

张晓琼, 乔琳, 胡利宗, 等. 棉花与模式植物 *DGAT2* 基因的鉴定与分析[J]. 江苏农业科学, 2016, 44(10): 69–73.  
doi:10.15889/j.issn.1002-1302.2016.10.015

# 棉花与模式植物 *DGAT2* 基因的鉴定与分析

张晓琼, 乔琳, 胡利宗, 廖立冰, 王秋霞, 解恒昌, 周琳

(周口师范学院生命科学与农学院, 河南周口 466001)

**摘要:** 二脂酰甘油酰基转移酶 II (*DGAT2*) 是催化三酰甘油生物合成的关键酶, 利用生物信息学手段在二倍体棉花和 6 种模式植物中共鉴定出 25 个 *DGAT2* 基因, 详细剖析这些基因的结构、序列和进化特征。基因结构分析表明, 绝大多数植物包含 8 个或 9 个外显子, 且内含子相位高度保守, 说明该基因结构起源于陆生植物。蛋白序列分析显示, 核心保守基序 1、2、3 与功能结构域相互重叠, 说明功能结构域起源于藻类植物。进化分析显示: 在棉花中, *DGAT2* 基因发生特异性扩增, 其分子机制为串联重复; 结合滑动窗口方法和位点模型选择压力检测结果, *DGAT2* 蛋白均受制于负选择作用。这些结果为棉花及模式植物 *DGAT2* 的功能研究提供帮助。

**关键词:** 二酰甘油酰基转移酶; 保守基序; 棉花; 模式植物; 进化

**中图分类号:** S562.03; Q78 **文献标志码:** A **文章编号:** 1002-1302(2016)10-0069-04

二酰甘油酰基转移酶(*DGAT*, EC 2.3.1.20)是催化三酰甘油(*TAG*)合成最后一步的关键酶<sup>[1]</sup>。目前, 在植物中存在 3 种类型的 *DGAT*, 即 *DGAT1*、*DGAT2* 和 *DGAT3*<sup>[2]</sup>。前人研究表明, *DGAT1*、*DGAT2* 和 *DGAT3* 隶属于不同基因家族, 其蛋白序列相似性极低, 但三者均具有二脂酰甘油酰基转移酶活性, 过表达这些基因均可不同程度提高植物种子含油量<sup>[3]</sup>。前人研究显示, *DGAT* 酶参与种子油脂合成、种子萌发与发育以及幼苗发育等众多生物学过程, 具有多样化的功能<sup>[4]</sup>。自 Lardizabal 等成功克隆 *MrDGAT2* 基因以来, 在人类、拟南芥和藻类等物种中均克隆出 *DGAT2* 基因<sup>[5-7]</sup>。*DGAT2* 基因功能研究显示, *DGAT2* 基因具有底物特异性, 优先利用特殊脂肪酸进行三酰甘油合成<sup>[8]</sup>。因此, *DGAT2* 基因的相关研究越来越受到人们的重视。棉花不仅是一种重要的经济作物, 还是天然纤维与油料作物。虽然在棉花纤维发育、纤维品质以及抗逆性等方面开展了许多研究, 但棉花 *DGAT2* 基因的研究仍未见报道。此外, 二倍体雷蒙德氏棉<sup>[9]</sup>和亚洲棉<sup>[10]</sup>全基因组序列测定已经被完成, 这为在棉花中研究 *DGAT2* 基因提供了数据资源。利用生物信息学手段鉴定棉花 *DGAT2* 基因, 并剖析这些基因的结构、保守基序、功能结构域以及系统进化关系。这些结果能为阐明棉花油脂积累的分子机制奠定基础, 为进一步研究棉花 *DGAT2* 基因的功能提供线索。

