

鸡 miR - 1749 生物信息学分析及真核表达载体构建

王顺红,王善禾,蒋可人,李 红,康相涛,孙桂荣*
(河南农业大学 牧医工程学院,河南 郑州 450002)

摘要: 为研究鸡 miR - 1749 功能及其前体 rs15860000 (C > T) 不同基因型对成熟 miRNA 生成的影响,采用 Mfold 软件预测 C > T 突变对 miR - 1749 前体二级结构能值的影响;选用 pcDNA3.1 - EGFP 质粒,构建 miR - 1749 前体不同等位基因的真核表达载体并转染 DF1 细胞,检测成熟 miR - 1749 的表达;利用 TargetScan 和 miRDB 软件预测 miR - 1749 的靶基因,并将靶基因的并集在 DAVID 数据库进行 GO 功能富集和 KEGG 信号转导通路分析。Mfold 预测结果显示,miR - 1749 前体 C > T 突变影响其前体二级结构自由能值;成功构建 pcDNA3.1 - EGFP - pre - miR - 1749 - C 和 pcDNA3.1 - EGFP - pre - miR - 1749 - T 真核表达载体,转染 DF1 细胞后,不同基因型重组质粒 miR - 1749 的表达量均显著高于空质粒组 ($P < 0.01$),T 等位基因成熟 miR - 1749 的表达量显著高于 C 等位基因 ($P < 0.05$)。2 个软件预测结果显示,miR - 1749 靶基因的并集有 250 个基因,这些基因显著富集于 Ras 蛋白信号转导调控、GTPase 介导信号转导调控、细胞稳态及胚后发育的生物学过程和 GnRH、MAPK 及细胞黏附分子的信号转导通路。

关键词: 鸡; miR - 1749; 真核表达载体; 单核苷酸多态性; 靶基因

中图分类号: S831.2 **文献标志码:** A **文章编号:** 1004 - 3268 (2015) 05 - 0129 - 05

Bioinformatics Analysis and Eukaryotic Expression Vector Construction of miR-1749 in Chicken

WANG Shunhong, WANG Shanhe, JIANG Keren, LI Hong, KANG Xiangtao, SUN Guirong*
(College of Animal Science and Veterinary Medicine, Henan Agricultural University, Zhengzhou 450002, China)

Abstract: The purpose of this study was to explore the function of miR-1749 and the effect of rs15860000 (C > T) polymorphism on mature miR-1749 biogenesis. The effect of rs15860000 on the secondary structure of miR-1749 precursor was calculated by the Mfold web server. miR-1749 expression vectors with different alleles were constructed using pcDNA3.1-EGFP plasmid and transfected into DF1 cells. Then, the expression of mature miR-1749 was detected. The potential targets of miR-1749 were predicted by TargetScan and miRDB and further evaluated by DAVID. The result showed that the C > T mutation caused the free energy change of miR-1749 precursor secondary structure. The pcDNA3.1-EGFP-pre-miR-1749-C and pcDNA3.1-EGFP-pre-miR-1749-T expression vector were constructed successfully. The expression of miR-1749 in cells transfected with the recombinant vector was significantly higher than cells transfected with the empty vector ($P < 0.01$), while miR-1749 expression of the cells transfected with pcDNA3.1-EGFP-miR-1749-T was significantly higher than the cells transfected with pcDNA3.1-EGFP-miR-1749-C ($P < 0.05$). A total of 250 targets were predicted by two softwares. These genes significantly focused on regulation of Ras protein signal transduction, regulation of small GTPase mediated signal transduction, cellular homeostasis and post-embryonic development biology process and GnRH signaling pathway, MAPK

收稿日期:2015 - 01 - 02
基金项目:国家自然科学基金项目(31201795);第 53 批博士后基金二等资助项目(2013M531675)
作者简介:王顺红(1989 -),女,河南浚县人,在读硕士研究生,研究方向:家禽遗传育种。E - mail:wang_shunhong@163.com
* 通讯作者:孙桂荣(1976 -),女,河南沈丘人,副教授,博士,主要从事家禽遗传育种研究。E - mail:grsun2000@126.com

signaling pathway and cell adhesion molecules (CAMs).

