

蓖麻 DNA 甲基转移酶序列与基因表达分析*

代梦媛, 高 梅, 郭丽芬, 邓剑川, 李文昌**

(云南省农业科学院 经济作物研究所, 云南 昆明 650205)

摘要: 植物 DNA 甲基化是与调控基因表达、保护基因组免受逆境胁迫等相关的表观遗传现象, 其中关键酶 DNA 甲基转移酶催化 DNA 序列维持甲基化和从头甲基化。本研究基于蓖麻全基因组、利用生物信息学分析方法、结合转录组数据和 RT-PCR 半定量表达分析, 鉴别出蓖麻 DNA 甲基转移酶氨基酸序列的结构特征、系统发生和表达形式。结果如下: 通过同源序列 BLAST 从蓖麻基因组数据库中获得 8 个 DNA 甲基转移酶基因; 系统进化分析将 8 个蛋白分为 4 个亚家族: 2 个 MET 同源蛋白、2 个 CMT 同源蛋白、3 个 DRM 同源蛋白和 1 个 DNMT2 同源蛋白, 进一步发现 DNA 甲基转移酶在系统进化上较为保守, 分化发生在单双子叶植物分化之后; 亚细胞定位分析发现 8 个 DNA 甲基转移酶都定位在细胞核中; 蛋白结构域比对结果显示 8 个蛋白的 C 端催化结构域都有 6 个保守 motif 参与甲基化修饰催化功能, 表明蛋白具有基本的催化甲基化功能, 而 N 端调节功能域的差异将蛋白分为四类, 其中 RcDNMT2 缺少 N 端调节结构域; 结合转录组数据和 RT-PCR 半定量表达分析发现蓖麻 DNA 甲基转移酶基因 (除 RcDRM2 外) 在发育胚乳中相对其他检测组织高表达, 推测是该时期的胚乳低甲基化诱导甲基化基因高表达。该结果有助于全面了解蓖麻 DNA 甲基转移酶特征及为从表观遗传深入开展蓖麻品种改良具有重要理论指导作用。

关键词: 蓖麻; DAN 甲基转移酶; 甲基化; 生物信息学; 基因表达

中图分类号: S 565.6 **文献标志码:** A **文章编号:** 1004-390X (2016) 04-0604-11

Sequence and Gene Expression Analysis of DNA Methyltransferase in Castor Bean (*Ricinus communis* L.)

DAI Mengyuan, GAO Mei, GUO Lifan, DENG Jianchuan, LI Wenchang

(Cash Crop Research Institute, Yunnan Academy of Agricultural Sciences, Kunming 650205, China)

Abstract: DNA methylation, which can protect plant genome from adversity stress and regulate gene expression, is one of epigenetic modification. DNA methyltransferase is key enzyme for maintaining DNA methylation and de novo methylation. In this research, 8 putative genes encoding DNA methyltransferase were identified and characterized from castor bean (*Ricinus communis*) genome. These genes have been classified into MET, CMT, DRM and DNMT2 subfamilies based on phylogenetic relationship and domain organization, and their structures, conserved domains, and subcellular locations were also characterized. In addition, the gene expression was inspected through the high-throughput transcriptional sequencing data and RT-PCR methods. Results showed these proteins encoded by 8 genes were located at nucleus, and catalytic C-terminal domain with 6 conserved motifs have common methylation function, while the N-terminal domain of these proteins showed diverse structural features

收稿日期: 2015-06-05

修回日期: 2015-08-26

网络出版时间: 2016-07-26 15:28

* 基金项目: 云南省重点新产品开发专项 (2016BB005)。

作者简介: 代梦媛 (1988—), 女, 云南玉溪人, 硕士, 研究实习员, 主要从事植物分子遗传与育种。

E-mail: mengyuandai@126.com

** 通信作者 Corresponding author: 李文昌 (1968—), 男, 云南大理人, 硕士, 研究员, 主要从事特种油料作物育种及栽培技术研究工作。E-mail: ynbioenergy@163.com

网络出版地址: <http://www.cnki.net/kcms/detail/53.1044.S.20160726.1528.004.html>

between different subfamilies, meanwhile, the lacking N-terminal domain presented in the RcDNMT2. Phylogenetically, the MET, CMT, DRM and DNMT2 were conservative and differentiated after mono- and dicot-cotyledon. The expression level of all genes in developing endosperm was higher than other detective tissues except RcDRM2, and hypothesized that endosperm hypomethylation can induce the high expression of these DNA methyltransferase genes during this period. These results contribute to understanding castor bean DNA methyltransferase comprehensively and provide a guideline for epigenetics breeding.

Keywords: *Ricinus communis*; DNA methyltransferase; methylation; bioinformatics; gene expression

DNA 甲基化是涉及多个重要发育过程的保守表观修饰现象,如 X 染色体失活、基因组印记以及保护基因组免受转座子、反转录转座子和病毒侵袭^[1],通过 DNA 甲基转移酶催化 S-腺苷-L-蛋氨酸上的甲基基团转移到 DNA 碱基位点上来实现^[2],植物中这种修饰作用主要发生在特定的序列形式 CpG、CpNpG (N 为任意核碱基)中胞嘧啶的第 5 个碳原子 (5 mC) 上^[3],植物基因组中有 30% 的 5 mC 被甲基化^[4],如拟南芥基因组中对称序列 CpG、CpNpG 和不对称序列 CpNpN 的甲基化水平分别为 24%、6.7% 和 1.7%^[5]。目前主要有 2 种 DNA 甲基化形式:维持甲基化和从头甲基化。DNA 复制结束后,甲基基团连接到半甲基化对称序列 (CpG、CpNpG) 胞嘧啶上叫作维持甲基化,该形式稳定存在于自身发育和世代传递之间;对非甲基化的 DNA 进行甲基化叫作从头甲基化,对称序列从头甲基化可以一步到位,随后可通过维持甲基化保持,而非对称序列甲基化较复杂,需频繁进行从头甲基化^[6-7]。目前在植物中已鉴定出 4 个 DNA 甲基转移酶亚家族^[8]:MET (Methyltransferase) 维持富含转录元件、重复序列和基因序列的异染色质区域的 CG 甲基化^[5,9];CMT (Chromo-methyltransferases) 为植物特有甲基转移酶,在 DNA 复制过程中靶向沉默转座子和异染色质以维持 CHG、CHH 序列甲基化,CMT 也作用于某些组蛋白修饰位点从头甲基化^[10],MET、CMT 与动物 DNMT1 同源;DRM (Domain rearranged methyltransferases) 与动物 DNMT3 同源,作用于 CG、CHG 和 CHH 序列从头甲基化进而催化 RNA 介导的 DNA 甲基化 (RdDM) 过程^[11];DNMT2 (DNA methyltransferase homologue 2) 在酵母、果蝇、动物和植物中都有发