## 1 材料与方法

### 1.1 物种抽样与基因序列下载

本研究涉及的物种包括亚洲棉、雷蒙德氏棉和其他 6 种

模式植物。序列下载步骤为: 在 NCBI 数据库直接检索并下载已知的 *DGAT2* 基因序列, 以这些序列为检索序列, 利用 BLAST 工具在默认参数条件下分别搜索相关数据库 Cotton-Gen (<http://www.cottongen.org/>)、Greenphyl V4 (<http://www.greenphyl.org/cgi-bin/index.cgi>) 和 Phytozome V9.1 (<http://www.phytozome.net/>), 最终获取棉花及其他模式植物的 *DGAT2* 基因。并利用 Pfam 工具<sup>[11]</sup>确定这些基因是否含有 DAGAT(PF03982)。

### 1.2 序列特征与蛋白特性分析

通过 DNA 和 cDNA 的比较, 棉花与模式植物 *DGAT2* 基因外显子和内含子的组织结构模式被鉴定, 由 GSDS 2.0 工具绘制<sup>[12]</sup>。MEME 服务器被用于剖析棉花与模式植物 *DGAT2* 蛋白的保守基序组成模式, 参数设置如下: 保守基序氨基酸长度为 6~180 bp; 保守基序数目最大值为 10; E 阈值为 0.000 01, 其他参数默认<sup>[13]</sup>。此外, 利用 HMMTOP 工具剖析 *DGAT2* 蛋白的跨膜结构域的数目与位置<sup>[14]</sup>。

### 1.3 系统进化与选择压力检测

为阐明 *DGAT2* 蛋白的系统进化关系, 利用 Clustal X 软件<sup>[15]</sup>在默认条件下对 *DGAT2* 蛋白进行序列比对, 其比对结果直接导入 MEGA 软件<sup>[16]</sup>中, 输出进化树并进行细微调整。在进化树的末端, 可鉴定出同源基因对, 同时基于滑动窗口方法剖析同源基因对的选择压力, 其计算由 DnaSP 5.10 完成, 参数设置为: 滑动窗口 100 bp 和步长 10 bp<sup>[17]</sup>。此外, 利用 Pal2nal 工具准备所有 *DGAT2* 基因的密码子比对序列<sup>[18]</sup>, 利用位点特异模型对 *DGAT2* 基因进行选择压力分析, 其计算过程由 PAML 3.15 软件包中的 CODEML 程序完成<sup>[19]</sup>。

## 2 结果与分析

### 2.1 棉花与模式植物 *DGAT2* 鉴定与蛋白序列特征

利用 Blast 工具在 2 种棉花与 6 种模式植物中共鉴定出 25 个 *DGAT2* 基因, 这些基因的位点标识符、染色体、正负链位置以及基因组匹配区段均被列出(表 1)。功能结构域分析表明, *DGAT2* 基因编码的蛋白均具有典型的功能结构域 DAGAT

收稿日期: 2015-08-28

基金项目: 周口师范学院高层次人才科研启动经费(编号: ZKNU2014109); 周口师范学院大学生科研创新基金(编号: ZK-NUD15037)。

作者简介: 张晓琼(1991—), 女, 河南安阳人, 硕士研究生, 主要从事植物遗传学研究。E-mail: zqxiaoqiong00@163.com。

通信作者: 胡利宗, 博士, 副教授, 主要从事植物遗传学研究。E-mail: hulizong@126.com。

(PF03982)。聚类分析显示,2 种棉花 DGAT2 与拟南芥和水稻 DGAT2 聚在一起形成了一个双子叶植物类群;低等植物 DGAT2 倾向于聚为一类,但有些低等植物 DGAT2 蛋白单独

形成一个分支(图 1-A)。此外,利用 HMMTOP 工具分析 DGAT2 蛋白的跨膜结构,结果显示,除 PaDGAT2c 外,棉花与模式植物 DGAT2 蛋白均具有至少 1 个跨膜结构域(表 1)。

