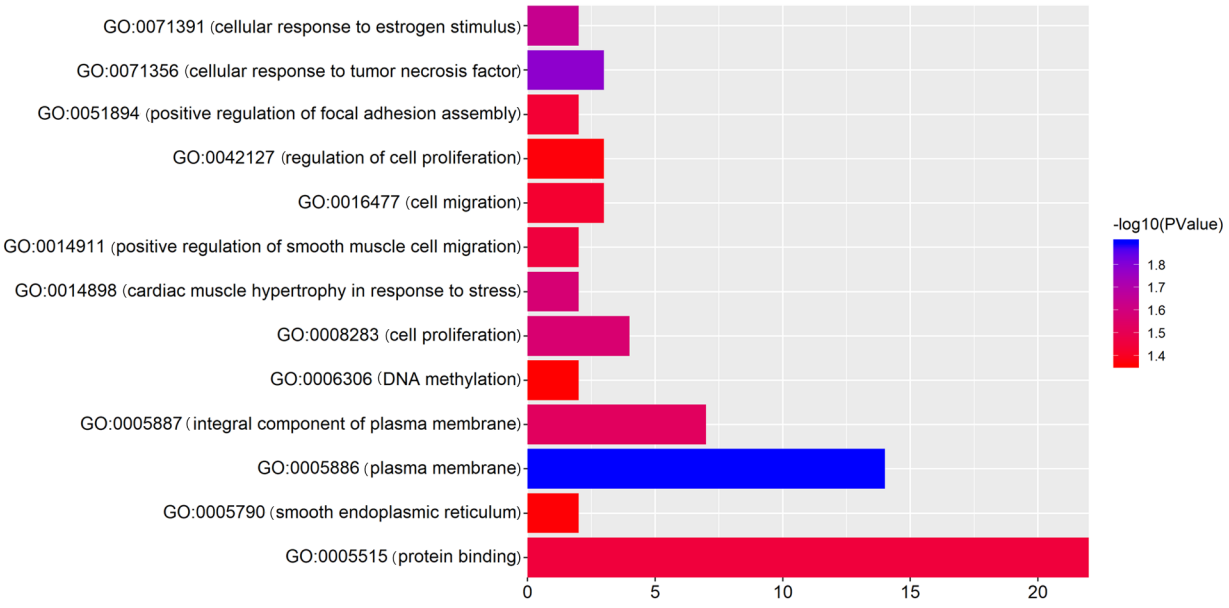

图5 靶基因GO富集分析可视化(条形图)

Figure 5 Visualization of GO enrichment analysis of target genes (barplot)

表 6 LncRNA-miRNA-mRNA ceRNA 网络中靶基因 KEGG 通路富集分析

Table 6 KEGG pathway enrichment analysis of target genes in lncRNA-miRNA-mRNA ceRNA network

KEGG pathway code	KEGG pathway	No. of enrich genes	*P*	Gene code
hsa05206	MicroRNAs in cancer	8	7.70E-09	90427|6541|5292|9252|2146|4609|6624|5328
hsa04010	MAPK signaling pathway	5	7.49E-05	2353|4217|627|9252|4609
hsa04668	TNF signaling pathway	3	0.003	2353|4217|9252
hsa04722	Neurotrophin signaling pathway	3	0.003	9252|4217|627
hsa05216	Thyroid cancer	2	0.004	7170|4609
hsa05219	Bladder cancer	2	0.007	9252|4609
hsa05221	Acute myeloid leukemia	2	0.011	5292|4609
hsa05210	Colorectal cancer	2	0.011	2353|4609
hsa04713	Circadian entrainment	2	0.023	2353|9252
hsa04725	Cholinergic synapse	2	0.027	3759|2353
hsa05200	Pathways in cancer	3	0.027	2353|7170|4609
hsa04210	Apoptosis	2	0.032	2353|4217
hsa04310	Wnt signaling pathway	2	0.032	6422|4609
hsa05161	Hepatitis B	2	0.032	2353|4609
hsa04261	Adrenergic signaling in cardiomyocytes	2	0.032	9252|7170
hsa04630	Jak-STAT signaling pathway	2	0.032	5292|4609
hsa04921	Oxytocin signaling pathway	2	0.032	3759|2353
hsa05202	Transcriptional misregulation in cancer	2	0.039	5328|4609
hsa04024	cAMP signaling pathway	2	0.045	2353|627
hsa05205	Proteoglycans in cancer	2	0.045	5328|4609

