

陶 虎,索效军,李晓峰,等. 山羊 miR-27a 靶基因预测及生物信息学分析[J]. 江苏农业科学,2016,44(10):73-75.  
doi:10.15889/j.issn.1002-1302.2016.10.016

# 山羊 miR-27a 靶基因预测及生物信息学分析

陶 虎,索效军,李晓峰,熊 琪,张 年,杨前平,刘 洋,陈明新

(湖北省农业科学院畜牧兽医研究所/动物胚胎工程及分子育种湖北省重点实验室,湖北武汉 430064)

**摘要:**繁殖力是决定羊养殖产业效益最重要的因素,卵泡的生长发育决定了排卵数,进而直接影响母羊的繁殖力。miRNA 是由内源性基因编码的单链非编码 RNA,参与靶基因的转录后调控。本研究通过对山羊 miR-27a 的靶基因进行预测及功能分析,为进一步研究其在山羊繁殖功能中的作用提供新途径。利用靶基因预测软件分析 miR-27a 的靶基因,将靶基因进行 GO 富集分析及 KEGG 信号通路分析。结果表明,预测靶基因集合分别富集在细胞增殖、转录调控、调控 RNA 代谢、正调控信号传导等分子功能上以及 ErbB、MAPK、mTOR 等信号通路中。

**关键词:**山羊;miR-27a;靶基因;信号通路;GO 分析

**中图分类号:**S827.2 **文献标志码:**A **文章编号:**1002-1302(2016)10-0073-03

我国是养羊大国,羊养殖产业在畜牧业中占有重要地位。繁殖和生长是决定规模化养殖产出的两大主要经济性状。在山羊中,母羊繁殖效益在总效益中占据较大比例,因此繁殖力是决定羊养殖产业效益最重要的因素之一。排卵率是产仔数性状的重要组分性状,卵巢中卵泡的生长发育决定了排卵数<sup>[1]</sup>。探索卵巢中的基因表达情况、发育规律、调控因素,对于揭示排卵率的遗传调控机制具有重要意义。

microRNA(miRNA)是由内源性基因编码的单链非编码 RNA,长度一般为 21~25 个核苷酸(nucleotide,nt),通过与靶基因 mRNA 的 3'-UTR 互补配对来调控其转录后表达水平<sup>[2]</sup>。已有研究表明,miR-21、miR-212、miR-132 在小鼠

颗粒细胞中呈显著的上调表达,且当 miR-21 表达量降低时会导致颗粒细胞凋亡<sup>[3]</sup>。在小鼠颗粒细胞中筛选出多个受 TGF- $\beta$ 1 信号调控的 miRNA,其中 miR-224 可靶向作用于 Smad4 基因来调控颗粒细胞的增殖<sup>[4]</sup>。前体 miRNA 发夹结构经 Dicer 加工形成成熟的 miRNA<sup>[5]</sup>。Dicer1<sup>-/-</sup> 雌鼠的卵母细胞中纺锤体发生异常,导致其不能完成第 1 次减数分裂<sup>[6-7]</sup>。对于颗粒细胞而言,miRNA 是一类发挥重要调控作用的小分子调控物。miR-27a 是 miR-27 家族的重要成员,该家族是 miRNA 中功能较为明显的家族之一。众多研究表明,miR-27a 在疾病尤其在肿瘤中发挥重要作用,而在生殖领域鲜见报道。研究 miR-27a 及其靶基因转录调控,对进一步揭示 miRNA 对山羊母性繁殖的调控作用具有重要意义。

## 1 材料与方法

### 1.1 获取 miRNA

从 miRNA 信息公共数据库 miRBase (<http://www.mirbase.org/>) 中获取不同物种的 miR-27a 前体序列和成熟序列。

### 1.2 同源性比对

利用 miR-27a 的前体序列,比较前体序列在不同物种

1451-1452.