Key words: chicken; miR-1749; eukaryotic expression vector; single nucleotide polymorphism; target genes

miRNA 是一类内源性、长约 22 个核苷酸的非编码 RNA,通过与靶基因 mRNA 的 3'非编码区(untranslated region, UTR)结合,在转录后水平降解靶标 mRNA 或抑制蛋白质翻译^[1]。在高等真核生物中,30% 以上的基因可能受 miRNA 的调控^[2]。miRNA 在动物体多个调节途径中发挥关键作用,包括动物的生长发育^[3]、细胞的增殖^[4]和凋亡^[5]。

随着 miRNA 研究的深入,研究者发现,miRNA 不仅参与了动物复杂性状表型形成的分子调控过程,而且 miRNA 基因的多态性可能影响 miRNA 的表达,导致 miRNA 功能异常,从而影响 miRNA 靶基因的表达和功能,进而引起生物表型性状的变异^[6-7]。

鸡是一种重要的经济禽类和模式生物。目前,研究者已经在鸡上发现了 740 个 miRNA 前体和 994 个 miRNA 成熟体(miRBase Release 21, Gallus - galus - 4.0)。Glazov 等^[8]通过高通量测序技术在鸡胚中发现 miR - 1749,但其功能和作用机制尚未见报道。笔者所在课题组(河南农业大学家禽种质资源创新工程研究中心)前期研究发现,miR - 1749 前体存在 rs15860000 (C > T)单核苷酸多态位点,且与鸡的体质量显著相关(数据未发表)。鉴于此,本研究拟采用生物信息学方法,通过对其潜在靶基因进行预测和功能富集分析来了解 miR - 1749 可能参与的生物学功能;利用在线软件 Mfold 预测 C > T 突变对其前体二级结构稳定性的影响;同时构建 miR - 1749 rs15860000 位点不同纯合基因型的真核表达载体,检测 C > T 突变对成熟 miRNA 表达量的影响,为进一步研究 miR - 1749 的生物学功能和靶基因鉴定提供参考。

1 材料和方法

1.1 试验材料

固始鸡与安卡鸡 F₂ 代资源群的 DNA、*E. coli* DH5 α 菌株、pcDNA3.1 - EGFP 载体和 DF1 细胞系由河南省家禽种质资源创新工程研究中心保存;限制性内切酶、T4 DNA 连接酶购自 NEB 公司;DNA 纯化回收试剂盒、无内毒素质粒提取试剂盒购自天根公司;Lipofectamine2000 转染试剂购自上海英骏生物科技有限公司;RNA 提取试剂、反转录试剂盒、荧光定量试剂购自宝生物工程(大连)有限公司;引物由生工生物工程(上海)有限公司合成;DNA 测序由北京六合华大基因科技股份有限公司完成。

1.2 miR - 1749 的生物信息学分析

在 NCBI 数据库检索鸡 miR - 1749,查找其在基因组上的位置,获得其前体序列。在前体序列两端各 100 多个碱基处设计引物^[9],扩增 miR - 1749 的前体序列。引物序列为:上游引物 5' - ccgCTCGAGGATACGTCAGTCTGTCTC - 3',下游引物 5' - tgcTCTAGATCCTTCAGCCACTGTGATCTG - 3'。2 条引物的 5'端分别添加 *Xho* I 和 *Xba* I 酶切位点(下划线字母为酶切序列,小写字母为保护碱基)。利用 Mfold 在线软件(<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>),装载 pre - miR - 1749 野生型及变异体序列,在默认条件下预测其二级结构^[10]。