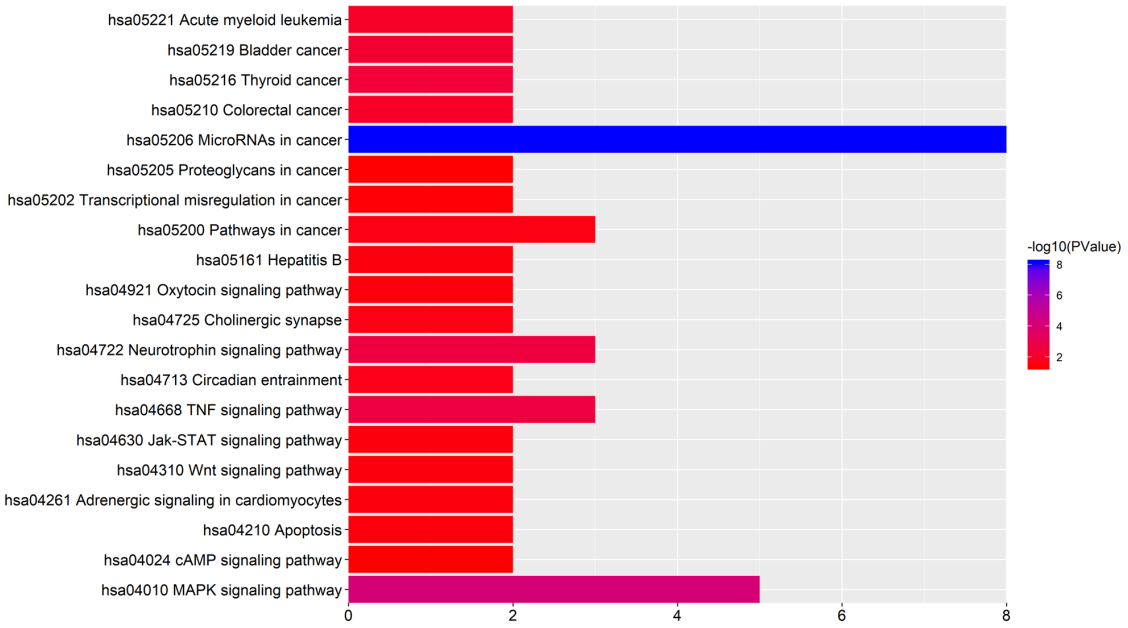

图6 靶基因KEGG通路富集分析可视化(条形图)

Figure 6 Visualization of KEGG pathway enrichment analysis of target genes (barplot)

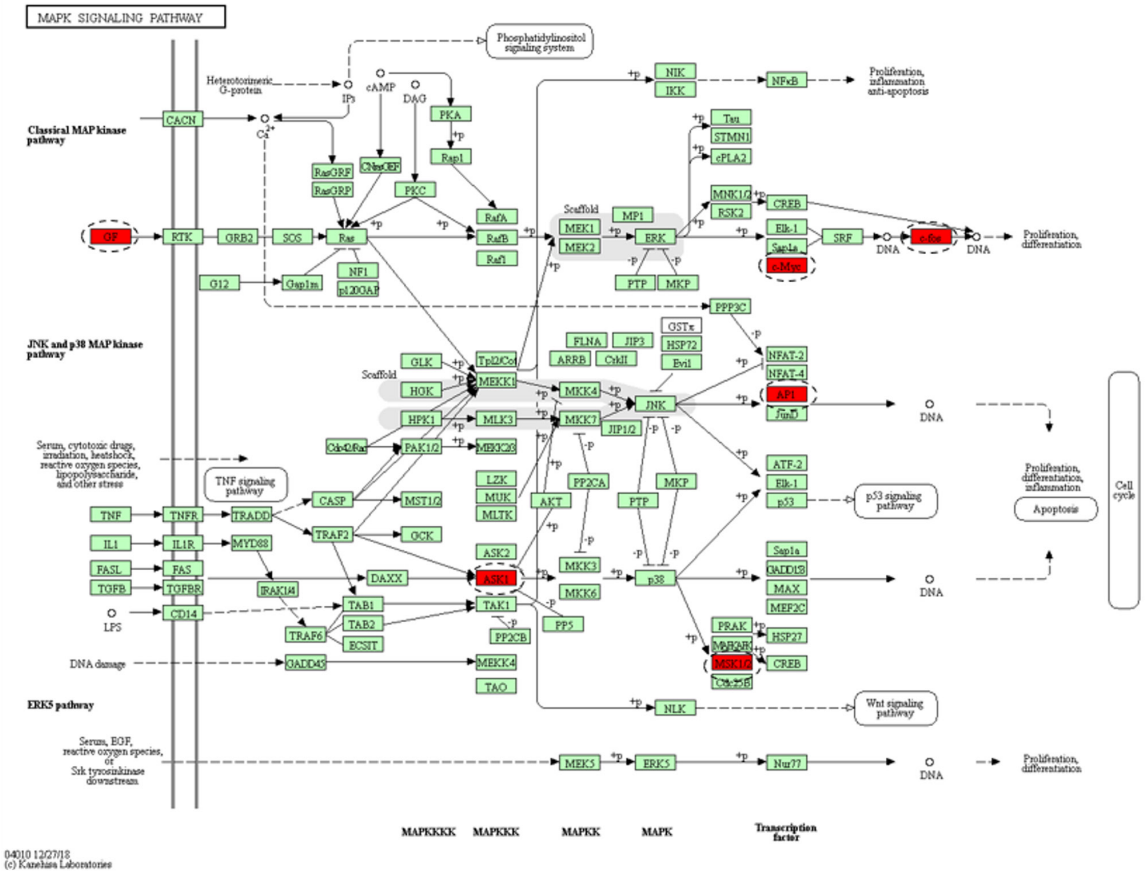

图7 MAPK信号通路图(虚线椭圆框：显著富集)

Figure 7 MAPK signaling pathway (ellipse dotted box: enriched significantly)