表 1 棉花与模式植物 *DGAT2* 基因的基本信息

基因名称	登录号	染色体:匹配区域(正负链)	读码框	跨膜数目
<i>GrDGAT2a</i>	Gorai. 007G064100	Chr07:4489790 – 4493013(1)	1005	3
<i>GrDGAT2b</i>	Gorai. 007G064200	Chr07:4501368 – 4505139(1)	993	2
<i>GrDGAT2c</i>	Gorai. 007G064300	Chr07:4507167 – 4509730(1)	984	2
<i>GrDGAT2d</i>	Gorai. 007G064500	Chr07:4519514 – 4523214(1)	984	2
<i>GaDGAT2a</i>	Cotton_A_10913	Chr6:56565901..56569083(–1)	906	1
<i>GaDGAT2b</i>	Cotton_A_10910	Chr6:56591542..56595248(–1)	993	2
<i>GaDGAT2c</i>	Cotton_A_10911	Chr6:56585693..56588139(–1)	1047	2
<i>GaDGAT2d</i>	Cotton_A_10912	Chr6:56577324..56579452(–1)	909	1
<i>AtDGAT2</i>	AT3G51520	Chr3:19110596 – 19112760(1)	945	2
<i>OsDGAT2a</i>	Os02g48350	Chr2:29598974 – 29601846(–1)	1017	1
<i>OsDGAT2b</i>	Os06g22080	Chr6:12809650 – 12812495(1)	1113	2
<i>PaDGAT2a</i>	MA_29382g0010	MA_29382:36266 – 42616	801	2
<i>PaDGAT2b</i>	MA_481978g0010	MA_481978:866 – 6178	591	2
<i>PaDGAT2c</i>	MA_9132531g0010	MA_9132531:1 – 3383	291	0
<i>PaDGAT2d</i>	MA_123860g0010	MA_123860:3803 – 4189	285	2
<i>SmDGAT2</i>	96204	Scf18:605329 – 606881(–1)	1023	2
<i>PpDGAT2a</i>	Phpat. 014G077300	Chr14:13165760 – 13169482(1)	942	2
<i>PpDGAT2b</i>	Phpat. 019G038400	Chr19:7016170 – 7018556(1)	1095	2
<i>PpDGAT2c</i>	Phpat. 021G046600	Chr21:8495995 – 8499080(1)	1092	2
<i>PpDGAT2d</i>	Phpat. 022G071500	Chr22:14003533 – 14006959(1)	930	2
<i>CrDGAT2a</i>	Cre03. g205050	Chr3:7551370 – 7554956(1)	984	2
<i>CrDGAT2b</i>	Cre06. g299050	Chr6:7351329 – 7354799(1)	1041	3
<i>CrDGAT2c</i>	Cre08. g377300	Chr8:3790075 – 3798402(–1)	3660	6
<i>CrDGAT2d</i>	Cre09. g386912	Chr9:3127456 – 3133016(1)	975	2
<i>CrDGAT2e</i>	Cre12. g557750	Chr12:7279129 – 7283104(–1)	996	3

利用 MEME 软件对 DGAT2 蛋白的保守基序进行分析,结果表明,在 25 个棉花和模式植物 DGAT2 蛋白中存在 10 个保守基序(Motif1 ~ 10),共形成 13 种保守基序组织模式,其中 6 个棉花 DGAT2 蛋白共享的保守基序组成模式为 Motif4 – 6 – 2 – 1 – 3,而 *AtDGAT2*、*OsDGAT2a/b* 和 *PpDGAT2d* 等共享的保守基序模式为 Motif9 – 6 – 2 – 1 – 3(图 1-B)。就单个保守基序而言,所有 DGAT2 蛋白保守基序被分为两大类型,即种属普遍类型和种属特异类型。保守基序 1、2、3、6、9 几乎存在于所有 DGAT2 蛋白中,属于种属普遍类型;但保守基序 4、5、7、8、10 表现出不同程度的种属特异性,属于种属特异类型。此外, DGAT2 蛋白的 DAGAT (PF03982) 功能结构域与保守基序 1、2、3、6、9 相互重叠,这说明植物的 5 个保守基序均可能与 DGAT2 活性高度相关。