[18]Suyama M,Torrents D,Bork P. PAL2NAL:robust conversion of protein sequence alignments into the corresponding codon alignments [J]. Nucleic Acids Research,2006,34:609-612.

[19]Yang Z. PAM4:Phylogenetic analysis by maximum likelihood[J]. Molecular Biology and Evolution,2007,24(8):1586-1591.

[20]朱红霞,胡利宗,邓小莉,等. 三种豆科植物 DGAT1 基因家族的分子特征与进化分析[J]. 生物技术通报,2011,3(10):163-166.

[21]Cao H. Structure-function analysis of diacylglycerol acyltransferase sequences from 70 organisms [J]. BMC Research Notes,2011,4(1):249.

[22]陈振玺,王鹏,盖江涛,等. 在基因家族背景下对四种植物中 DGAT2 的鉴定和序列分析[J]. 植物科学学报,2015,33(2):186-194.

收稿日期:2016-04-08

基金项目:国家自然科学基金(编号:31501932);湖北省农业科学院青年科学基金(编号:2015NKYJJ14);湖北省农业科学院竞争性计划(编号:2016jzxj030)。

作者简介:陶 虎(1986—),男,安徽合肥人,博士,助理研究员,主要从事草食家畜分子育种研究。Tel:(027)87380139;E-mail:taohu1986@hotmail.com。

通信作者:陈明新,研究员,主要从事草食家畜育种研究。E-mail:316824556@qq.com。

[13]Bailey T L,Williams N,Misleh C,et al. MEME:discovering and analyzing DNA and protein sequence motifs [J]. Nucleic Acids Research,2006,34:369-373.

[14]Tusnady G E,Simon I. The HMMTOP transmembrane topology prediction server[J]. Bioinformatics,2001,17(9):849-850.

[15]Thompson J D,Gibson T J,Plewniak F. The CLUSTAL\_X windows interface:flexible strategies for multiple sequence alignment aided by quality analysis tools[J]. Nucleic Acids Research,1997,25(24):4876-4882.

[16]Tamura K,Dudley J,Nei M,et al. MEGA4:molecular evolutionary genetics analysis (MEGA) software version 4.0 [J]. Molecular Biology and Evolution,2007,24(8):1596-1599.

[17]Librado P,Rozas J. DnaSP v5:a software for comprehensive analysis of DNA polymorphism data [J]. Bioinformatics,2009,25(11):

间的保守程度。利用 Clustal X 软件进行序列比对,运用 Clustal Omega( <http://www.ebi.ac.uk/Tools/msa/clustalo/>) 软件制作不同物种间 miR-27a 前体序列的进化树。

1.3 靶基因预测

分别利用 miRanda( <http://www.microrna.org/microrna/home.do>)、miRDB( <http://mirdb.org/miRDB/index.html>)、TargetScan( <http://www.targetscan.org/>)等靶位点分析软件分析 miR-27a 的靶基因和靶位点的保守性。将结果的交集作为进一步研究的基因集合,筛选 3 个软件均能预测到且在 4 种物种中均保守的靶基因作为候选靶基因。通过分析靶基因的功能进一步分析 miR-27a 的生物学功能。

1.4 靶基因的分析

利用 Gene Ontology( <http://geneontology.org/>)、DAVID( <https://david.ncifcrf.gov/home.jsp>)数据库对预测得到的 miR-27a 靶基因进行功能富集分析,利用 KEGG( <http://www.kegg.jp/>)数据库进行信号通路分析,预测结果在 GO 类别上的高频率注释以及有统计学意义的信号调控通路。

2 结果与分析

2.1 miR-27a 序列的获得和同源性比对

利用 miRBase 数据库检索到 miR-27a 的成熟序列,利用 Clustal X 在线软件进行序列比对,发现该序列在人、小鼠、大鼠、猪、马中完全一致,在牛、山羊、狗中只相差最后 1 个碱基,因此 miR-27a 的成熟序列具有很高的保守性(图 1)。利用 Clustal Omega 软件分析上述物种的 miR-27a 前体序列,构建进化树。结果发现,小鼠、大鼠、猪的 miR-27a 前体序列比较接近,牛与山羊的比较接近(图 2)。