现,不同于其他甲基转移酶而具有较宽的底物特异性和重要的细胞功能^[6]。

在逆境胁迫下,生物体启动各种 DNA 甲基转移酶催化甲基基团迅速、可逆地对 DNA 进行修饰以保护基因组,显然表观遗传修饰优先于基因组进化^[12],而且由于存在一套识别半甲基化序列 (一条链被甲基化修饰) 的系统能够转化成全甲基化状态 (两条链都被甲基化修饰)^[4],表观遗传修饰可以在代与代之间遗传。有关研究表明大麻能有效抵抗重金属胁迫与其根部高甲基化密切相关,其根部甲基化程度较重金属敏感植物三叶草高 2.3 倍^[13];低温诱导可使拟南芥 DNA 甲基化水平降低,促进花期提前,机理类似于春化作用^[14];5-azacytidine 处理亚麻也能降低 DNA 甲基化水平,获得矮化和提早开花株系,且可稳定遗传至第 3 代^[15]。近年来众多作物育种科学家开始重视表观遗传修饰对作物遗传改良的影响,深入对植物甲基化分子机制的研究,对进一步筛选抗性、矮化及早熟株系具有重要的意义。

蓖麻 (*Ricinus communis* L.) 为大戟科双子叶一年生或多年生草本植物,是世界十大油料作物之一^[16],种子含油量在 50% 以上,蓖麻油可与酒精任意比例混合用于汽车发动机燃料,且燃烧充分、环境污染小,是当代最具潜力的生物柴油;蓖麻油可合成综合性能良好的尼龙 11,广泛用于汽车、电子电器、军工等领域^[17];蓖麻油中 90% 为蓖麻油酸,因其低温不凝固、高温不燃烧的特性,被用作航天航空润滑油。蓖麻具有耐盐碱、耐干旱、耐贫瘠、适应性强等优点,可以种植在边际土地上而不与粮食争地,所以蓖麻有较大的种植潜力。中国为世界第二大蓖麻种植国,虽然种植面积大,但单产较低成为制约蓖麻种植产业推广的

因素之一。因此推动世界蓖麻产业发展急需加强蓖麻遗传改良育种的研究,蓖麻优良种源性状包括:高产,株型矮、早熟和种子含油量等^[19],目前良种选育方法主要是杂交育种、诱变育种、抗性育种和杂优利用^[18],大多通过改变 DNA 序列而获得改良品种,少有涉及到 DNA 序列修饰,原由是有关蓖麻表观遗传的研究甚少,没有足够的理论依据支持深入重大研究工作的开展。本文利用生物信息学从蓖麻基因组数据中挖掘出 DNA 甲基转移酶相关基因并进行了分类及功能域比对,随后对各个基因组织表达和时间表达进行分析,旨在为进一步研究 DNA 甲基化与蓖麻品种改良奠定分子基础,对从表观遗传学上展开深入研究具有重要理论指导作用。

1 材料与方法

1.1 蓖麻 DNA 甲基转移酶的获取与确定

以拟南芥 DNA 甲基转移酶蛋白序列(下载自 TAIR 数据库, <http://www.arabidopsis.org/>)作为基序,在蓖麻基因组数据库(Phytozome v9.0 数据库)中进行 Blastp 同源序列比对,以 $P < 0.001$, 相似度 $> 40\%$ 为显著同源标准确定候选蛋白,随后将候选蛋白序列在拟南芥数据库中一一进行同源比对,进一步确认序列真实性,利用 SMART (<http://smart.embl-heidelberg.de/>) 分别检测分析蓖麻 DNA 甲基转移酶的功能结构域存在与否。

1.2 DNA 甲基转移酶分子进化树构建

利用候选蓖麻 DNA 甲基转移酶蛋白序列在 Phytozome v9.0 数据库中比对获得其他物种的同源序列, Clustal W 软件进行氨基酸多序列比对分析, MEGA 5.0 软件中选择 Neighbor-Join 法构建系统进化树,设定 Bootstrap 的值为 1 000 进行校正,最终生成系统发育树。

1.3 蓖麻 DNA 甲基转移酶的理化特性和结构分析

利用 ExPASy Proteomics Server 工具 (<http://www.expasy.org/tools>) 中提供的 ProtParam 和 ProtScale 软件进行酶蛋白的分子量、理论等电点、亲水性和稳定性分析;利用 SignalP 4.0 程序 (<http://www.cbs.dtu.dk/services/SignalP-4.0/>) 分析信号肽;蛋白亚细胞定位首先利用 Predotar1.03 (<http://urgi.versailles.inra.fr/predotar/>

<http://urgi.versailles.inra.fr/predotar/>) 软件进行初步预测,排除蛋白存在线粒体、质体和内质网后再利用 ProtComp v. 9.0 (<http://linux1.softberry.com/berry.phtml>) 软件进一步确定其具体位置;利用 cNLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) 软件预测蛋白核定位信号;利用 Pfam 数据库 (<http://pfam.xfam.org/>) 和 MEME 在线软件 (<http://meme.nbcr.net/meme/>) 分别预测 DNA 甲基转移酶蛋白结构域(domain)和基序(motif);利用 ExPASy 工具中的 SOPMA 软件预测 α -螺旋(α -helix, H) β -转角(β -turn, T) 无规则卷曲(random coil, C) 以及延伸链(extended strand, E) 等二级结构。

1.4 蓖麻 DNA 甲基转移酶基因的表达分析

试验所用蓖麻为宿生野生品种,根据蓖麻种子发育形态学描述^[20],授粉后早期胚乳和胚为液泡化状态,20 d 胚乳形成而胚仍为液泡化,因此我们采集的样品包括:15 d 种子,20 d 胚乳,25 d 胚和胚乳,50 d 胚和胚乳,样品分离处理后立即放于液氮中保存。提取纯化总 RNA:使用大连宝生物工程有限公司的植物 Total RNA 提取试剂及柱式纯化试剂盒抽提并纯化各个样品的总 RNA,通过琼脂糖凝胶电泳检测 RNA 完整性,并使用 Nanodrop 仪器检测 RNA 浓度和质量。反转录 cDNA:利用北京全式金生物技术有限公司的 TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix 试剂盒对分离纯化的 poly A + RNA 进行反转录。半定量 RT-PCR (semi-quantitative real time PCR):利用 Primer 5.0 对蓖麻 8 个 DNA 甲基转移酶的 CDS 序列设计 PCR 引物(表 1),引物全由上海捷瑞生物工程有限公司合成;RT-PCR 所用模板为各个样品反转录的 cDNA,反应条件为 94 $^{\circ}\text{C}$ 预变性 5 min, 94 $^{\circ}\text{C}$ 变性 30 s、50 $^{\circ}\text{C}$ 退火 30 s、72 $^{\circ}\text{C}$ 延伸 50 s、35 个循环,72 $^{\circ}\text{C}$ 延伸 5 min。