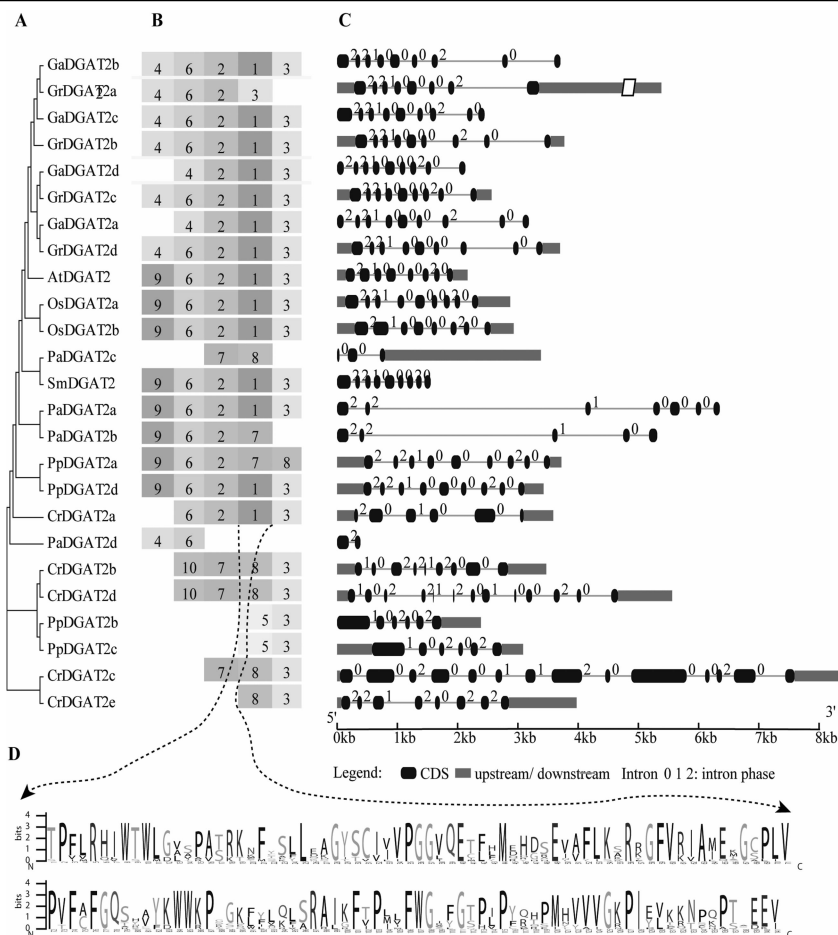
2.2 棉花与模式植物 *DGAT2* 基因的外显子与内含子组成模式

通过比较 DNA 与 CDS 序列,棉花与模式植物 *DGAT2* 基因的结构特点均被阐明,并鉴定其内含子相位。根据外显子数目,这些 *DGAT2* 基因结构被分为三大类型:(1)8 个外显子类型,这种类型的基因包括 *GrDGAT2a*、*GrDGAT2d* 和 *AtDGAT2* 等 5 个基因;(2)9 个外显子类型,这种类型的基因包括 *GrDGAT2b*、*GaDGAT2a* 和 *OsDGAT2a* 等 11 个基因;(3)2、3、5、6、7、13 个或 14 个外显子类型,这类基因包括 *CrDGAT2a*、*PpDGAT2b* 和 *PaDGAT2c* 等 9 个基因(图 1-C)。结合 DGAT2