2.2 miR-27a 靶基因预测

ADORA2B PLK2 TRIM23 NAP1L2 ACVR1C GCC2 AQP11 FBXW7 SLC6A1 RNF139 EYA4 XKRX EN2 SEMA7A PARD6B HOXB8 AFF4 DCUN1D4 DNAJC13 PDS5B ARFGEF1 USP42 NIPAL4 NRK STK39 FAM193B SNAP25 PPP1CC NETO1 CDS1 HOXA5 CKAP4 PPARG EYA1 ARHGEF26 TSC1 KCNK2 TNRC18 PSM1A1 WSB1 B4GALT3 ITSN2 MMD PAQR9 AKIRIN1 PELL2 GRIK3 CCNK CTH NEK6 NR2F6 PAX9 LONRF1 APPBP2 GRIA4 GLRA2 APAF1 CEP135 CHKA B3GNT7 RBBP6 EDEM3 SFRP1 CREB1 PPP1R3F RPGRIP1L PALM2 FAM134C TMBIM6 DCX POGLUT1 ING5 SHC4 IRF4 BBX ZMAT3 ELFN2 NCOA7 NR5A2 ZHX1 DKK2 SLC35F1 HOXA10 HEG1 MAL2 BDPI DNAJC27 PHB PDIA5 CACNA2D3 CDH11 FZD7 CCNG1 TRAPPC8 LIMK1 LPAR6 CD2AP GOSR2 RSAD2 TAB3 GATC NR2F2 CACNA1A PLXND1 ARHGAP1 MTSSIL PDE7B RARA SUGLG2 MATN3 RAPGEF2 FAM184A SPRY2 NEURL4 RGS1 DOTIL SLC7A11 GRM5 CCNJ EGFR GPD1 GNS RSNBL1 ZBTB42 RPS6KA5 NPEPPS ABCA1 PNKD MAG13 CREM GFPT2 ABL2 LYPD3 PDHX RREB1 NRP2 FAM133B SMAD9 FLRT3 TADA1 TET1 MAP3K14 CNOT7 EEPD1 GATA2 UNKL BMI1 CSRP2 ADD1 HCN4 ISL1 CYP39A1 SPTY2D1 MEPCE GLIPR2 LIFR VPS26B ZCCHC24 TMEM170B MFSD2A HLX GFPT1 CSF1 RNF144A GLTP FAM65B HBEGF CBFA2T3 NKAIN1 BRSK1 DNAJC3 CDK18 TLK2 POM121 CSDC2 FAM78A NLN NRBF2 FRS3 FOXP2 FAM126B ALG9 CSRNPI MBTD1 ARMC8 PLEKHH1 BEND4 LDLR RET ADAMTS10 FOXPA4 RBPMS2 TMCC1 KLHL31 GALNT7 NFE2L2 SMOC2 AK2 NR1D2 HAPLN1 EIF2S2 VIP DLL4 NXT2 SLC39A11 EBF3 KPNA3 SLITRK1 SATB2 MAPK14 FGF1 MYEF2 E2F7 GPR126 OAF TPPP HSP90AA1 BAG2 PDGFRA ST14 PK1A CCM2 HNRNPF GXYLT1 LBH OPA1 TMUB1 ELMO1 STAG1 AMMECRIL SOS1 SEMA6D GALNT3 PHLPP2 MED14 MDF1 FOXO1 NOVA1 FAM102A SSI18L1 DTNA ATP11C SIK1 UBE2N BRPF3 UBE4A UBE2V1 SEC24A EIF5 OTX2 KHLH4 GRIA3 UBR5 ARX VAV3 GPAM TTC39A NCOA5 PPARA FAM84B ST6GALNAC3 SYDE1 CLCN5 MAN2A1 ONECUT2 GSPT1 PANK1 ZFH3 NEURL1B SIX1 SEMA6A NRARP UNC13C PDPK1 SCAF11 EHD3 LITAF CABP1 PDE3B PDE3A SLC22A23 ADCY6 RCAN2 STX16 THRB ATP2B1 MEIS2 YWHAB EML1 VEGFC NGFR TNRC6A MKNK2 MAML2 CDC25B HIC1 PAK6 FLRT2 SGPP1 SLC6A6 C2CD2 MFHAS1 MAP2K7 UBE2F NFX1 AGRN ZFH3A PAPP4 CBF6 CABLES2 ITGA5 SRSF1 NPTX2 ZFP36L1 TRIL H3F3B DCP2 SLC35F3 ARRDC4 NDUFS4 RGL2 PDE10A KBTBD8 PGAP1 FAM13A INPP5J MARCKS JMJD1C PPAP2B FOSB MYT1 TGFBR1 NOL4 SGMS1 HOXC6 PCNX LSM12 UNC80 SOCS6 RAB11FIP1 ZFP36L2 NGFRAP1 ELL2 DCAF7 CPEB3 MIER3 KCNN3 RELN CECR6 CACNB2 GAB1 CHD6 NAV2 TGFB3 RMND5A RORA TMEM9B FRYL BCORL1 NEO1 DYNLL2 TROVE2 RPS6KB1 ZBTB34 EDAR NAA50 FBXO33 SAMD10 NAA15 UBE2W MED12L ADAM19 TNPO1 ANKIB1 LOX NRXN1 SH3RF1 HYOU1 MED13 RICTOR SNX18 NFAT5 ZBTB39 ARID1B E2F6 USP46 LARP4 KCMF1 CCNY RNGTT