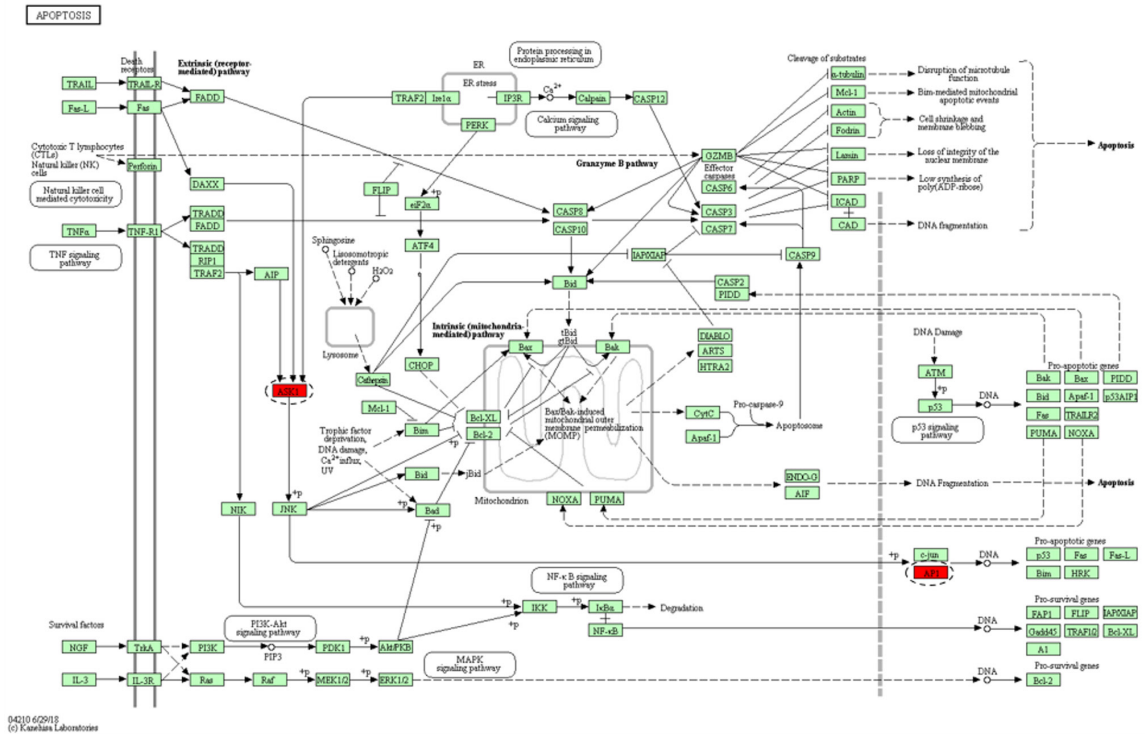

图8 凋亡信号通路图(虚线椭圆框：显著富集)

Figure 8 Apoptosis signaling pathway (ellipse dotted box: enriched significantly)