2 结果与分析

2.1 蓖麻 DNA 甲基转移酶基因的获取确定与结构域分析

利用拟南芥 11 个 DNA 甲基转移酶的氨基酸序列在蓖麻基因组数据库中进行同源序列比对获得 8 个相应的蓖麻同源蛋白,比对参数列于表 2 中,从蓖麻基因组中发掘出 2 个 MET、2 个 CMT、

3 个 DRM 和 1 个 DNMT2 同源序列，其中 MET 编码序列较长，分别是 1 542 和 1 584 个氨基酸。所
有序列与拟南芥同源序列的同源性都在 40% 以上，E 值小于 1E - 100。

表 1 PCR 引物序列

Tab. 1 PCR primer sequences

引物名称 primer name	引物序列 primer sequence (5' - 3')	扩增长度/bp PCR product length	退火温度/℃ annealing temperature
ReCMT1-F	CCTATTGAGGGCTTGAGTGA	431	52. 2
ReCMT1-R	GTGTTGGCAATGGATACTGG		52. 5
ReCMT2-F	TTTCGGTTGCGTGTCTTCTT	487	54. 0
ReCMT2-R	GCACTTGCTCCTTTGTTGGA		54. 6
ReDNMT2-F	TTCACCTCCTTGTCAGCCCTAC	421	57. 7
ReDNMT2-R	GGCTCACATGACTGGAGCAA		55. 9
ReDRM1-F	TGCTCTGGGTAGGCAGGAAT	442	56. 1
ReDRM1-R	CCGCCAACAACAAGATCAAAT		53. 6
ReDRM2-F	CCAAGACCACCACTCACCAT	475	55. 2
ReDRM2-R	GCGGTATCCCAAGTCTATGC		53. 9
ReDRM3-F	GTCCATATTAGATGCAATGCCACA	460	56. 0
ReDRM3-R	AATGCCAAGGCGGTGTAAAG		54. 8
ReMET1-F	CCACGCAGGTTAAGATTCCG	429	54. 7
ReMET1-R	ATCCCTCTGACAAGCCACCA		56. 3
ReMET2-F	TCTTATCTTTCGCTGATTACTTCC	468	53. 2
ReMET2-R	TTAGGACCACTGCATCTCCC		54. 4

表 2 蓖麻 DNA 甲基转移酶基因的获取与定位信息

Tab. 2 Obtaining and locating DNA methyltransferase genes from *R. communis*

拟南芥 DNA 甲基转移酶 DNA methyltransferase of <i>Arabidopsis</i>	基因座位 gene loci	编码蛋白长度/aa length of coded proteins	同源性/% homology	S 值 ^a S value	E 值 ^a E value	拟南芥同源基因 ^b homologue from <i>Arabidopsis</i> ^b
AtMET1	29983. m003308	1 542	62. 3	1 818. 1	0	AT5G49160/MET1
AtMET2b	29609. m000606	1 584	54. 8	1 609. 7	0	AT4G14140/MET2
AtCMT1	28582. m000332	845	52. 7	808. 9	0	AT1G80740/CMT1
AtCMT2	29827. m002677	734	63. 8	946	0	AT4G19020/CMT2
AtDRM1	29631. m001043	479	58. 5	500	4. 6E-171	AT5G15380/DRM1
AtDRM2	29917. m001982	686	56. 00	638. 3	0	AT5G14620/DRM2
AtDRM3	29889. m003366	678	40. 50	678	9. 20E-138	AT3G17310/DRM3
AtDNMT2	29848. m004665	404	66. 80	404	7. 70E-175	AT5G25480/DNMT2

注：a. 利用 Blast 比对蓖麻蛋白质数据库；b. Blast 比对拟南芥蛋白质数据库。

Note: a. Blast with *Ricinus communis* database; b. Blast with *Arabidopsis* protein database.

如图 1 所示：所有 DNA 甲基转移酶 C 端催化区域都包含 MTase 结构域（DNA methylase domain），而 N 端调控区域的差异将蓖麻 DNA 甲基转移酶分成 4 类，这种 N 端差异均存在于哺乳动物和植物中^[21]。通过 MEME 分析发现所有 DNA 甲基转移酶蛋白在 C 端催化区域都有 6 个高度保守的基

序：Motif X 和 Motif I 为 S - 腺苷 - L - 蛋氨酸结合亚结构，MotifIV、MotifVI、MotifVIII、MotifIX 为胞嘧啶 - 5 甲基转移酶功能催化位点，其中 MotifIV 为激活位点，MotifVI 是靶向胞嘧啶结合位点，MotifVIII 是 DNA 中和区域，Motif IX 是靶序列位点识别区域^[8,10]。通过数据库 Pfam 搜索获得蓖麻 2 个 MET

同源蛋白 29983. m003308 和 29609. m000606 有 3 种结构域: BAH 结构域 (bromo-adjacent homology domain, BAH domain) 具有蛋白-蛋白互作调控功能, 在 DNA 甲基化链接、复制和转录调控中有重要作用^[22]; RFD 结构域 (replication foci domain, RFD domain) 与 BAH 结构域位于 N 端都具有调控作用, 植物 RFD 结构域可以抑制来源于复制复合体的未甲基化 CpG 二核苷酸的催化基序与 DNA 结合^[10]; MTase 结构域催化 S-腺苷-蛋氨酸中甲基基团转移到 DNA 致使胞嘧啶甲基化^[23]。蓖麻 2 个 CMT 同源蛋白 28582. m000332 和 29827. m002677 结构域分析显示: N 端调控区域有 1 个 BAH 结构域, C 端 MTase 结构域的 Motif IV 和 Motif I 之间存在 1 个 Chromo 结构域 (chromatin modifier domain, Chromo domain), 该结构域对引导蛋白结合到异染色质上至关重要, 由此表明 CMT 在修饰异染色质 DNA 中起到重要作用^[8, 24]; 与拟南芥、水稻、玉米和大豆