1.3 miRNA 表达载体的构建与鉴定

1.3.1 重组质粒连接 将纯化的 miR - 1749 的 PCR 产物和 pcDNA3.1 - EGFP 质粒分别用 *Xho* I 和 *Xba* I 限制性核酸内切酶进行双酶切,酶切产物纯化后用 T4 DNA 连接酶 16 ℃ 连接过夜。

1.3.2 重组质粒转化 取 10 μ L 连接产物转化 100 μ L DH5 α 感受态细胞,涂布于 LB 平板(含氨苄青霉素 100 μ g/mL),37 ℃ 恒温培养箱培养过夜。

1.3.3 重组质粒阳性克隆筛选和测序 挑取单克隆菌落接种于含氨苄青霉素(100 μ g/mL)的 LB 液体培养基中,37 ℃ 恒温摇床培养 12 ~ 16 h,进行菌液 PCR 并用 1% 琼脂糖凝胶电泳检测,选取阳性克隆菌株进行测序鉴定;miR - 1749 前体 rs15860000 (C > T) 不同等位基因的重组质粒分别命名为 pcDNA3.1 - EGFP - miR - 1749 - C 和 pcDNA3.1 - EGFP - miR - 1749 - T。

1.4 细胞培养及转染

DF1 细胞在含 10% 胎牛血清的 DMEM/f12 培养基、37 ℃ 和 5% CO₂ 的饱和湿度下培养。选择生长状态良好的细胞,以 3.5×10^5 个/孔接种于细胞培养板 12 孔板中,细胞生长至融合 70% 时,分别将 pcDNA3.1 - EGFP - miR - 1749 - C、pcDNA3.1 - EGFP - miR - 1749 - T 用脂质体 Lipofectamine2000 转染细胞,pcDNA3.1 - EGFP 空质粒作为阴性对照,操作按照 Lipofectamine2000 说明书进行。

1.5 实时荧光定量 PCR 检测重组质粒过表达效果

转染后 48 h 收集细胞,Trizol 法抽提总 RNA,紫外分光光度计检测 RNA 含量和纯度,用反转录试剂盒获得 cDNA。miR - 1749 定量和内参 5S rRNA 引物购自上海吉玛生物制药有限公司。定量 PCR 体

系为 20 μ L;2 \times UltraSYBR Mixture 10 μ L、上下游引物 2 μ L(10 μ mol/L)、cDNA 1 μ L,加去 RNA 酶水至总体积 20 μ L;反应程序为:95 $^{\circ}$ C 预变性 3 min;95 $^{\circ}$ C 变性 12 s,60 $^{\circ}$ C 退火 40 s,40 个循环。采用 $2^{-\Delta\Delta C_t}$ 法计算 miR - 1749 的相对表达量^[11],利用 SPSS 20.0 软件进行单因素方差分析。

1.6 miR - 1749 靶基因预测和功能分析

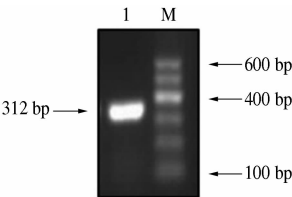
利用 TargetScan (<http://www.targetscan.org/>)^[12] 和 miRDB (<http://mirdb.org/miRDB/>)^[13] 在线软件预测 miR - 1749 的靶基因,取两者预测结果的集合,应用 DAVID (<http://david.abcc.ncifcrf.gov/>)^[14] 分析工具对 miR - 1749 的靶基因进行基于 GO 的功能富集和基于 KEGG 的信号转导通路分析。