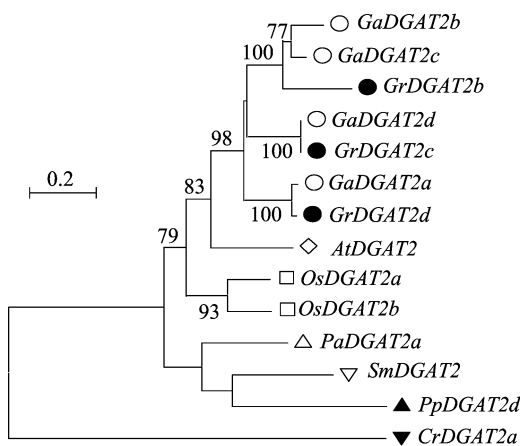
蛋白聚类结果可知,位于同一进化分支的基因往往具有相同或相似的基因结构,甚至外显子和内含子长度、内含子相位具有高度保守性;棉花 DGAT2 蛋白聚为一类,具有高度相似的基因结构,并且这些基因的内含子相位高度保守;此外,有些物种 DGAT2 蛋白单独形成一个分支,具有独特的基因结构如 *PaDGAT2d*、*CrDGAT2a* 和 *PaDGAT2b*,它们分别拥有 2、6、5 个外显子,并且其内含子相位变异较大。总之,棉花与模式植物 *DGAT2* 基因的结构多数以 8 个或 9 个外显子形式存在,尤其是在高等植物中该基因的结构具有较高的保守性。

2.3 棉花与模式植物 *DGAT2* 基因的系统进化关系

在构建棉花与模式植物 *DGAT2* 基因的进化树前,移走覆盖度低于 80%、相似度低于 30% 的序列,进而构建由亚洲棉、雷蒙德氏棉和水稻等 8 种植物的 14 个 DGAT2 蛋白形成的进化树。系统进化分析表明,藻类植物 *CrDGAT2a* 蛋白单独形成 1 个进化分支,与其他植物具有较远的亲缘关系;其他苔藓类、蕨类、裸子与被子植物 DGAT2 蛋白明显被分为 2 个进化分支,其中被子植物 DGAT2 蛋白形成 1 个进化分支,藓类、蕨类和裸子植物 DGAT2 形成另一个进化分支(图 2)。此外,在进化树末端分支上,存在 4 对同源基因,其中 2 对属于直系同源基因,2 对属于旁系同源基因;旁系同源基因包括 *GaDGAT2b/GaDGAT2c* 和 *OsDGAT2a/OsDGAT2b*,它们起源于物种特异的基因扩增;直系同源基因包括 *GrDGAT2c/GaDGAT2d* 和 *GaDGAT2a/GrDGAT2d*,它们起源于物种分化事件(图 2)。



A—进化树；B—保守基序组成；C—基因结构；D—保守基序1的隐马尔可夫模型

图1 植物 *DGAT2* 基因的进化关系与序列特征图2 棉花与模式植物 *DGAT2* 基因的系统进化树

#### 2.4 基于同源对水平的棉花与模式植物 *DGAT2* 基因的选择压力

为考察进化树末端的 4 对同源基因之间是否经历适应性进化,本研究利用 DnaSP 5.10 软件对 4 对 *DGAT2* 同源基因进行滑动窗口分析,参数设置:窗口尺寸为 100 bp,步移长度为 10 bp。结果显示,4 对同源基因的  $\omega$  平均值从大到小为 0.44 (*GrDGAT2d*/*GaDGAT2a*) > 0.31 (*GaDGAT2b*/*GaDGAT2c*) > 0.16 (*OsDGAT2b*/*OsDGAT2a*) > 0.15 (*GaDGAT2d*/*GrDGAT2c*)。就

旁系同源基因而言,亚洲棉旁系基因对 *GaDGAT2b*/*GaDGAT2c* 和水稻旁系同源基因对 *OsDGAT2b*/*OsDGAT2a* 的每个步移窗口对应的进化速率具有异质性,但所有的步移窗口的  $\omega$  值均小于 1,这充分说明这 2 对旁系同源基因均受控于负选择作用,也揭示了基因重复之后并未发生适应性进化(图 3 - A)。同理,在 2 种棉花物种中,直系同源基因对 *GrDGAT2d*/*GaDGAT2a* 和 *GaDGAT2d*/*GrDGAT2c* 的  $\omega$  均不大于 1,这说明物种分化产生重复基因受到负选择作用,这可能与功能限制密切相关(图 3 - B)。