人 UUCACAGUGGCUAAGUUCCGC  
小鼠 UUCACAGUGGCUAAGUUCCGC  
大鼠 UUCACAGUGGCUAAGUUCCGC  
猪 UUCACAGUGGCUAAGUUCCGC  
马 UUCACAGUGGCUAAGUUCCGC  
牛 UUCACAGUGGCUAAGUUCCG-  
山羊 UUCACAGUGGCUAAGUUCCG-  
狗 UUCACAGUGGCUAAGUUCCG-  
\*\*\*\*\*

图1 miR-27a 在不同物种间的保守性

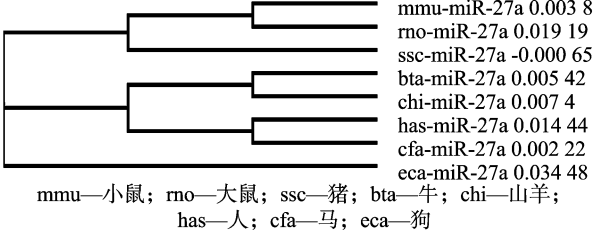


图2 miR-27a 在不同物种中的进化关系

利用 miRNA 靶基因预测软件 miRanda、miRDB、TargetScan(表 1)进行靶基因联合预测,分别得到 miR-27a 靶基因为 8 902、697、786 个。筛选 3 种软件预测结果的交集(共 390 个),作为进一步分析的基因集合(图 3)。

表 1 miRNA 靶基因预测软件

软件名称	预测依据	网址
TargetScan	种子区互补性	<a href="http://www.targetscan.org/">http://www.targetscan.org/</a>
miRDB	miRNA-靶标 复合体热力学 稳定性	<a href="http://www.mirdb.org/miRDB/">www.mirdb.org/miRDB/</a>
miRanda	互补性	<a href="http://www.microrna.org/microrna/home.do">http://www.microrna.org/microrna/home.do</a>

图3 miR-27a 靶基因交集

2.3 miR-27a 靶基因 GO 分析和 Pathway 分析

将预测得到的 390 个靶基因进行 GO 功能富集分析( $P < 0.05$ ),结果表明,靶基因富集在细胞增殖、转录调控、调控 RNA 代谢、调控信号传导、细胞分化、染色质修饰和组织、基因表达调控、转运、神经发育和分化等生物学过程和功能上