2 结果与分析

2.1 鸡 miR - 1749 前体的克隆与结构分析

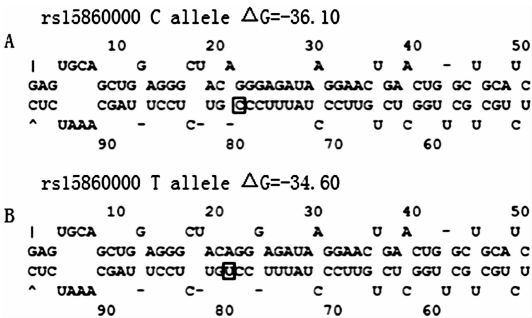
鸡 miR - 1749 基因位于 2 号染色体上,前体序列为 99 bp。PCR 扩增得到 miR - 1749 前体及其侧翼序列的 DNA 片段共 312 bp(图 1)。采用 Mfold 软件对 miR - 1749 前体 rs15860000 (C > T) 不同基因型二级结构进行预测,结果表明,不同基因型的 miR - 1749 前体都能够形成具有发夹的二级结构,符合 miRNA 的结构特征(图 2),但 pre - miR - 1749 - C 的二级结构自由能值为 -36.10 kcal/mol,pre - miR - 1749 - T 的二级结构自由能为 -34.60 kcal/mol。rs15860000 C > T 导致 miR - 1749 前体二

级结构的自由能值变化 1.5 kcal/mol,使 miR - 1749 前体二级结构的稳定性增强,可能会影响成熟 miR - 1749 的生成。



M. DNA Maker I ; 1. miR - 1749 基因

图 1 miR - 1749 基因 PCR 扩增



方框表示突变碱基,下同

图 2 pre - miR - 1749 前体 rs15860000 (C > T) 不同基因型的二级结构

2.2 重组质粒的鉴定

菌液 PCR 产物经琼脂糖凝胶电泳显示,扩增得到的 DNA 片段大小符合预期结果,初步证明 pre - miR - 1749 片段已插入到载体中。重组质粒测序结果(图 3)与 NCBI 中的序列一致,表明重组质粒构建成功。

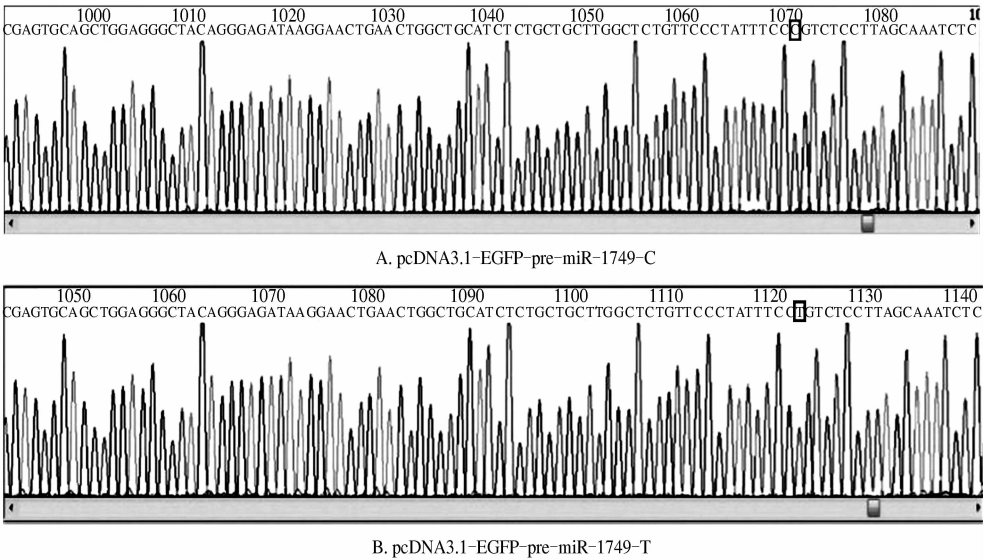


图 3 pre - miR - 1749 rs15860000 (C > T) 不同基因型重组质粒测序结果

2.3 重组质粒转染 DF1 细胞后 miR - 1749 的表达

分别将 pcDNA3.1 - EGFP - miR - 1749 - C/T 重组质粒和 pcDNA3.1 - EGFP 空质粒转染 DF1 细

胞,24 h 后,荧光显微镜观察到绿色荧光,表明 DF1 细胞中有重组质粒导入和表达(图 4)。提取总 RNA 进行 miR - 1749 基因的实时荧光定量 PCR,结