相似^[6, 10], 在 BAH 结构域之前, 28582. m000332 比 29827. m002677 多 105 个氨基酸, 并在该区域发现 1 个核定位信号位点 (NLS)。Pafm 分析显示蓖麻 DRM 和 DNMT2 同源蛋白仅有 MTase 结构域, 且 DNMT2 同源蛋白 (29848. m004665) 缺少 N 端调控区域。蓖麻 3 个 DRM 同源蛋白的 Motif I 和 Motif X 相邻, 同拟南芥、水稻、烟草、玉米和大豆 DRM 一致^[6, 10], 相关研究发现 DRM 的 MTase 结构域中 Motif 呈循环排列致使 Motif I 和 Motif X 在折叠结构中更接近, 并能维持 MTase 结构域的稳定^[7, 10]。通过 cNLS Mapper 分析发现: MET 同源蛋白 (29983. m003308)、CMT 同源蛋白 (28582. m000332) 以及 3 个 DRM 同源蛋白 (29631. m001043、29917. m001982 和 29889. m003366) 都有至少 1 个核定位信号序列 (NLS), 其中 29631. m001043 有 2 个 NLS, 蛋白序列存在多个 NLS 可调控蛋白高效运送到核进而控制蛋白功能^[25]。

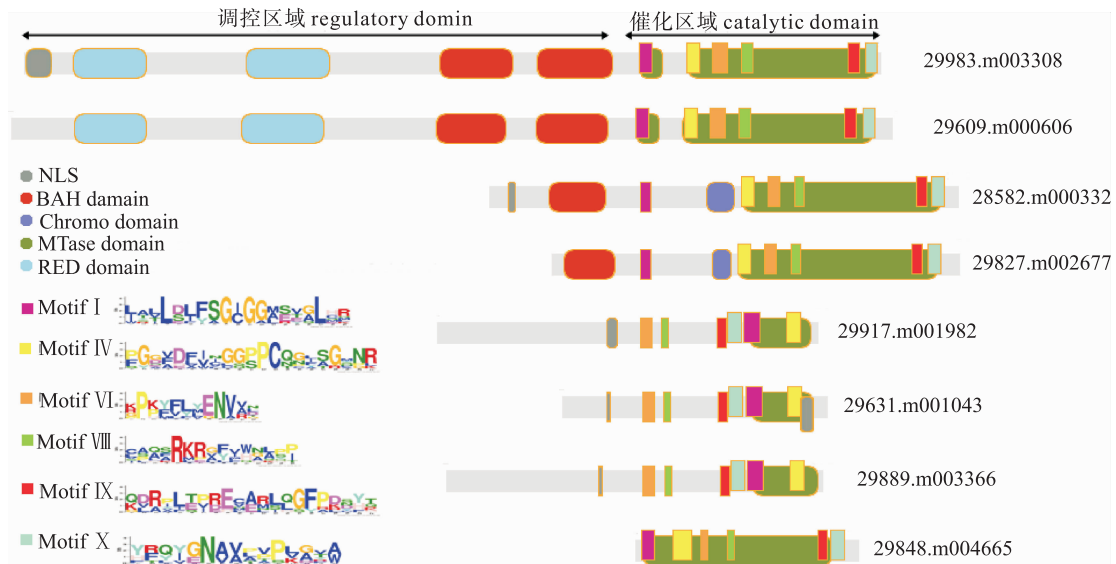


图 1 蓖麻 DNA 甲基转移酶蛋白结构域
Fig. 1 Predicted protein domain of DNA methyltransferase of *R. communis*

2.2 蓖麻 DNA 甲基转移酶蛋白的系统进化分析

选择与蓖麻甲基转移酶相似性较高的其他物种的同源蛋白用于进一步分析蓖麻 DNA 甲基转移酶的进化关系 (表 3), 用 MEGA 5.0 软件中的 Neighbor-Joining 法构建其他系统发育树。如图 2 结果所示: DNA 甲基转移酶明显聚类成 4 个独立分支: CMT、MET、DNMT2 和 DRM。其中 CMT 形成 3 个小分支: CMT1、CMT2 和 CMT3, 蓖麻

28582. m000332 在 CMT2 支系上, 故命名为 RcC-MT1, 29827. m002677 在 CMT2 支系, 命名为 RcC-MT2; 蓖麻 MET 同源蛋白 29983. m003308 和 29609. m000606 分化较小, 分别命名为 RcMET1 和 RcMET2; 在选择 11 种植物中都分别仅鉴定出 1 个 DNMT2 同源蛋白, 蓖麻 29848. m004665 命名为 RcDNMT2; DRM 同源蛋白 29889. m003366 与拟南芥 AtDRM3 类属于同一个小分支, 故命名为

RcDRM3, 29917. m001982 和 29631. m001043 分化较小, 分别命名为 RcDRM1 和 RcDRM2。蓖麻 DNA 甲基转移酶在单双子叶植物中有明显的分化, 表明 DNA 甲基转移酶的分化在单子叶和双子叶进

化分开后; 在各个分支上与同科植物木薯的亲缘关系较近, 与之相似, 拟南芥与同科植物瘦果的同类酶也显示出较近的亲缘关系, 表明该基因在进化过程中较为保守。

表 3 进化树分析所用物种的基本信息

Tab. 3 Basic information of related species in analyzing the phylogenetic tree

物种名称 species name	简写名称 short	分类 classification	物种名称 species name	简写名称 short	分类 classification
蓖麻 castor	Rco	D	大豆 soybean	Gma	D
木薯 manioc	Mes	D	毛果杨 <i>P. trichocarpa</i>	Ptr	D
拟南芥 <i>Arabidopsis</i>	Ath	D	玉米 maize	Zma	M
瘦果 capsella	Cru	D	水稻 rice	Osa	M
可可 cacao	Tca	D	高粱 sorghum	Sbi	M
番茄 tomato	Sly	D			

注: D. 双子叶植物; M. 单子叶植物。
Note: D. Dicotyledon; M. Monocotyledon.