(表 2)。利用 DAVID 在线数据库对 miR-27a 的靶基因进行 Pathway 伏击分析( $P < 0.05$ ),结果表明,miR-27a 的靶基因富集于 ErbB 信号通路、MAPK 信号通路、Insulin 信号通路、Neurotrophin 信号通路、前列腺癌信号通路、黏着斑信号通路、mTOR 信号通路、T 细胞受体信号通路(表 3)。

表 2 miR-27a 靶基因的 GO 富集分析结果

GO 号	GO 注释	P 值	基因名称
0042127	Regulation of cell proliferation	0.033	VIP、ING5、BMII、EGFR、PDSSB、TGFBRI、CSF1、PHB 等
0045449	Regulation of transcription	$2.38 \times 10^{-6}$	BMII、PPARA、ZBTB34、THRB、E2F6、E2F7、PPARG 等
0051252	Regulation of RNA metabolic process	$2.28 \times 10^{-5}$	PPARA、THRB、E2F6、E2F7、PPARG、NAA15、FOXO1 等
0045597	Positive regulation of cell differentiation	0.035	VEGFC、SMAD9、THRB、LIMK1、MAPK14、HLX、CSF1 等
0016568	Chromatin modification	$5.01 \times 10^{-4}$	ING5、BMII、SATB2、PHB、TADA1、ARID1B、DOT1L 等
0006325	Chromatin organization	0.002 4	ING5、BMII、SATB2、PHB、TADA1、ARID1B、NAP1L2 等
0010628	Positive regulation of gene expression	$2.82 \times 10^{-6}$	ING5、PPARA、THRB、CSF1、PPARG、ONECUT2 等
0010629	Negative regulation of gene expression	0.037	MDF1、PPARA、SATB2、THRB、E2F6、PHB、ZHX1 等
0006350	Transcription	$2.18 \times 10^{-4}$	BMII、PPARA、ZBTB34、THRB、E2F6、E2F7、PPARG 等
0048666	Neuron development	$8.94 \times 10^{-6}$	EGFR、NRP2、PARD6B、RET、CREB1、ONECUT2 等
0030182	Neuron differentiation	$6.84 \times 10^{-6}$	NRP2、ONECUT2、BRSK1、RORA、ARX、GATA2 等
0030528	Transcription regulator activity	$8.70 \times 10^{-8}$	PPARA、THRB、E2F6、E2F7、PPARG、FOXO1、RORA 等

表 3 miR-27a 预测靶基因的 Pathway 分析结果

信号通路	P 值	基因名称
ErbB signaling pathway	0.001	EGFR、PAK6、SOS1、GAB1、HBEGF、RPS6KB1、MAP2K7、ABL2、SHC4
MAPK signaling pathway	0.002	EGFR、TGFBRI、MKNK2、CACNB2、CACNA2D3、CDC25B、ACVR1C、RPS6KA5、MAPK14、SOS1、PDGFRA、FGF1、RAPGEF2、MAP3K14、MAP2K7、CACNA1A
Insulin signaling pathway	0.006	PDPK1、TSC1、SOS1、MKNK2、PDE3B、FOXO1、PDE3A、RPS6KB1、PPP1CC、SHC4
Neurotrophin signaling pathway	0.011	RPS6KA5、MAPK14、SOS1、GAB1、YWHAB、NGFRAP1、NGFR、MAP2K7、SHC4
Prostate cancer	0.022	EGFR、PDPK1、HSP90AA1、CREB1、SOS1、PDGFRA、FOXO1
Focal adhesion	0.027	EGFR、PAK6、VEGFC、PDPK1、VAV3、ITGA5、SOS1、PDGFRA、RELN、PPP1CC、SHC4
mTOR signaling pathway	0.039	VEGFC、PDPK1、TSC1、RPS6KB1、RICTOR
T cell receptor signaling pathway	0.050	PAK6、VAV3、MAPK14、SOS1、NFAT5、MAP3K14、MAP2K7

