

# 麻疯树油体钙蛋白 *JcCale26.8* 基因克隆及序列分析

蒲芝雨, 黄建辉, 李国泽, 丁勇\*

(西南林业大学云南省高校林木生物技术重点实验室, 云南昆明 650224)

**摘要** [目的]克隆麻疯树(*Jatropha curcas*)油体钙蛋白基因 *JcCale26.8* 全长 cDNA 序列,探究麻疯树油体钙蛋白及其基因的结构与功能。[方法]应用 RNA-seq 测序和 PCR 技术,从麻疯树种子中克隆获得 1 个油体钙蛋白家族基因 *JcCale26.8*,并进行测序和序列分析。[结果]*JcCale26.8* 基因序列含有 6 个外显子和 5 个内含子,内含子剪接位点符合真核生物基因的 GT-AG 法则。*JcCale26.8* 基因 mRNA 序列的完整开放阅读框长 717 bp,推测编码由 238 个氨基酸组成、相对分子量为 26.8 kD 的油体钙蛋白 *JcCale26.8*。*JcCale26.8* 蛋白主要由  $\alpha$ -螺旋和无规则卷曲结构构成,并具有油体钙蛋白典型结构特征,与来自蓖麻、川桑和异色山黄麻等多个不同物种的 Caleosin 蛋白具有较高的同源相似性。[结论]该研究可为麻疯树油体钙蛋白基因的表达调控与功能研究奠定基础。

**关键词** 种子;油体;麻疯树;油体钙蛋白;基因克隆

中图分类号 S794.9 文献标识码 A

文章编号 0517-6611(2022)12-0086-07

doi: 10.3969/j.issn.0517-6611.2022.12.022



开放科学(资源服务)标识码(OSID):

## Cloning and Sequence Analysis of Caleosin Gene *JcCale26.8* in *Jatropha curcas*

PU Zhi-yu, HUANG Jian-hui, LI Guo-ze et al (Key Laboratory of Forest Biotechnology in Yunnan, Southwest Forestry University, Kunming, Yunnan 650224)

**Abstract** [Objective] To clone the full-length cDNA sequence of *JcCale26.8* gene from *J. curcas* and investigate the structure and function of caleosin and its gene in *J. curcas* seeds. [Method] This study used RNA-sequencing and PCR technology to clone a caleosin family gene *JcCale26.8* from *J. curcas* seeds. Then the sequence of gene *JcCale26.8* were sequenced and analyzed. [Result] The DNA sequence of *JcCale26.8* gene contains 6 exons and 5 introns. The splicing sites of introns conform to the GT-AG rule of eukaryotic genes. The full-length mRNA comprised 717 bp nucleotides consisting of an open reading frame encoded a putative *JcCale26.8* comprising 238 amino acid residues with molecular weight of 26.8 kD. *JcCale26.8* protein is mainly composed of  $\alpha$ -helix and random coil structure, and has the typical structural characteristics of caleosin. It has high homology and similarity with caleosin proteins from many different species such as *Ricinus communis*, *Morus notabilis* and *Trema orientalis*. [Conclusion] The study can lay the experimental foundation for the study of the expression regulation and function of caleosin gene in *J. curcas*.

**Key words** Seed; Oil body; *Jatropha curcas*; Caleosin; Gene cloning

油体是脂质(主要是三酰甘油)储存库,主要存在于种子和衰老的叶片中<sup>[1]</sup>。油体不仅在植物生长发育过程中发挥重要作用,还能参与胁迫响应和激素信号通路等过程<sup>[2]</sup>。油体由外部完整的磷脂单分子层和围绕着疏水核心的膜蛋白组成<sup>[3]</sup>,这些膜蛋白包括油质蛋白(Oleosin)、油体钙蛋白(Caleosin)和油体固醇蛋白(Steroleosin)<sup>[4]</sup>。Oleosin 是植物油体特有,且最为丰富的结合蛋白<sup>[5-6]</sup>,与油体大小和种子萌发率密切相关,并能增强植物种子的抗冻性<sup>[4]</sup>。Caleosin 作为油体上的微量结合蛋白,最早发现于水稻(*Oryza sativa*)颖果发育时期的细胞膜上<sup>[7]</sup>,其在植物种子萌发过程中参与脂质降解<sup>[8]</sup>、油体形成<sup>[9]</sup>及维持油体的稳定<sup>[10]</sup>,在拟南芥(*Arabidopsis thaliana*)的非种子组织中,Caleosin 还表现出过氧羟基脂肪酸羟化环氧化酶活性<sup>[11]</sup>,参与抵抗逆境胁迫<sup>[11]</sup>和植物开花调节<sup>[12]</sup>等过程。