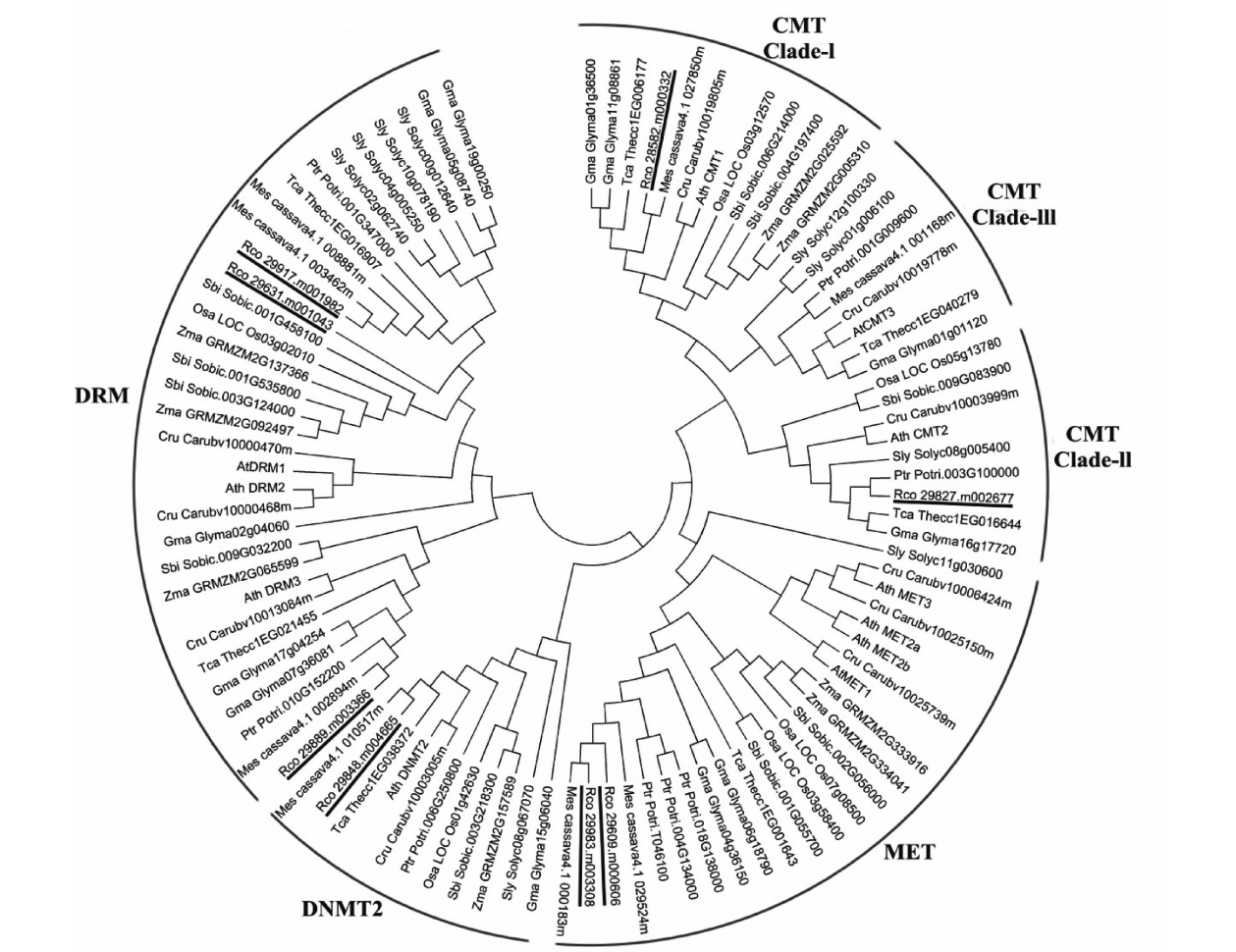


图 2 蓖麻DNA甲基转移酶蛋白系统发育树

Fig. 2 Phylogenetic tree of DNA methyltransferase from *R. communis* and other species

2.3 蓖麻 DNA 甲基转移酶蛋白理化性质、二级结构分析和亚细胞定位预测

如表 4 所示：蓖麻 RcMET1 和 RcMET2 分别编码 1 542 和 1 548 个氨基酸序列，相对分子质量分别为 173 812.5 和 179 867 ku，为分子量较大的 2 种酶蛋白；另外 6 种甲基转移酶蛋白编码氨基酸数量在 400 到 850 之间，相对分子量在 10 万 ku 以内，为分子量较小的酶蛋白。RcCMT2 的酸性氨基酸较碱性氨基酸多 1 个，pI 值为 6.95，为中性稍偏酸性蛋白质；RcDRM2 的 pI 值为 9.21，为碱性蛋白质；由 pI 值预测其余 6 种酶蛋白为酸性蛋白质。对蛋白质亲/疏水性

进行预测，正值越大疏水性越强，负值越大亲水性越强，由此推断 8 种酶蛋白皆为亲水性蛋白。稳定性预测表明：RcCMT1、RcCMT2 和 RcDNMT2 为稳定蛋白，其余 5 个酶蛋白均为不稳定蛋白。以上理化性质的分析对蓖麻 DNA 甲基转移酶蛋白的分离提纯具有重要的意义。信号肽预测结果显示：8 种氨基酸均不存在信号肽，表明它们不经过蛋白转运，而是在胞内作用于甲基化底物。利用 SOPMA 对蓖麻甲基转移酶蛋白的二级结构进行预测，结果如表 5 所示：8 种蛋白皆主要由无规则卷曲和 α-螺旋构成，而 β-转角和延伸链散布于整个蛋白中。

表 4 蓖麻 DNA 甲基转移酶基因编码蛋白序列的理化性质
Tab. 4 Physical and chemical properties of protein coded by DNA methyltransferase genes from *R. communis*

基因座位 gene loci	蛋白 protein	分子量/ku melocular weight	碱性氨基酸 alkaline amino acids	酸性氨基酸 acidic amino acids	pI	亲水性 hydrophilicity	稳定性 stability	信号肽 signal peptide
28582. m000332	RcCMT1	95 119. 6	112	133	5. 29	- 0. 573	36. 63 稳定 stable	no
29827. m002677	RcCMT2	83 319. 7	94	95	6. 95	- 0. 451	38. 96 稳定 stable	no
29848. m004665	RcDNMT2	45 680. 6	40	46	6. 21	- 0. 383	34. 78 稳定 stable	no
29917. m001982	RcDRM1	77 363. 7	78	103	5. 24	- 0. 379	42. 66 不稳定 unstable	no
29631. m001043	RcDRM2	54 776. 9	66	53	9. 21	- 0. 405	43. 32 不稳定 unstable	no
29889. m003366	RcDRM3	77 373. 8	80	103	5. 46	- 0. 606	43. 43 不稳定 unstable	no
29983. m003308	RcMET1	173 812. 5	205	226	6	- 0. 548	43. 45 不稳定 unstable	no
29609. m000606	RcMET2	179 867	209	221	6. 31	- 0. 428	42. 05 不稳定 unstable	no