#### 2.5 基于类群水平的棉花与模式植物 *DGAT2* 基因的选择压力

以亚洲棉、雷蒙德氏棉和水稻等 8 种植物的 14 个 *DGAT2* 基因为研究对象,利用位点模型对 *DGAT2* 基因进行选择压力检测,鉴定出正选择位点,并利用贝叶斯后验概率推断每个正选择位点概率值。似然比检验表明,  $2\Delta\ln L = 271.62 > 13.277$  ( $M0/M3$ ),这说明  $M3$  极显著优于  $M0$  ( $P < 0.01$ ),因此,这些 *DGAT2* 蛋白位点间的进化速率明显具有异质性;就  $M2a$  和  $M1a$  比较而言,似然比检验值  $2\Delta\ln L = 0 < 5.991$ ,说明  $M2a$  并不优于  $M1a$  模型, *DGAT2* 基因可能未经历适应性进化;同理,  $2\Delta\ln L = 0 < 5.991$  ( $M7/M8$ ),说明  $M8$  模型并不优于  $M7$  模型,进而可推知棉花与模式植物 *DGAT2* 基因中并没有鉴定到正选择氨基酸位点(表 2)。结合似然比测验和贝叶斯推断分

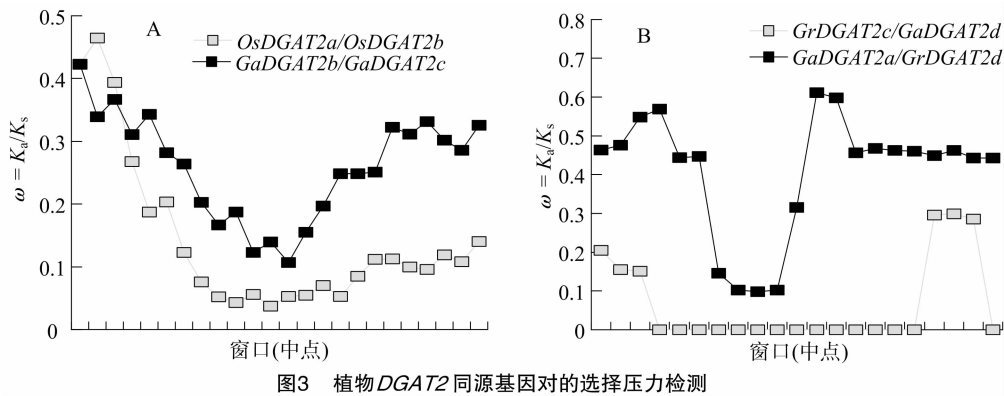


图3 植物 *DGAT2* 同源基因对的选择压力检测

表 2 基于位点模型检测植物 *DGAT2* 基因的选择压力检测

模型	自由参数数目	参数评价	lnL	似然模型对	df	2ΔlnL
M0:单比率	27	$\omega_0 = 0.16$	-5 907.99			
M3:高散型	31	$p_0 = 0.23, p_1 = 0.44, p_2 = 0.33, \omega_0 = 0.0054, \omega_1 = 0.11, \omega_2 = 0.40$	-5 761.93	M0/M3	4	271.62 **
M1a:中性	28	$p_0 = 0.78, p_1 = 0.22, \omega_0 = 0.12, \omega_1 = 1.00$	-5 839.13			
M2a:选择	30	$p_0 = 0.78, p_1 = 0.20, p_2 = 0.02, \omega_0 = 0.12, \omega_1 = 1.00, \omega_2 = 1.00$	-5 839.13	M1a/M2a	2	0
M7:β	28	$p = 0.61, q = 2.67$	-5 760.96			
M8:β&ω	30	$p = 0.61, q = 2.67, p_0 = 0.99999, p_1 = 0.00001, \omega = 1.00$	-5 760.96	M7/M8	2	0