麻疯树(*Jatropha curcas*)又名黄肿树(广东)、芙蓉树、假花生(广西)、高桐(云南)、吗哄罕(傣名)、桐油树(台湾)、南洋油桐(日本),为大戟科(Euphorbiaceae)麻疯树属(*Jatropha*)植物,主要分布于热带和亚热带地区,在我国半野生或栽培种植面积较广,资源丰富<sup>[13]</sup>。成熟麻疯树种子的含油

量高达 36%<sup>[14]</sup>,野生麻疯树种仁的最高含油量约 60%,超过油菜和大豆等常见的油料作物<sup>[15]</sup>,因此开展麻疯树油体结合蛋白及其基因结构与功能研究对于其在油供工业方面的开发利用具有重要意义。而基因克隆与序列分析是开展不同生物目标基因功能研究的前提<sup>[16-18]</sup>。目前,Oleosin 和 Caleosin 油体结合蛋白基因已经在水稻<sup>[7]</sup>、芝麻(*Sesamum indicum*)<sup>[9]</sup>、油菜(*Brassica napus*)<sup>[19]</sup>、拟南芥<sup>[20]</sup>和蓖麻(*Ricinus communis*)<sup>[21]</sup>等不同植物成熟种子中被克隆获得。有关麻疯树中 Oleosin 及其基因序列和结构特征已有研究<sup>[22-24]</sup>,但关于麻疯树 Caleosin 及其基因的研究鲜见报道。笔者以麻疯树种子为材料,应用 RNA-seq 和 PCR 技术克隆麻疯树 Caleosin 基因序列并进行分析,为完善对麻疯树油体膜蛋白的认识及后续解析 Caleosin 在麻疯树中的基因表达特征和功能发挥奠定试验基础。

## 1 材料与方法

**1.1 材料与试剂** 试验材料为 3 年生麻疯树植株发育 40 d 的麻疯树种子,取自西双版纳热带植物园,液氮速冻后放置于 -80 °C 冰箱中保存备用。

RNA 提取 Trizol 试剂和新型快速植物基因组 DNA 提取试剂盒为 OMEGA 公司产品,反转录(5 × All-In-One RT MasterMix)和 PCR(Kodacq 2 × PCR MasterMix with dye)试剂盒为爱必梦(abm)公司产品。

## 1.2 试验方法

**1.2.1 核酸提取与检测。**将麻疯树种子种壳去除,取种仁于

**基金项目** 云南省大学生创新创业训练计划项目(201810677018);云南省高校林木生物技术重点实验室开放基金项目(51700201)。

**作者简介** 蒲芝雨(1998—),女,贵州遵义人,硕士研究生,研究方向:植物生物化学与分子生物学。\*通信作者,硕士,副教授,从事植物生物技术研究。

**收稿日期** 2021-09-19

液氮中研磨并用于核酸提取。总 RNA 的提取按照 Trizol 试剂使用说明书进行,基因组 DNA 的提取按照 OMEGA 公司新型快速植物基因组 DNA 提取试剂盒使用说明书进行;基因组 DNA 和总 RNA 的完整性和纯度分别用 1.0% 琼脂糖凝胶电泳和 Nano Photometer 分光光度计 (IMPLEN CA USA) 检测。

**1.2.2 麻疯树种子 RNA-seq 测序及目标基因挖掘。**检测合格的总 RNA 样本送至安诺优达基因科技(北京)有限公司,构建文库后利用 Illumina 平台进行 RNA-seq 测序,原始下机序列(Raw Reads)通过去低质量序列、去接头污染等过程完成数据处理得到高质量的序列(Clean Reads)。对质控后的高质量序列进行组装,再基于组装序列结果进行 ORF 及功能预测分析,从 RNA-seq 数据库中挖掘具有潜在 Caleosin 功能结构域的目标 cDNA 序列。

**1.2.3 麻疯树 *JcCale26.8* 基因 cDNA 序列克隆验证。**在目标基因 cDNA 序列 ORF 区两侧设计特异性引物对 F(5'-AG-ATGGCTACTAGAACGGACG-3') 和 R(5'-GCCACCATT-GCTTACATTCT-3'),并委托上海生工生物工程技术有限公司合成。取总 RNA 样品为模板,按照反转录试剂盒说