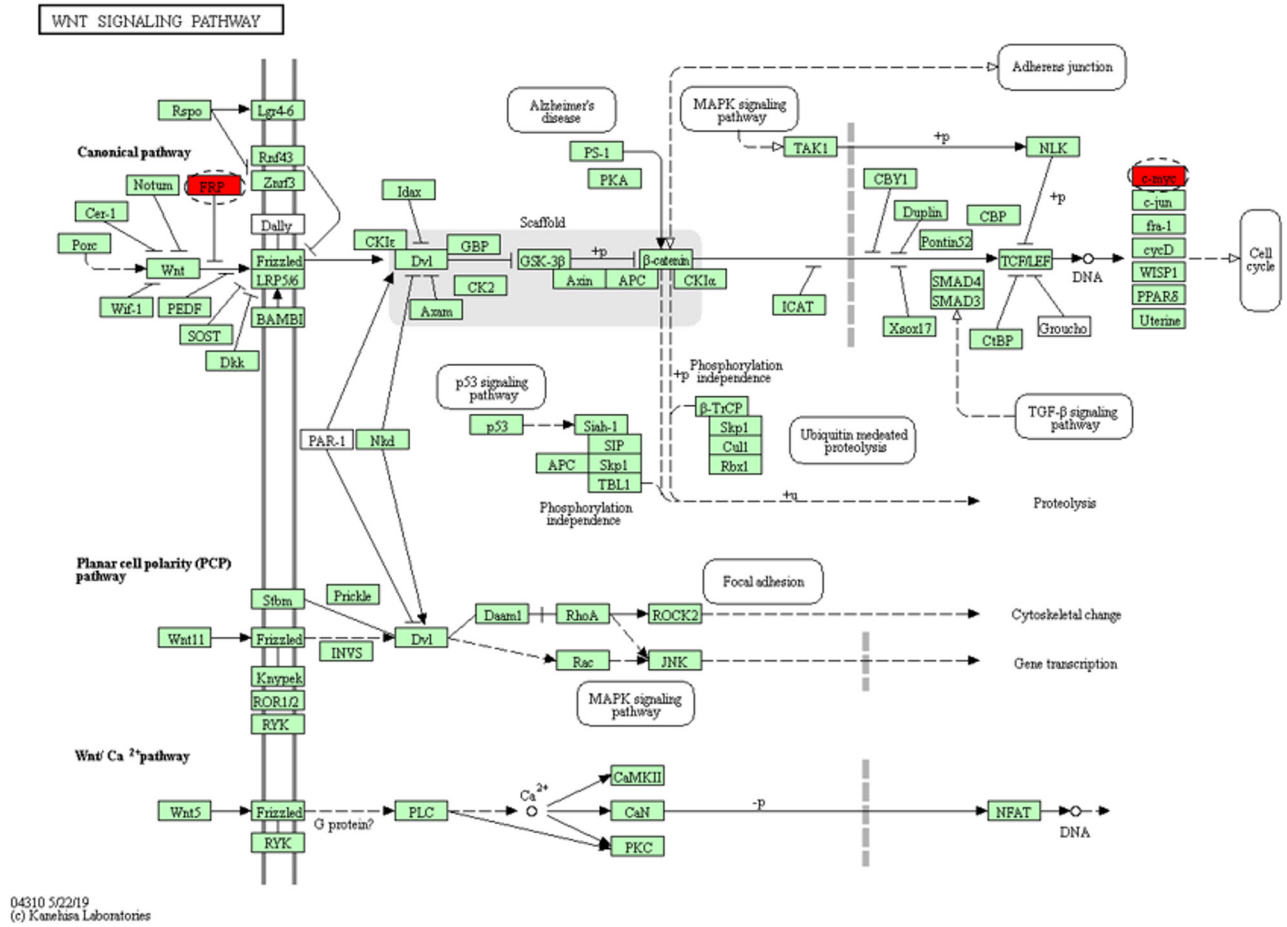

图9 Wnt信号通路图(虚线椭圆框：显著富集)

Figure 9 Wnt signaling pathway(ellipse dotted box: enriched significantly)

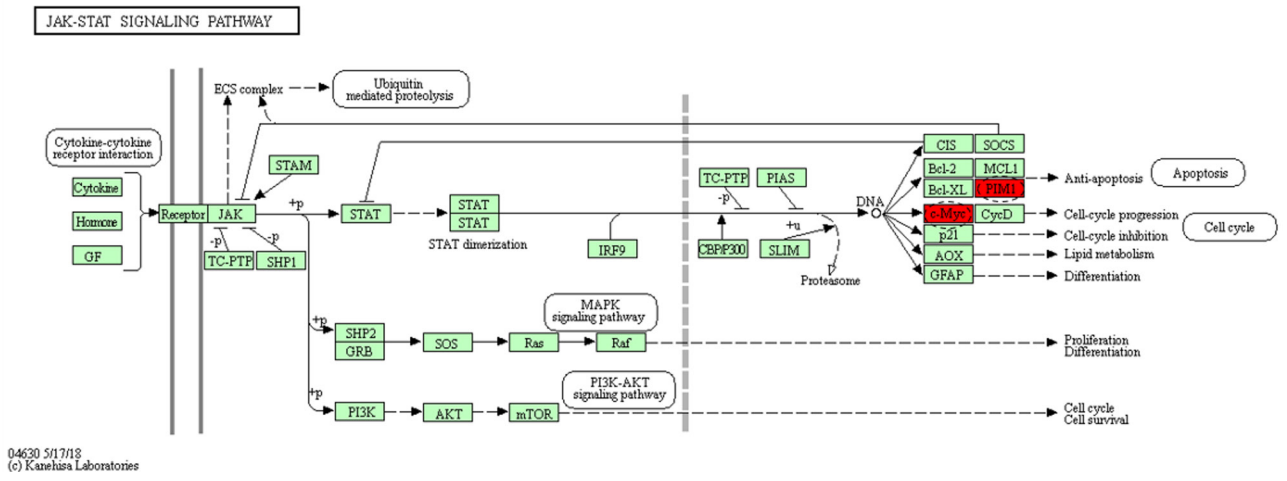

图10 JAK-STAT信号通路图(虚线椭圆框：显著富集)

Figure 10 JAK-STAT signaling pathway(ellipse dotted box: enriched significantly)