表 5 蓖麻 DNA 甲基转移酶基因编码蛋白的二级结构分析
Tab. 5 Secondary structure analysis of protein coded by DNA methyltransferase genes from *R. communis*

基因座位 gene loci	蛋白 protein	α-螺旋 alpha helix (Hh)	β-转角 beta turn (Tt)	延伸链 extended strand (Ee)	无规则卷曲 random coil (Cc)
28582. m000332	RcCMT1	244 (28. 88%)	140 (16. 57%)	51 (6. 04%)	410 (48. 52%)
29827. m002677	RcCMT2	227 (30. 93%)	121 (16. 49%)	37 (5. 04%)	349 (47. 55%)
29848. m004665	RcDNMT2	134 (33. 17%)	62 (15. 35%)	22 (5. 45%)	186 (46. 04%)
29917. m001982	RcDRM1	270 (39. 36%)	59 (8. 6%)	33 (4. 81%)	324 (47. 23%)
29631. m001043	RcDRM2	203 (42. 38%)	51 (10. 65%)	29 (6. 05%)	196 (40. 92%)
29889. m003366	RcDRM3	265 (39. 09%)	56 (8. 26%)	31 (4. 57%)	326 (48. 08%)
29983. m003308	RcMET1	523 (33. 92%)	244 (15. 82%)	80 (5. 19%)	695 (45. 07%)
29609. m000606	RcMET2	598 (37. 75%)	276 (17. 42%)	93 (5. 87%)	617 (38. 95%)

利用 ProtComp 9.0 对蓖麻 8 种 DNA 甲基转移酶进行亚细胞定位预测分析，结果如表 6 所示。所有蛋白定位于细胞核的得分都在 8.4 以上；所有蛋白在高尔基体上都有预测到定位得分，其中 RcDRM2 和 RcMET1 得分较高，分别为 1.44 和 1.10；RcDNMT2 和 RcDRM1 在细胞膜上定位得分分别为 0.09 和 0.03，其余蛋白在膜上定位得分均为 0.00；RcDRM2 和 RcMET1 在液泡中的定位得分分别为 0.08 和 0.06，其余蛋白在液泡中没有定位；除 RcMET1 在胞外定位，其余蛋白没有预测到定位于胞外。结果预示可能所有的 DNA 甲基转移酶都在细胞核中发挥功能。

表 6 蓖麻 DNA 甲基转移酶基因编码蛋白的亚细胞定位分析
Tab. 6 Subcellular localization prediction of protein coded by DNA methyltransferase genes from *R. communis*

基因座位 gene loci	细胞核 nuclear	细胞膜 plasma membrane	胞外 extracellular	高尔基体 golgi	液泡 vacuolar
28582. m000332	9. 98	0	0	0.02	0
29827. m002677	9. 97	0	0	0.03	0
29848. m004665	9. 9	0.09	0	0.01	0
29917. m001982	9. 96	0.03	0	0.01	0
29631. m001043	8. 48	0	0	1.44	0.08
29889. m003366	9. 97	0	0	0.03	0
29983. m003308	8. 8	0	0.03	1.1	0.06
29609. m000606	9. 95	0	0	0.05	0

2.4 蓖麻 DNA 甲基转移酶基因的表达分析

为检测 8 个甲基转移酶基因在组织间的表达情况，我们从 GenBank 数据库中下载蓖麻 5 个组织转录组数据（www. ncbi. nlm. nih. gov/sra accession SRX007402 to SRX007408）^[26]。表达谱分析（图 3A）发现：除 RcDRM2 仅在雄花中表达外，其余 7 个基因在所有检测组织中皆可检测到表达量，且在发育胚乳中的表达量比其他组织高；除 RcMET2 外，其余 6 个基因的表达量在后期胚乳中均下调，RcMET1 和 RcCMT1 在前期胚乳中的表达量显著高于其余检测组织；RcDNMT2 在后期胚乳、萌芽种子和叶片中的表达量差异不大。

为进一步检测甲基转移酶在蓖麻种子发育阶段的表达情况，取授粉后 15 d 种子、20 d 胚乳、25 d 胚和胚乳、50 d 胚和胚乳作为不同时期的种子材料，通过 RT-PCR 检测各个基因在各个时期的表达量发现：RcDRM2 仅在 15 d 的种子中表达；RcDRM3 在种子发育过程中没有显著变化；RcDRM1、RcDNMT2、RcCMT1、RcCMT2、Rc-

MET1 和 RcMET2 在 15 d 种子中表达量低，随种子发育到 25 d（前期胚乳）表达量最高，随后至 50 d（后期胚乳）RcCMT1、RcCMT2 和 RcMET1 的表达量显著降低，而 RcMET2 在 50 d 胚乳中表达量上调，该结果与转录组数据一致；RcCMT1、RcCMT2 和 RcMET2 在 50 d 胚中的表达量较 25 d 胚显著下调；25 d 种子中 RcMET1 在胚乳中表达量比胚中高，而 RcMET2 在胚乳中表达量比胚中低；50 d 种子中 RcMET1 在胚中表达量比胚乳高，而 RcMET2 在胚中表达量比胚乳低。