注:lnL 表示最大似然值的对数; \*\* 表示似然比检验极显著。

析可知,棉花与模式植物 *DGAT2* 基因均受制于负选择作用,这意味着 *DGAT2* 具有高度保守性,这可能与严格的功能限制相关。

3 结论与讨论

*DGAT2* 基因广泛存在于真菌、藻类、植物和动物基因组中,尤其是在藻类植物中通常以多拷贝形式存在<sup>[1,3,20-21]</sup>,这说明 *DGAT2* 基因起源较早,是一个古老的基因家族。与藻类植物相比,被子植物中 *DGAT2* 基因数目较少,绝大多数物种仅包括 1 个或 2 个成员。但令人感兴趣的是,在 2 种二倍体棉花中均发现了 4 个 *DGAT2* 基因,并且在 2 种二倍体棉花中 *DGAT2* 基因均以串联重复方式排列在同一个染色体区段内,这说明棉花 *DGAT2* 基因发生了物种特异性串联重复。基因结构和蛋白保守基序分析显示,所有物种 *DGAT2* 序列特征具有较高的多样性,但在被子植物中 *DGAT2* 基因的结构与蛋白基序均表现出高度保守性,这说明被子植物 *DGAT2* 蛋白受到严格的功能限制,具有高度相似的生物学功能,这与前人的研究结果<sup>[21-22]</sup> 基本一致。此外,水稻和棉花的 *DGAT2* 同源基因对选择压力检测显示,同源基因对分化后并没有发生适应性进化。同时,基于位点模型对 *DGAT2* 蛋白进行选择压力分析,结果表明植物的 14 个 *DGAT2* 蛋白均受控于负选择作用,受到严格的功能限制作用<sup>[21]</sup>。总之,二倍体棉花基因组中存在 4 个 *DGAT2* 基因,这些基因呈现串联重复特征,并且棉花 *DGAT2* 基因受纯净选择作用,这与其功能重要性密切相关。这些结果揭示了 *DGAT2* 基因的序列特征与进化关系,为其功能研究提供了理论基础。

参考文献:

[1] Weselake R J, Taylor D C, Rahman M H, et al. Increasing the flow of carbon into seed oil[J]. Biotechnol Advances, 2009, 27(6): 866-878.

[2] 唐桂英,柳展基,单 雷. 二酰基甘油酰基转移酶(DGAT)研究进展[J]. 中国油料作物学报, 2010, 32(2): 320-328.

[3] Turchetto-Zolet A C, Maraschin F S, de Moraes G L, et al. Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis[J]. BMC Evolutionary Biology, 2011, 11(1): 263.

[4] Chai Y, Hao X, Yang X, et al. Validation of DGAT1-2 polymorphisms associated with oil content and development of functional markers for molecular breeding of high-oil maize[J]. Molecular Breeding, 2012, 29(4): 939-949.

[5] Lardizabal K D, Mai J T, Wagner N W, et al. DGAT2 is a new diacylglycerol acyltransferase gene family purification, cloning, and expression in insect cells of two polypeptides from *Mortierella ramanniana* with diacylglycerol acyltransferase activity[J]. Journal of Biological Chemistry, 2001, 276(42): 38862-38869.

[6] 房逢立, 吴 洪, 周志刚. 缺刻缘绿藻二酰甘油酰基转移酶 2 (DGAT2) 的基因特性与功能鉴定[J]. 水产学报, 2013, 37(8): 1162-1172.

[7] 袁峥嵘, 柳小春, 马海明, 等. 二脂酰甘油酰基转移酶 2 (DGAT2) 基因研究进展[J]. 遗传, 2008, 30(3): 289-294.

[8] Li R, Yu K, Hildebrand D F. DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants[J]. Lipids, 2010, 45(2): 145-157.

[9] Paterson A H, Wendel J F, Gundlach H, et al. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres[J]. Nature, 2012, 492(7429): 423-427.

[10] Li F, Fan G, Wang K, et al. Genome sequence of the cultivated cotton *Gossypium arboreum*[J]. Nature Genetics, 2014, 46(6): 567-572.

[11] Bateman A, Birney E, Cerruti L, et al. The Pfam protein families database[J]. Nucleic Acids Research, 2002, 30(1): 276-280.

[12] 郭安源, 朱其慧, 陈 新, 等. GSDS: 基因结构显示系统[J]. 遗传, 2007, 29(8): 1023-1026.

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

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):

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. PAML4: 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.