### 3 结论与讨论

miRNA 是一种非编码的小 RNA,通过与靶基因 3'-UTR 特异性结合,从而抑制转录后基因的表达,在调控细胞周期、基因表达、生殖发育等过程中具有重要作用<sup>[8]</sup>。miRBase 数据库至今已收录的 miRNA 达 24 521 条,为本研究奠定了基础。

目前,miRNA 的研究已取得了很大进展,对于各种生命活动的调控作用也进行了大量研究。GO 富集分析、KEGG 信号通路分析等生物信息学分析技术的出现对 miRNA 的研究起到了关键性的支持作用,帮助研究者从海量的 miRNA 数据中筛选出有价值的信息。Gene Ontology 可分为分子功能、生物过程、细胞组成 3 个部分。蛋白质或基因可通过 ID 对应或序列注释的方法找到与之对应的 GO 号,从而进行功能类别或细胞定位<sup>[9]</sup>。KEGG 信号通路分析能够将特定 miRNA 靶基因的信号通路进行整合归类,也可设定相关参数,根据实际需要需要对结果进行提取<sup>[10]</sup>。

miR-27a 研究的常规思路是通过靶基因预测网站(miRGen、TargetScan 等),利用已知的 miRNA 来预测下游靶基因。利用上述预测手段得到可能与之作用的 miRNA,通过合成 miRNA 的类似物和抑制物,与所构建靶基因的荧光载体共转染细胞,确定能够被 miR-27a 显著下调的靶基因,从而完成验证。

miR-27a 的靶基因参与了细胞增殖、转录调控、调控 RNA 代谢、正调控信号传导等一系列重要的生命活动过程,预示 miR-27a 在山羊繁殖过程中可能发挥着重要的作用。

### 参考文献:

- [1] 张春艳. 山羊繁殖性状的影响因素和遗传规律及分子调控机制研究[D]. 武汉:华中农业大学,2010.
- [2] Gu S, Jin L, Zhang F, et al. Biological basis for restriction of microRNA targets to the 3'untranslated region in mammalian mRNAs[J]. Nat Struct Mol Biol, 2009, 16(2): 144-150.
- [3] Carletti M Z, Fiedler S D, Christenson L K. MicroRNA 21 blocks apoptosis in mouse periovulatory granulosa cells[J]. Biol Reprod, 2010, 83(2): 286-295.
- [4] Yao G, Yin M, Lian J, et al. MicroRNA-224 is involved in transforming growth factor-beta-mediated mouse granulosa cell proliferation and granulosa cell function by targeting Smad4[J]. Mol Endocrinol, 2010, 24(3): 540-551.
- [5] Park J E, Heo I, Tian Y, et al. Dicer recognizes the 5' end of RNA for efficient and accurate processing[J]. Nature, 2011, 475(7355): 201-205.
- [6] Murchison E P, Stein P, Xuan Z, et al. Critical roles for Dicer in the female germline[J]. Genes Dev, 2007, 21(6): 682-693.
- [7] Tang F, Kaneda M, O'Carroll D, et al. Maternal microRNAs are essential for mouse zygotic development[J]. Genes Dev, 2007, 21(6): 644-648.
- [8] Ambros V. The functions of animal microRNAs[J]. Nature, 2004, 431: 350-355.
- [9] 崔晓钢. 基于 RNA-seq 与 small RNA-seq 进行奶牛产奶性状功能基因挖掘及生物信息学预测牛新 miRNA[D]. 北京:中国农业大学,2015.
- [10] 杨洋, 阚清, 张盼, 等. miRNA-126 靶基因预测及其相关信号通路的生物信息学分析[J]. 中国当代儿科杂志, 2013, 15(3): 227-232.