明书进行第一链 cDNA 的合成。取反转录产物 cDNA 模板 2.0  $\mu\text{L}$ , R 和 F 引物(10  $\mu\text{mol}/\mu\text{L}$ )各 0.5  $\mu\text{L}$ , Kodaq 2 PCR MasterMix 12.5  $\mu\text{L}$ , Nuclease-free  $\text{H}_2\text{O}$  9.5  $\mu\text{L}$ , 建立 25  $\mu\text{L}$  PCR 反应体系。PCR 反应程序:94  $^\circ\text{C}$  预变性 3 min, 94  $^\circ\text{C}$  变性 30 s, 53.3  $^\circ\text{C}$  退火 30 s, 72  $^\circ\text{C}$  延伸 45 s, 38 个循环, 72  $^\circ\text{C}$  延伸 10 min。

**1.2.4 麻疯树 *JcCale26.8* 基因 DNA 序列克隆。**取“1.2.1”步骤中提取的高质量 gDNA 样品 3  $\mu\text{L}$  为模板, R 和 F 为引物对, 进行 *JcCale26.8* 基因 DNA 序列的 PCR 扩增, 反应体系和扩增条件同“1.2.3”。

**1.2.5 PCR 产物检测回收与克隆测序。**步骤“1.2.3”和“1.2.4”中 PCR 产物采用 1.0% 琼脂凝胶电泳检测, 目的片段回收后与 pMD19-T 载体连接, 转化 DH5a 感受态细胞, 经复苏后涂布在 LB 固体培养基上(含 Amp、IPTG 和 X-Gal)培养, 蓝白斑筛选阳性克隆, 委托上海生工生物技术服务有限公司测序。

**1.2.6 *JcCale26.8* 基因的生物信息学分析。**应用表 1 显示的网站对 *JcCale26.8* 基因及推导编码的蛋白质序列进行对应项目内容的分析。

表 1 *JcCale26.8* 基因生物信息学分析项目与网址  
Table 1 *JcCale26.8* bioinformatics analysis project and web site

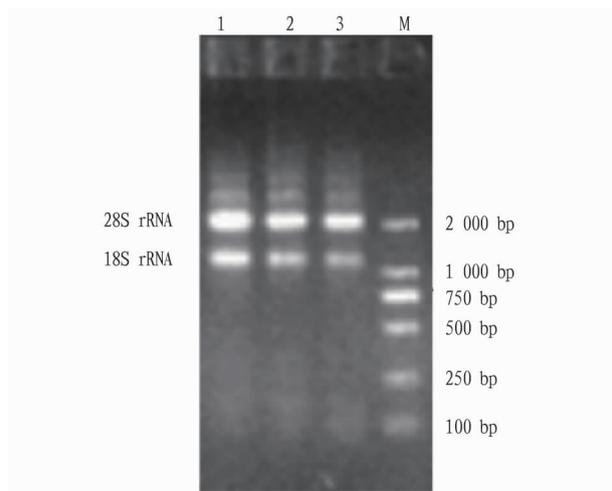
序号 No.	分析项目 Analysis item	网址 The URL
1	核酸序列相似性搜索	Blastn ( <a href="https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&amp;PAGE_TYPE=BlastSearch&amp;LINK_LOC=blasthome">https://blast.ncbi.nlm.nih.gov/Blast.cgi? PROGRAM = blastn&amp;PAGE_TYPE = BlastSearch&amp;LINK_LOC = blasthome</a> )
2	开放阅读框分析	ORF Finder ( <a href="https://www.ncbi.nlm.nih.gov/orffinder/">https://www.ncbi.nlm.nih.gov/orffinder/</a> )
3	氨基酸序列相似性搜索	Blastp ( <a href="https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&amp;PAGE_TYPE=BlastSearch&amp;LINK_LOC=blasthome">https://blast.ncbi.nlm.nih.gov/Blast.cgi? PROGRAM = blastp&amp;PAGE_TYPE = BlastSearch&amp;LINK_LOC = blasthome</a> )
4	多序列相似性比对	DNAMAN ( <a href="http://www.lynnon.com/">http://www.lynnon.com/</a> )
5	核酸和氨基酸序列编辑整理	Uedit32 ( <a href="https://www.ultraedit.com/">https://www.ultraedit.com/</a> )
6	理化性质	ProtParam ( <a href="http://web.Expasy.org/protparam/">http://web.Expasy.org/protparam/</a> )
7	PolyA 信号预测	<a href="http://www.softberry.com/berry.phtml?topic=polyah&amp;group=programs&amp;subgroup=promoter">http://www.softberry.com/berry.phtml? topic = polyah&amp;group = programs&amp;subgroup = promoter</a>
8	保守结构域分析	CD-Search ( <a href="https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi">https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi</a> )
9	信号肽分析	( <a href="http://www.cbs.dtu.dk/services/SignalP-3.0/">http://www.cbs.dtu.dk/services/SignalP-3.0/</a> )
10	氨基酸疏水性分析	( <a href="http://web.expasy.org/protscale/">http://web.expasy.org/protscale/</a> )
11	跨膜结构域分析	TMHMM ( <a href="http://www.cbs.dtu.dk/services/TMHMM/">http://www.cbs.dtu.dk/services/TMHMM/</a> )
12	蛋白质二级结构预测	( <a href="https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html">https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl? page = npsa_sopma.html</a> )
13	蛋白质三维结构预测	Phyer2 ( <a href="http://www.sbg.bio.ic.ac.uk/phyre2/html/">www.sbg.bio.ic.ac.uk/phyre2/html/</a> )
14	系统发育树构建	Dnapars 和 protpars ( <a href="http://evolution.gs.washington.edu/phylip.html">http://evolution.gs.washington.edu/phylip.html</a> )