果表明,转染不同基因型重组质粒组 miR - 1749 的表达量显著高于阴性对照组 ($P < 0.01$) (图 5),并且转染 pcDNA3.1 - EGFP - miR - 1749 - T 组 miR -

1749 的表达量显著高于转染 pcDNA3.1 - EGFP - miR - 1749 - C 组 ($P < 0.05$),表明 rs15860000 (C > T) 影响成熟 miR - 1749 的表达。

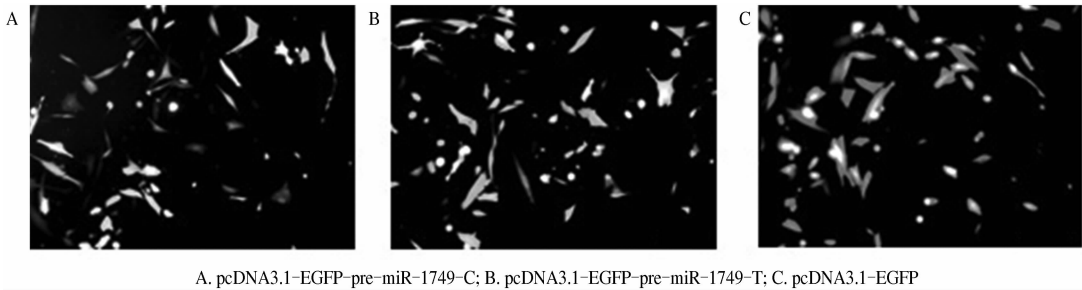


图 4 细胞转染后 24 h 荧光显微镜检测绿色荧光蛋白

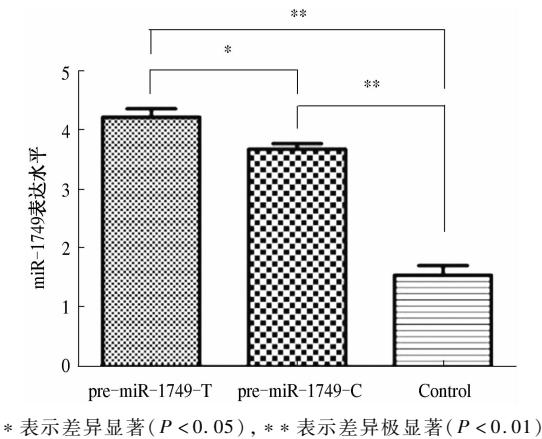


图 5 miR - 1749 在 DF1 细胞中的表达

2.4 miR - 1749 靶基因预测和功能分析

为探讨 miR - 1749 的功能,对其靶基因进行预测和分析。TargetScan 预测得到 miR - 1749 的靶基因 31 个,miRDB 预测得到 miR - 1749 的靶基因 229 个。2 种预测方法的交集得到 10 个共同靶标,并集得到靶基因 250 个。对预测靶基因的并集进行 GO 功能富集和 KEGG 分析,GO 功能富集分析结果显示,生物学过程显著富集于 Ras 蛋白信号转导调控、GTPase 介导信号转导调控、细胞稳态和胚后发育 ($P < 0.05$) (表 1);KEGG 分析显示,miR - 1749 靶基因显著富集于 GnRH、MAPK 和细胞黏附分子信号通路 ($P < 0.05$) (表 2)。

表 1 鸡 miR - 1749 预测靶基因的 GO 生物学过程富集分析

GO 号	GO 名称	P 值	基因名称
0046578	Ras 蛋白信号转导调控	0.014	VAV3、ARHGEF18、PSD3、IQSEC3、USP6NL、KALRN
0051056	小 G 蛋白介导的信号转导调控	0.031	VAV3、ARHGEF18、PSD3、IQSEC3、USP6NL、KALRN
0019725	细胞稳态	0.035	SLC9A8、AIFM3、GNAQ、PKD2、TXNRD1、ITPR1
0009791	胚后发育	0.053	GNAQ、SLC18A2、ITPR1