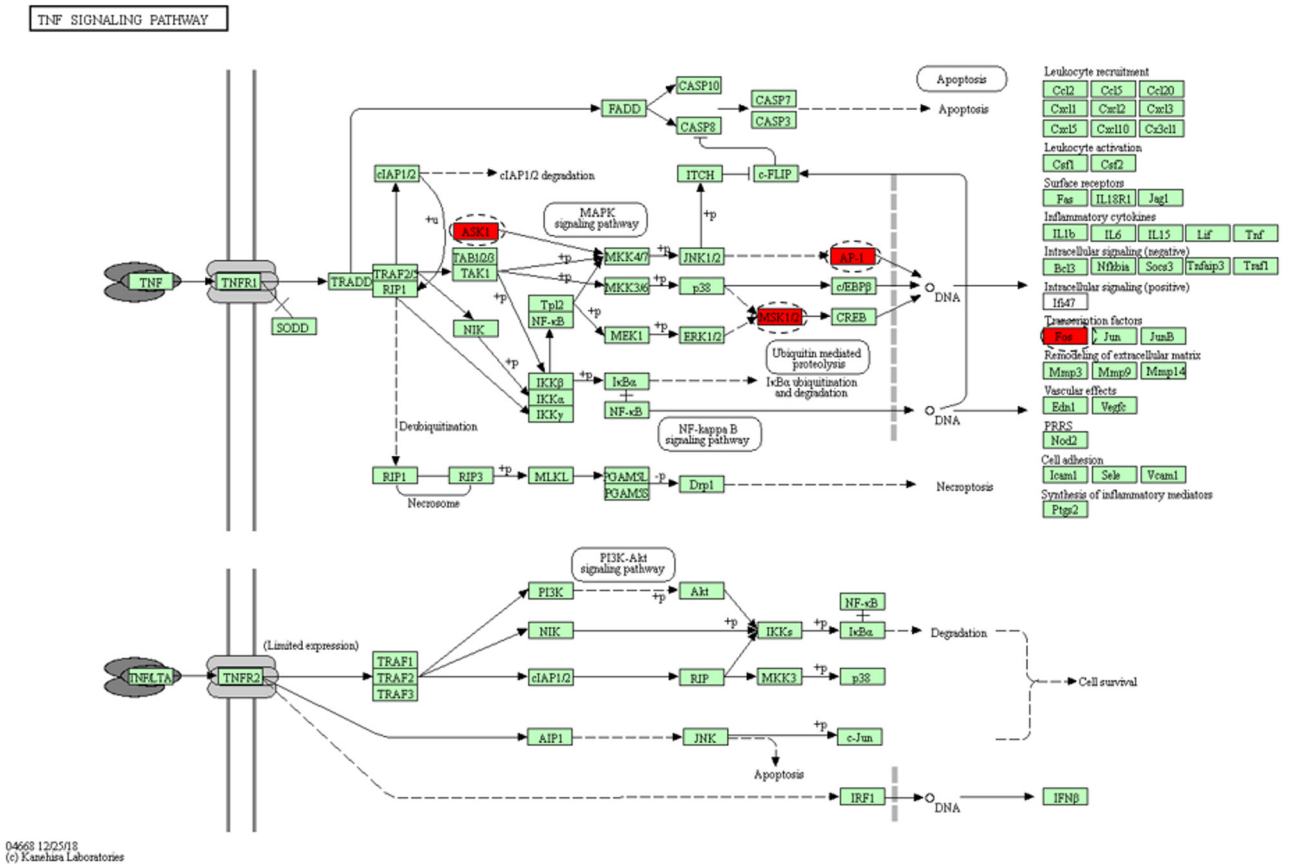

图11 TNF信号通路图(虚线椭圆框：显著富集)

Figure 11 TNF signaling pathway(ellipse dotted box: enriched significantly)

3 讨论

心力衰竭是各种器质性心脏病的临床终末期，其防治已成为心血管病研究领域“最后的战场”。由于其发病机制尚不明确，目前仍缺乏最佳的治疗手段，预后与恶性肿瘤接近。心力衰竭早期多表现为心肌肥厚，是心脏对高负荷状态(如血流动力学应激、高血压)、损伤(如心肌梗死)或基因突变等的适应性调节，通常以心肌细胞增多和心室壁增厚为特征。但若不良刺激长期存在，最终会导致心力衰竭。近些年研究[7]表明：心力衰竭的发生与心肌肥厚、重构、炎症、离子通道等密切相关。心肌肥厚是心力衰竭等心血管疾病的独立危险因素。越来越多的证据[8-9]表明：lncRNA基因调控机制参与心肌肥厚及心力衰竭的发生发展。LncRNA可减少缺血性心肌损伤、心肌肥厚和心肌纤维化，促进心肌血管生成，减少心力衰竭的进展。LncRNA可通过“海绵吸附”机制作用于miRNA，与mRNA竞争抑制miRNA活性，从而微调miRNA介导的靶基因(mRNA)的表达调控[10]。在lncRNA-miRNA-mRNA轴上，“不安分”的lncRNA可能参与多种病理生理过程和基因表达，影响心血管疾病的发生发展。如H19 lncRNA可作用于CaMKIIδ，抑制心肌肥厚[11]。LncRNA Chast最近被证实[12]可促进心肌肥厚，而gapmeR介导的Chast沉默可显著抑制病理性心肌肥厚。因而，探讨肥厚心肌细胞lncRNA表达谱及下游基因表达调控对进一步了解心力衰竭发病机制有着深远的意义。然而，全面研究心肌肥厚基因调控水平却受到人类心肌细胞可用性的限制。既往研究[13]发现：ET-1刺激的hiPSC-CMs可作为人类心肌肥厚的研究模型。