## 2 结果与分析

**2.1 总 RNA 提取和 *JcCale26.8* 基因的 cDNA 序列克隆分析** 植物组织总 RNA 样品的纯度和完整性是基因克隆等分子生物学试验成功的关键因素<sup>[25-26]</sup>。该试验获得的发育 40 d 麻疯树种子总 RNA 完整性好,显示 28S 和 18S rRNA 条带清晰,点样孔无亮斑,无弥散拖尾,且前者条带的亮度约为后者的 2 倍(图 1);总 RNA 的  $A_{260}/A_{280}$  比值在 1.8~2.0,说明纯度好且无蛋白质、多糖和酚类物质等杂质污染<sup>[27]</sup>。符合 RNA-seq 建库测序和基因克隆试验要求。试验构建的麻疯树种子 RNA-seq 数据库共获得 33 902 个 Unigene, GC 含量占 40.72%,总碱基数达 33 168 384 个,从中挖掘获得 2 条 cDNA 序列具有潜在 Caleosin 功能结构域(表 2),序列 S1 和 S2

均具有完整 ORF 区域。以具有较长 ORF 区的序列 S1 为目标进行克隆验证分析。按“1.2.3”进行 RT-PCR 扩增反应,琼脂糖凝胶电泳结果(图 2)显示,在对应 DL2000 分子量标准的 750 bp 条带处出现 1 条符合预期大小的主条带,经胶回收 T/A 克隆测序,获得与转录组数据库 Unigene 序列 S1 引物间完全一致的长 738 bp 的 cDNA 序列。长 1 026 bp 的目标序列 S1 推导编码分子量为 26.8 kD 的 Caleosin 蛋白,因此命名为 *JcCale26.8* 基因。该试验获得 *JcCale26.8* 基因 cDNA 序列全长 1 026 bp, 1~43 bp 为 5' UTR, 761~1 026 bp 为 3' UTR, 位于 826~831 bp 的一致性序列 AATAAA 和 2 个 U-rich 单元(位于 957~964 bp 和 1 016~1 021 bp 处)为预测的 PolyA 信号位点;44~760 bp 为完整 ORF,推测编码由 238 个