表 2 鸡 miR - 1749 预测靶基因的 KEGG 通路富集分析

收录号	KEGG 通路	P 值	基因名称
gga04912	GnRH 信号通路	0.018 0	GNAQ、MAP3K3、FSHB、CACNA1D、ITPR1
gga04010	MAPK 信号通路	0.020 9	RPS6KA3、FGFR3、MAP3K3、TAOK1、RRAS2、NFATC2、CACNA1D、AKT3
gga04514	细胞黏附分子	0.037 6	NCAM2、CNTN2、NFASC、CD4、SDC2

3 结论与讨论

研究 miRNA 生物学功能的关键在于其靶基因的确定。目前,对 miRNA 靶基因的研究主要是先通过生物信息学的方法预测,再进行试验验证。miRNA 靶基因的预测软件有很多,但是能对鸡 miRNA 进行预测的只有 TargetScan 和 miRDB。本研究采用 TargetScan 和 miRDB 2 种方法预测 miR - 1749 的靶基因,并用基因功能注释经典数据库 DAVID 对该 miRNA 的靶基因进行 GO 功能注释和 KEGG 信号转导通路分析。miR - 1749 的靶基因显著富集于 Ras 蛋白信号转导调控、小 G 蛋白介导信号转导调控、

细胞稳态及胚后发育的生物学过程,GnRH、MAPK 及细胞黏附分子的信号转导通路。其中 Ras 蛋白与小 G 蛋白属于同一家族,在细胞的生长过程中起着重要作用,Ras 在细胞增殖分化信号从激活的跨膜受体传递到下游蛋白激酶的过程中处于中枢地位^[15];小 G 蛋白具有 GTP 酶活性,参与调控多种生命活动^[16]。

miRNA 在细胞中的功能研究通常采取过表达或抑制其表达的方法。过表达有 2 种方式:一种是采用 miRNA 的模拟物,该方法使用方便、效果稳定,但其价格昂贵,且不能在活体动物中研究其表达和基因的功能^[17-18];另一种是构建真核表达载体,该

方法相对耗费低、可以在细胞及动物个体中稳定表达,有利于进行 miRNA 功能的长期研究^[19]。本研究从鸡基因组 DNA 中扩增出 miR-1749 的前体序列,将其克隆到真核表达载体上,成功构建了 pcDNA3.1-EGFP-miR-1749-C/T 真核表达载体,且能成功表达。该载体标记真核表达基因新霉素和绿色荧光蛋白,在转染细胞时,可用新霉素进行稳转细胞的筛选;同时可通过荧光显微镜观察绿色荧光监测细胞转染和基因表达情况,有利于后期试验的进行。miRNA 真核表达载体符合 miRNA 的生物合成过程和作用原理,可以很好地模拟其在体内的作用机制,能够更真实地反映 miRNA 的生物学功能。

miRNA 的多态性影响其加工过程和成熟 miRNA 的生成,进而影响其与靶基因的作用及对靶基因选择,影响基因沉默的特异性^[20]。Duan 等^[21]对 miR-125a 研究发现,miR-125a 成熟体上的突变能阻碍其初级转录本到前体的加工,使成熟 miR-125a 的生成减少,从而降低其对靶基因 Lin-28 的翻译抑制作用。本试验中生物信息学分析预测发现,rs15860000 C>T 突变改变了 miR-1749 前体二级结构稳定性,且显著影响成熟 miR-1749 的表达,暗示该突变可能通过影响前体 miR-1749 二级结构的稳定性而使其成熟 miR-1749 的表达发生变化,从而影响其对靶基因的选择与作用效率,最终可能导致动物表型等发生变化。