本研究采用生物信息学方法，通过筛选GEO数据库中心肌肥厚相关的芯片数据，发现与无ET1刺激的hiPSC-CMs相比，ET1刺激的hiPSC-CMs中19个lncRNA和717个mRNA表达差异具有统计学意义。其中，lncRNA表达上调的有4个，下调的有15个；mRNA表达上调的有312个，下调的有405个。进一步对差异表达的lncRNA进行miRNA及下游靶基因的预测，并成功构建了

lncRNA-miRNA-mRNA ceRNA调控网络，从中我们发现，上调的AC017002，MIR210HG分别与9，10个miRNA关联密切，下调的LINC00342与9个miRNA关联密切。研究[14-17]发现lncRNA MIR210HG可“海绵吸附”miRNA，调控基因表达，影响骨肉瘤、肝细胞癌、乳腺浸润性癌等恶性肿瘤的发生发展，并可作为肿瘤性疾病的预后标志物。而其在心肌肥厚的病理机制中是否可以类似的作用方式影响疾病过程尚无文献报道。LINC00342被发现可靶向作用于miR-203a-3p而调节非小细胞肺癌癌细胞的生长及代谢，而其在急性心肌梗死发生发展中也扮演着重要角色[18-19]。AC017002是否参与基因表达调控，尚未见报道。本研究进一步预测miRNA下游的靶基因发现，在ceRNA网络中有18个mRNA显著上调，即TPM3，PIM1，AMOTL2，INPP5F，FSCN1，KANK4，LITAF，SLC7A1，KCNJ2，GBA2，ENPEP，BDNF，TNFRSF21，FOS，LIMS1，SPRY4，MYC，PLAU；14个mRNA显著下调，即SEMA6D，ARL15，RPS6KA5，NR3C2，SFRP1，TGFBR3，BMF，EZH2，MAP3K5，SLC16A1，MXI1，GPAM，ZWINT，BHLHB9。再对靶基因进行GO富集分析、KEGG通路富集分析，初步探索了该基因芯片的基因表达差异及ceRNA调控网络。

早期的心肌肥厚是心肌细胞对外界刺激的代偿反应，是一种适应性的改变，以保证机体对心输出量增加的需求。心肌肥厚包括心肌细胞肥大、心肌间质细胞增殖以及细胞外基质重建等多方面复杂的病理生理过程，即心肌重构，期间心肌细胞内蛋白合成显著增加、细胞体积增大,肌节数量增多并伴有纤维组织增生。GO富集分析发现：靶基因被富集到13个不同的GO子集中，包括BP，CC和MF 3个方面，其中各自最显著富集的子集分别为细胞增殖、质膜和蛋白结合。上述各子集中富集基因数目均较多，提示这些BP，CC和MF及其富集的基因或蛋白分子在心肌肥厚的发生、发展中扮演着重要的角色。进一步行KEGG通路富集分析发现：靶基因被富集到MAPK信号通路、细胞凋亡、Wnt信号通路、JAK-STAT信号通路、TNF信号通路等20个信号通路上。MAPK信号通路是一种经典的炎症信号通路，广泛参与心肌肥厚、心肌重构、心力衰竭等心血管疾病的病理机制[20]。国外有研究[21]报道：在右心室肥大伴心力衰竭的小鼠模型中，p38 MAPK的抑制剂PH797804可提高右心室的功能，并抑制心肌纤维化。在主动脉结扎诱导的大鼠心肌肥厚模型中，使用MAPK的抑制剂如吡格列酮也可减弱心肌肥厚[22]。心肌肥厚是导致心力衰竭的主要危险因素。在导致心力衰竭的病理重塑过程中，心肌细胞失代偿性肥大、凋亡或坏死、纤维化导致心肌细胞死亡以及心肌细胞的进行性功能障碍是明显的。心肌细胞肥大、凋亡或死亡，是一个渐变或逐渐演进的过程，与心力衰竭的发生密切相关[23]。在腹主动脉缩窄术诱导的心肌肥厚大鼠模型中，非甾体类抗炎药塞来昔布可抑制心肌细胞的炎症反应、凋亡及氧化应激，从而减轻心肌肥厚[24]。AngII受体抑制剂替米沙坦可以剂量依赖方式通过抑制活化T细胞核因子核易位、A型脑钠肽和B型脑钠肽的表达及心肌细胞凋亡等方式来改善因后负荷过重导致的心肌细胞肥大[25]。上述研究均表明：MAPK及凋亡信号通路在心肌肥厚、心力衰竭的发生发展中起重要作用。KEGG富集分析发现：在MAPK或凋亡信号通路中，ASK1(即MAP3K5)表达下调，AP1转录因子(通常以jun与Fos家族成员组成的同源或异源二聚体表达其活性)表达上调，MSK1/2(即RPS6KA5)表达下调，c-MYC表达上调，参与心肌细胞细胞增殖分化及凋亡等。Wnt信号通路是参与心肌肥大的关键通路之一，该通路的阻断剂CGX1321可抑制腹主动脉横缩术诱导的活化T细胞的核因子的核易位和磷酸化c-JUN表达的上调，最终减轻心肌肥厚、凋亡和纤维化，其有可能成为治疗心肌肥厚患者的新型药物[26]。跨膜受体重组蛋白FZD1(Frizzled-1)免疫疗法可阻断Wnt信号通路受体，削弱小鼠心梗后心室肥大并改善心脏的功能[27]。本研究发现在Wnt信号通路中，分泌的卷曲相关蛋白1(secreted frizzled-related protein 1，SFRP1)表达下调，c-MYC表达上调，参与细胞周期，与心肌肥大密切相关。有报道[28]指出，以盐酸水苏碱干预大鼠可阻断NF-κB及JAK-STAT信号通路而抑制炎症反应及氧化应激，从而改善异丙肾上腺素诱导的心肌肥厚。同样，降脂药辛伐他汀也可调整JAK-STAT信号通路而阻止丙肾上腺素诱导的心肌肥厚[29]。这些都提示JAK-STAT信号通路在心肌肥厚的发生发展中同样起着不容小觑的作用。而通路富集分析发现JAK-STAT信号通路中PIM1及MYC表达均上调，参与细胞周期及抗凋亡作用。国内有研究者[30]报道：组织蛋白酶B是心肌肥厚必要的调节器，其通过调节TNF-α/ASK1/JNK信号通路而参与心肌重塑。这间接提示TNF信号通路在心肌肥厚病理机制中同样扮演着举足轻重的角色。在该通路中，显著富集的基因有ASK1，