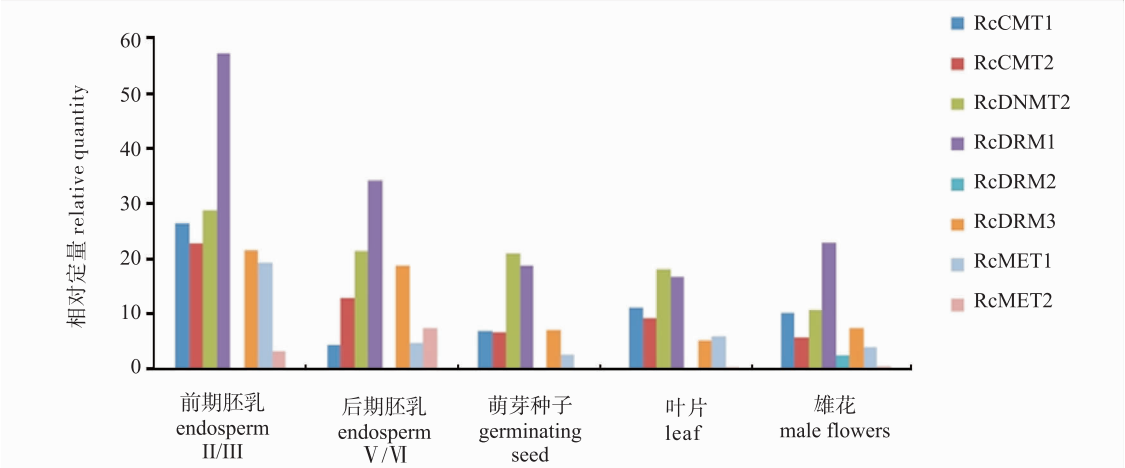
3 讨论

通过蛋白结构域预测发现蓖麻甲基转移酶的结构域与其他物种同类酶是一致的，所以应该具有相似的维持甲基化和从头甲基化建立功能。在植物中 DNMT2 缺少 N 端调节结构域，不具有建立或维持甲基化的重要作用，对其功能研究较为复杂而未阐明^[27]，所以 RcDNMT2 有待进一步研究。通过核定位信号预测发现 RcMET1、RcCMT1、RcDRM1、RcDRM2 和 RcDRM3 均有核定位信号序列（NLS），NLS 是蛋白质通过主动运输进入核孔的充要条件^[28]，所以推断这 5 个蛋白是在细胞核内发挥生物学功能，如 DNA 复制过程中维持甲基化、从头甲基化修饰 DNA 或染色质等。亚细胞定位预测表明 8 个蛋白皆定位在细胞核，所以推测 RcMET2、RcCMT2 和 RcDNMT2 的核定位信号可能属于 cNLS（经典核定位信号）以外的类型^[28]，有待进一步探究。

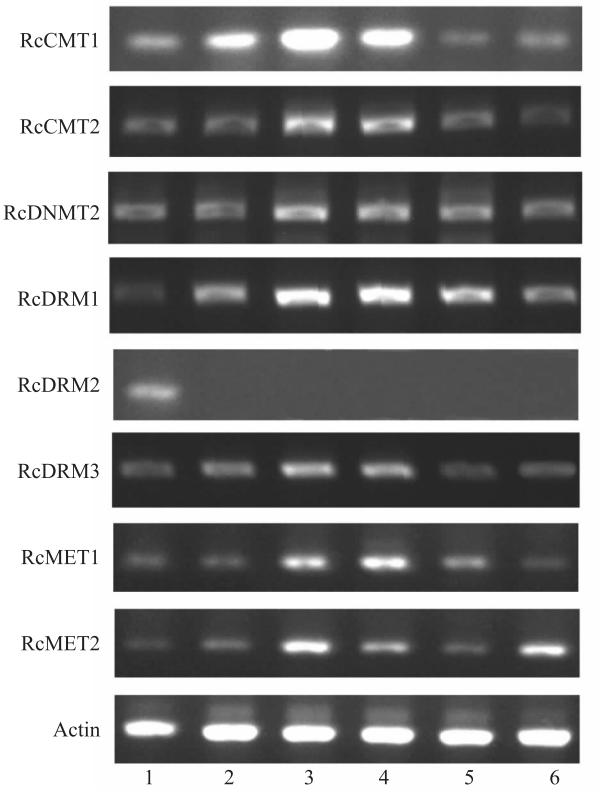
基因表达谱通常与基因功能鉴定相关。蓖麻甲基转移酶在种子发育中期有较高的表达，表明种子在该时期甲基化状态发生很大变化，有研究对授粉后 35 d 蓖麻胚乳进行高通量重亚硫酸盐测序后发现胚乳特异性低甲基化^[29]，该时期去甲基化相关基因 DNA 糖基化酶 RcDME 和 RcROS1 在蓖麻胚乳中高表达^[30]，所以推测胚乳低甲基化可能诱导甲基化相关基因在胚乳中特异性表达增强。RcCMT1 和 RcCMT2 表达模式相似，预示可能存在功能冗余基因。RcDRM 和 RcMET 亚家族中的基因表达模式有明显差异，预示可能有亚功能或假功能基因^[10]，RcMET1 和 RcMET2 在种子发育前后期呈相反的表达模式，认为可能在发育种子中功能互补。RcMET1 在前期胚乳中高表达，胚乳是植物表观印迹发生的主要组织^[31]，且有研究发现 MET1 在水稻、玉米和拟南芥的胚乳中都有较高的表达量^[6, 27]，

MET1 参与维持甲基化并调控胚乳基因组印记^[32]，所以 RcMET1 在胚乳中比胚中表达水平高。RcMET2 在后期胚乳中高表达，功能可能参与维持其他非印记基因甲基化。另外我们对比蓖麻与豆科植物同一亚家族的基因表达模式也发现差异，表明甲

基转移酶功能存在种系差异。DNMT2 有较宽的底物特异性^[6]，我们发现 RcDNMT2 的表达模式没有组织特异性，所以 RcDNMT2 生物学功能可能较为复杂而不仅限于 DNA 甲基化，如拟南芥中发现 AtDNMT2 作用于组蛋白去乙酰化^[33]。



A. 转录组数据分析蓖麻DNA甲基转移酶基因表达
Expression analysis of DNA methyltransferase genes from Brown's study



B. RT-PCR半定量检测蓖麻DNA甲基转移酶基因表达
RT-PCR analysis of DNA methyltransferase genes expression level in various tissues

注：1.15 d种子;2. 20 d胚乳;3. 25 d胚;4. 25 d胚乳;5. 5.0 d胚;6. 50 d胚乳。
Note: 1.15 days seed; 2. 20 days endosperm;3. 25 days embryo;4. 25 days endosperm; 5. 50 days embryo; 6. 50 days endosperm.

图 3 蓖麻DNA甲基转移酶基因组织特异表达
Fig. 3 Expression analysis of DNA methyltransferase genes from different tissues in *R. communis*

本研究通过同源序列比对的方法从蓖麻基因组中鉴别出8个DNA甲基转移酶,2个MET同源蛋白、2个CMT同源蛋白、3个DRM同源蛋白和1个DNMT2同源蛋白。结构域分析显示8个甲基转移酶蛋白的C端均是保守的功能结构域,N端存在相异的调节结构域,而RcDNMT2缺少调节结构域;系统进化分析表明DNA甲基转移酶在进化上较为保守,分化可能发生在单双子叶植物分化之后;氨基酸序列理化性质分析为蓖麻甲基转移酶蛋白的提取纯化提供了依据;基因表达分析发现在蓖麻DNA甲基转移酶基因在发育种子中高表达与该时期的胚乳低甲基化水平有关。该研究为深入研究蓖麻DNA甲基化调控的表观遗传学奠定了分子基础,为开展蓖麻表观遗传相关的育种研究提供了重要的理论依据。

[参考文献]