本研究成功构建了鸡 miR-1749 不同基因型的真核表达载体,分析了其可能参与的生物学活动,为进一步研究 miR-1749 基因的功能奠定基础。

参考文献:

- [1] Filipowicz W, Bhattacharyya S N, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? [J]. *Nature Reviews Genetics*, 2008, 9(2):102-114.
- [2] Taft R J, Pheasant M, Mattick J S. The relationship between non-protein-coding DNA and eukaryotic complexity [J]. *Bioessays*, 2007, 29(3):288-299.
- [3] Wienholds E, Plasterk R H A. MicroRNA function in animal development [J]. *FEBS Letters*, 2005, 579(26):5911-5922.
- [4] Chen J F, Mandel E M, Thomson J M, et al. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation [J]. *Nature Genetics*, 2006, 38(2):228-233.
- [5] Cimmino A, Calin G A, Fabbri M, et al. miR-15 and miR-16 induce apoptosis by targeting BCL2 [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102(39):13944-13949.
- [6] Borel C, Antonarakis S E. Functional genetic variation of human miRNAs and phenotypic consequences [J]. *Mammalian Genome*, 2008, 19(7/8):503-509.
- [7] Mishra P J, Mishra P J, Banerjee D, et al. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: Introducing microRNA pharmacogenomics [J]. *Cell Cycle*, 2008, 7(7):853-858.
- [8] Glazov E A, Cottee P A, Barris W C, et al. A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach [J]. *Genome Research*, 2008, 18(6):957-964.
- [9] Chen C Z, Li L, Lodish H F, et al. MicroRNAs modulate hematopoietic lineage differentiation [J]. *Science*, 2004, 303(5654):83-86.
- [10] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction [J]. *Nucleic Acids Res*, 2003, 31(13):3406-3415.
- [11] Schmittgen T D, Livak K J. Analyzing real-time PCR data by the comparative CT method [J]. *Nature Protocols*, 2008, 3:1101-1108.
- [12] Friedman R C, Farh K K H, Burge C B, et al. Most mammalian mRNAs are conserved targets of microRNAs [J]. *Genome Research*, 2009, 19(1):92-105.
- [13] Wang X. miRDB: A microRNA target prediction and functional annotation database with a wiki interface [J]. *RNA*, 2008, 14(6):1012-1017.
- [14] Huang D W, Sherman B T, Lempicki R A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources [J]. *Nature Protocols*, 2008, 4(1):44-57.
- [15] Schaich M, Illmer T. *Mdr1* gene expression and mutations in Ras proto-oncogenes in acute myeloid leukemia [J]. *Leukemia & Lymphoma*, 2002, 43(7):1345-1354.
- [16] Haubruck H, Prange R, Vorgias C, et al. The ras-related mouse ypt1 protein can functionally replace the *YPT1* gene product in yeast [J]. *The EMBO Journal*, 1989, 8(5):1427-1432.
- [17] Hao L, Patel P C, Alhasan A H, et al. Nucleic acid-gold nanoparticle conjugates as mimics of microRNA [J]. *Small*, 2011, 7(22):3158-3162.
- [18] Younger S T, Corey D R. Transcriptional gene silencing in mammalian cells by miRNA mimics that target gene promoters [J]. *Nucleic Acids Research*, 2011, 39(13):5682-5691.
- [19] Attar M, Arefian E, Nabuini M, et al. MicroRNA 17-92 expressed by a transposon-based vector changes expression level of cell-cycle-related genes [J]. *Cell Biology International*, 2012, 36(11):1005-1012.
- [20] Duan S, Mi S, Zhang W, et al. Comprehensive analysis of the impact of SNPs and CNVs on human microRNAs and their regulatory genes [J]. *RNA Biol*, 2009, 6(4):412-425.
- [21] Duan R, Pak C H, Jin P. Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA [J]. *Human Molecular Genetics*, 2007, 16(9):1124-1131.