MSK1/2及AP1，参与细胞炎症等生理过程。此外，肥厚心肌细胞差异表达基因也显著富集于癌症中的miRNAs、甲状腺癌、膀胱癌、急性髓系白血病、结直肠癌、癌症中转录错误调节、癌症中蛋白聚糖等多个信号通路上，提示心肌肥厚、心力衰竭与癌症等疾病可能享有共同发病机制，值得进一步探讨。

虽然GO和KEGG富集分析为探索心肌肥厚乃至心力衰竭的发病机制提供了丰富的结果；然而，其具体调控机制仍不明确，需要更进一步的研究来阐明这些差异基因的作用。本研究发现lncRNA-miRNA-mRNA ceRNA调控网络可能为研究心肌肥厚的发病机制提供了一个新的视角。笔者实验室后期将进行更深入的研究。

综上所述，本研究探讨了肥厚心肌细胞lncRNAs的表达谱，并构建了ceRNA调控网络，该网络中有3个lncRNA(表达上调的为AC017002、MIR210HG，表达下调的为LINC00342)可能以竞争性“海绵吸附”miRNA的方式调控下游mRNA的表达水平，使AP1，c-MYC和PIM1表达上调，ASK1(即MAP3K5)，MSK1/2(即RPS6KA5)和SFRP1表达下调，并通过在MAPK信号通路、细胞凋亡、Wnt信号通路、JAK-STAT信号通路、TNF信号通路等途经中发挥不同的调控作用，从而参与心肌肥厚的病理生理过程。这些基因或通路有可能成为潜在的治疗靶点，研发阻断上述lncRNA或下游靶基因的相关药物可能对治疗心肌肥厚有着深远的临床意义。本研究丰富了对心肌肥厚发病机制的认识，并为研发新型治疗药物提供实验理论依据。

参考文献

1. Heger J, Schulz R, Euler G. Molecular switches under TGFbeta signalling during progression from cardiac hypertrophy to heart failure[J]. Br J Pharmacol, 2016, 173(1): 3-14.
2. Akhade VS, Pal D, Kanduri C. Long noncoding RNA: genome organization and mechanism of action[J]. Adv Exp Med Biol, 2017, 1008(1): 47-74.
3. Weidle UH, Birzele F, Kollmorgen G, et al. Long non-coding RNAs and their role in metastasis. Cancer Genomics Proteomics, 2017, 14(3): 143-160.
4. Shen S, Jiang H, Bei Y, et al. Long non-coding RNAs in cardiac remodeling[J]. Cell Physiol Biochem, 2017, 41(5): 1830-1837.
5. Uchida S, Dimmeler S. Long noncoding RNAs in cardiovascular diseases[J]. Circ Res, 2015, 116(4): 737-750.
6. Bayzigitov DR, Medvedev SP, Dementyeva EV, et al. Human Induced pluripotent stem cell-derived cardiomyocytes afford new opportunities in inherited cardiovascular disease modeling[J]. Cardiol Res Pract, 2016, 2016: 3582380.
7. Braunwald E. Research advances in heart failure: a compendium[J]. Circ Res, 2013, 113(6): 633-645.
8. Li Y, Liang Y, Zhu Y, et al. Noncoding RNAs in cardiac hypertrophy[J]. J Cardiovasc Transl Res, 2018, 11(6): 439-449.
9. Wang K, Liu F, Zhou LY, et al. The long noncoding RNA CHRF regulates cardiac hypertrophy by targeting miR-489[J]. Circ Res, 2014, 114(9): 1377-1388.
10. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition[J]. Nature, 2014, 505(7483): 344-352.
11. Liu L, An X, Li Z, et al. The H19 long noncoding RNA is a novel negative regulator of cardiomyocyte hypertrophy[J]. Cardiovasc Res, 2016, 111(1): 56-65.
12. Viereck J, Kumarswamy R, Foinquinos A, et al. Long noncoding RNA Chast promotes cardiac remodeling[J]. Sci Transl Med, 2016, 8(326): 326ra22.
13. Burridge PW, Keller G, Gold JD, et al. Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming[J]. Cell Stem Cell, 2012, 10(1): 16-28.
14. Wang Y, Li W, Chen X, et al. MIR210HG predicts poor prognosis and functions as an oncogenic lncRNA in hepatocellular carcinoma[J]. Biomed Pharmacother, 2019, 111(1): 1297-1301.
15. Li J, Wu QM, Wang XQ, Zhang CQ. Long noncoding RNA miR210HG sponges miR-503 to facilitate osteosarcoma cell invasion and metastasis[J]. DNA Cell Biol, 2017, 36(12): 1117-1125.
16. Li XY, Zhou LY, Luo H, et al. The long noncoding RNA MIR210HG promotes tumor metastasis by acting as a ceRNA of miR-1226-3p to regulate mucin-1c expression in invasive breast cancer[J]. Aging (Albany NY), 2019, 11(15): 5646-5665.
17. He Z, Dang J, Song A, et al. Identification of LINC01234 and MIR210HG as novel prognostic signature for colorectal adenocarcinoma[J]. J Cell Physiol, 2019, 234(5): 6769-6777.
18. Chen QF, Kong JL, Zou SC, et al. LncRNA LINC00342 regulated cell growth and metastasis in non-small cell lung cancer via targeting miR-203a-3p[J]. Eur Rev Med Pharmacol Sci, 2019, 23(17): 7408-7418.
19. Shen LS, Hu XF, Chen T, et al. Integrated network analysis to explore the key mRNAs and lncRNAs in acute myocardial infarction[J]. Math Biosci Eng, 2019, 16(6): 6426-6437.
20. Liu R, Molkentin JD. Regulation of cardiac hypertrophy and remodeling through the dual-specificity MAPK phosphatases (DUSPs)[J]. J Mol Cell Cardiol, 2016, 101(1): 44-49.
21. Kojonazarov B, Novoyatleva T, Boehm M, et al. p38 MAPK inhibition