- [1] COLOT V, ROSSIGNOL J L. Eukaryotic DNA methylation as an evolutionary device [J]. *Bioessays*, 1999, 21 (5): 402.
- [2] YODER J A, WALSH C P, BESTOR T H. Cytosine methylation and the ecology of intragenomic parasites [J]. *Trends in Genetics*, 1997, 13 (8): 335.
- [3] 陈欣,王子成. 植物DNA甲基转移酶 [J]. *生命的化学*, 2009, 29 (4): 534.
- [4] TSAFTARIS A S, POLIDOROS A N. DNA methylation and plant breeding [J]. *Plant Breeding Reviews*, 1999, 18: 87.
- [5] COKUS S J, FENG S, ZHANG X, et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning [J]. *Nature*, 2008, 452 (7184): 215.
- [6] SHARMA R, MOHAN SINGH R, MALIK G, et al. Kapoor M. Rice cytosine DNA methyltransferases gene expression profiling during reproductive development and abiotic stress [J]. *FEBS Journal*, 2009, 276 (21): 6301.
- [7] CAO X, SPRINGER N M, MUSZYNSKI M G, et al. Conserved plant genes with similarity to mammalian de novo DNA methyltransferases [J]. *Proceedings of the National Academy of Sciences*, 2000, 97 (9): 4979.
- [8] PAVLOPOULOU A, KOSSIDA S. Plant cytosine-5 DNA methyltransferases; structure, function, and molecular evolution [J]. *Genomics*, 2007, 90 (4): 530.
- [9] LISTER R, O'MALLEY R C, TONTI-FILIPPINI J, et al. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis* [J]. *Cell*, 2008, 133 (3): 523.
- [10] GARG R, KUMARI R, TIWARI S, et al. Genomic Survey, Gene expression analysis and structural modeling suggest diverse roles of DNA methyltransferases in Legumes [J]. *PLoS One*, 2014, 9 (2): e88947.
- [11] CAO X, AUFSATZ W, ZILBERMAN D, et al. Role of the DRM and CMT3 methyltransferases in RNA-Directed DNA methylation [J]. *Current Biology*, 2003, 13 (24): 2212.
- [12] PENG H, ZHANG J. Plant genomic DNA methylation in response to stresses: potential applications and challenges in plant breeding [J]. *Progress in Natural Science*, 2009, 19 (9): 1037.
- [13] AINA R, SCORBATI S, SANTAGOSTINO A, et al. Specific hypomethylation of DNA is induced by heavy metals in white clover and industrial hemp [J]. *Physiologia Plantarum*, 2004, 121 (3): 472.
- [14] BURN J, BAGNALL D, METZGER J, et al. DNA methylation, vernalization, and the initiation of flowering [J]. *Proceedings of the National Academy of Sciences*, 1993, 90 (1): 287.
- [15] FIELDS M. Heritable effects of 5-azacytidine treatments on the growth and development of flax (*Linum usitatissimum*) genotypes and genotypes [J]. *Genome*, 1994, 37 (1): 1.
- [16] 傅福勤. 特用油源作物——蓖麻 [M]. 北京: 中国环境科学出版社, 1990.
- [17] 朱志伟, 杜文书, 胡国胜. 蓖麻油深加工产物——尼龙11的合成及应用 [J]. *中国油脂*, 2007, 32 (3): 57.
- [18] 孙振钧, 吕丽媛, 伍玉鹏. 蓖麻产业发展: 从种植到利用 [J]. *中国农业大学学报*, 2012, 17 (6): 204.
- [19] 邱雨俊, 刘爱忠. 我国优势油脂类生物柴油植物蓖麻的种质资源发掘和生物柴油利用 [J]. *生命科学*, 2014, 26 (5): 503.
- [20] GREENWOOD J, BEWLEY J. Seed development in *Ricinus communis* (castor bean). I. Descriptive morphology [J]. *Canadian Journal of Botany*, 1982, 60 (9): 1751.
- [21] FINNEGAN E J, DENNIS E S. Isolation and identification by sequence homology of a putative cytosine methyltransferase from *Arabidopsis thaliana* [J]. *Nucleic Acids Research*, 1993, 21 (10): 2383.
- [22] CALLEBAUT I, COURVALIN J C, MORNON J P. The BAH (bromo-adjacent homology) domain: a link be-

- tween DNA methylation, replication and transcriptional regulation [J]. *FEBS Letters*, 1999, 446 (1): 1893.
- [23] MALONE T, BLUMENTHAL R M, CHENG X. Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyl-transferases, and suggests a catalytic mechanism for these enzymes [J]. *Journal of Molecular Biology*, 1995, 253 (4): 618.
- [24] EISENBERG J C. Molecular biology of the chromo domain; an ancient chromatin module comes of age [J]. *Gene*, 2001, 275 (1): 19.
- [25] LUO M, PANG C, GERKEN A E, et al. Multiple nuclear localization sequences allow modulation of 5-lipoxygenase nuclear import [J]. *Traffic*, 2004, 5 (11): 847.
- [26] BROWN A P, KROON J T, SWARBRECK D, et al. Tissue-specific whole transcriptome sequencing in castor, directed at understanding triacylglycerol lipid biosynthetic pathways [J]. *PLoS One*, 2012, 7 (2): e30100.
- [27] QIAN Y, XI Y, CHENG B, et al. Genome-wide identification and expression profiling of DNA methyltransferase gene family in maize [J]. *Plant Cell Reports*, 2014, 33 (10): 1661.
- [28] 赵元茵, 王元忠, 曹念, 等. 核定位信号及其分析策略 [J]. *中国生物化学与分子生物学报*, 2009, 25 (8): 683.
- [29] XU W, DAI M Y, LI F, et al. Genomic imprinting, methylation and parent-of-origin effects in reciprocal hybrid endosperm of castor bean [J]. *Nucleic Acids Research*, 2014: gku375.
- [30] 代梦媛, 敖涛, 徐伟, 等. 蓖麻 DNA 糖基化酶的序列与基因表达分析 [J]. *中国油料作物学报*, 2014, 36 (2): 181.
- [31] 张美善, 刘宝. 植物胚乳发育的表观遗传学调控 [J]. *植物学报*, 2012, 47 (2): 101.
- [32] XIAO W, GEHRING M, CHOI Y, et al. Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase [J]. *Developmental Cell*, 2003, 5 (6): 891.
- [33] SONG Y, WU K, DHAUBHADEL S, et al. Arabidopsis DNA methyltransferase AtDNMT2 associates with histone deacetylase AtHD2s activity [J]. *Biochemical and Biophysical Research Communications*, 2010, 396 (2): 187.