improves heart function in pressure-loaded right ventricular hypertrophy[J]. Am J Respir Cell Mol Biol, 2017, 57(5): 603-614.
22. Wei WY, Ma ZG, Xu SC, Zhang N, Tang QZ. Pioglitazone protected against cardiac hypertrophy via inhibiting AKT/GSK3beta and MAPK signaling pathways[J]. PPAR Res, 2016, 2016: 9174190.
23. Banerjee P, Chander V, Bandyopadhyay A. Balancing functions of annexin A6 maintain equilibrium between hypertrophy and apoptosis in cardiomyocytes[J]. Cell Death Dis, 2015, 6(1): e1873.
24. Zhang C, Wang F, Zhang Y, et al. Celecoxib prevents pressure overload-induced cardiac hypertrophy and dysfunction by inhibiting inflammation, apoptosis and oxidative stress[J]. J Cell Mol Med, 2016, 20(1): 116-127.
25. Li X, Lan Y, Wang Y, et al. Telmisartan suppresses cardiac hypertrophy by inhibiting cardiomyocyte apoptosis via the NFAT/ANP/BNP signaling pathway[J]. Mol Med Rep, 2017, 15(5): 2574-2582.
26. Jiang J, Lan C, Li L, et al. A novel porcupine inhibitor blocks WNT pathways and attenuates cardiac hypertrophy[J]. Biochim Biophys Acta Mol Basis Dis, 2018, 1864(10): 3459-3467.
27. Fan J, Qiu L, Shu H, et al. Recombinant frizzled1 protein attenuated cardiac hypertrophy after myocardial infarction via the canonical Wnt signaling pathway[J]. Oncotarget, 2017, 9(3): 3069-3080.
28. Zhao L, Wu D, Sang M, et al. Stachydrine ameliorates isoproterenol-induced cardiac hypertrophy and fibrosis by suppressing inflammation and oxidative stress through inhibiting NF-kappaB and JAK/STAT signaling pathways in rats[J]. Int Immunopharmacol, 2017, 48(1): 102-109.
29. Al-Rasheed NM, Al-Oteibi MM, Al-Manee RZ, et al. Simvastatin prevents isoproterenol-induced cardiac hypertrophy through modulation of the JAK/STAT pathway[J]. Drug Des Devel Ther, 2015, 9(1): 3217-3229.
30. Wu QQ, Xu M, Yuan Y, et al. Cathepsin B deficiency attenuates cardiac remodeling in response to pressure overload via TNF-alpha/ASK1/JNK pathway[J]. Am J Physiol Heart Circ Physiol, 2015, 308(9): H1143-H1154.

本文引用：尤红俊, 赵倩倩, 常凤军, 韩稳琦, 程功, 寿锡凌, 刘富强, 刁佳宇. 心肌肥厚模型：内皮素1刺激人诱导的多能干细胞源性心肌细胞lncRNA表达谱及ceRNA网络的生物信息学分析[J]. 临床与病理杂志, 2020, 40(11): 2789-2804. doi: 10.3978/j.issn.2095-6959.2020.11.001

Cite this article as: YOU Hongjun, ZHAO Qianqian, CHANG Fengjun, HAN Wenqi, CHENG Gong, SHOU Xiling, LIU Fuqiang, DIAO Jiayu. Bioinformatics analysis of lncRNA expression profiles and competing endogenous RNA network of myocardial hypertrophy model: human induced pluripotent stem cell derived cardiomyocytes stimulated by endothelin 1[J]. Journal of Clinical and Pathological Research, 2020, 40(11): 2789-2804. doi: 10.3978/j.issn.2095-6959.2020.11.001