氨基酸组成的 Caleosin 蛋白 JcCale26.8。

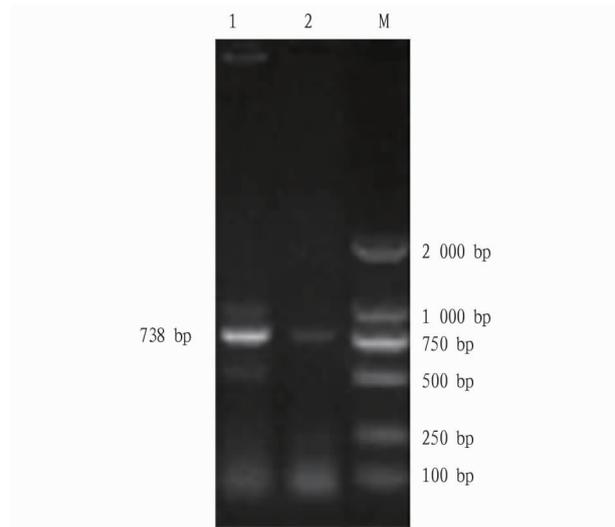


注: M.DL 2000 Marker; 1,2,3. 麻疯树种子总 RNA

Note: M.DL 2000 Marker; 1,2,3. Total RNA of *Jatropa curcas* seeds

图 1 麻疯树种子总 RNA 电泳

Fig.1 Total RNA electrophoresis of *Jatropa curcas* seeds



注: M.DL 2000 Marker; 1,2.cDNA 扩增

Note: M.DL 2000 Marker; 1,2.cDNA amplification

图 2 *JcCale26.8* 基因 cDNA 序列的 PCR 产物电泳

Fig.2 Electrophoresis of cDNA PCR products of *JcCale26.8* gene

表 2 麻疯树种子 RNA-seq 数据库中 2 条具 Caleosin 功能结构域的 cDNA 序列信息

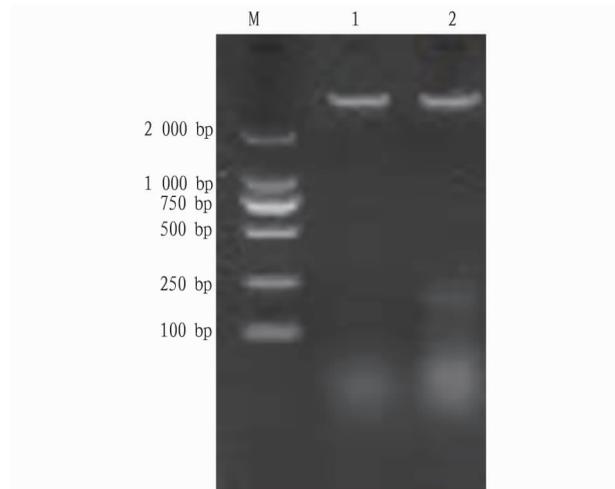
Table 2 Two sequences information of caleosin-domain containing protein from RNA-seq database of *Jatropa curcas* seeds

序列 Sequence	核苷酸序 列长度 Nucleotide sequence length//bp	开放阅 读框 ORF bp	氨基酸序 列长度 Amino acid sequence length//aa	蛋白质家族 数据库名称 PFAM_Name	蛋白质家族 数据库描述 PFAM_Description	蛋白质功能注 释数据库描述 eggNOG_Description
S1	1 026	44~760	238	Caleosin	Caleosin related protein	Caleosin related protein
S2	1 357	730~1 089	119	Caleosin	Caleosin related protein	—

**2.2 *JcCale26.8* 基因 DNA 序列克隆与分析** 该试验提取的麻疯树种子基因组 DNA 完整性好、纯度高。DNA 样品条带清晰单一(图 3),点样孔无亮斑,无弥散拖尾,其  $A_{260}/A_{280}$  值为 1.96。以提取的高质量 DNA 样品为模板,按“1.2.4”进行 PCR 扩增获得大于 2 000 bp 的 DNA 条带(图 4),克隆测序结果为 2 018 bp 的 *JcCale26.8* 基因组 DNA 序列。经 DNA 和 mRNA 序列比对分析(图 5)显示,*JcCale26.8* 基因含有 6 个外显子和 5 个内含子。内含子位于 *JcCale26.8* 基因 DNA 序列 116~395、546~1 008、1 095~1 271、1 367~1 464、1 591~1 674 bp,每个内含子序列 5'端为 GT、3'端为 AG,均符合真核生物基因内含子剪接位点的 GT-AG 规则。

**2.3 *JcCale26.8* 基因序列分析** 通过 NCBI 网站 Blast 工具和 CD-Search 服务器对麻疯树 *JcCale26.8* 基因编码的 238 个氨基酸序列进行相似性搜索分析和功能域分析,结果(图 6)表明,目的基因编码的麻疯树 28.6 kD 的 Caleosin 蛋白为 Caleosin 蛋白超家族成员,具有典型的 Caleosin 功能结构域。该蛋白序列与 NCBI 登录的来自麻疯树的未知蛋白 JCGZ 02382(KDP40384.1)和过氧羟基脂肪酸羟化环氧氧化酶(Peroxygenase,POG)异构体 X1(XP 012069904.1)的氨基酸序列具有 100%的一致性,表明该试验成功克隆获得的麻疯树 Caleosin 蛋白可能具有 POG 酶活性。

*JcCale26.8* 和其他物种 Caleosin 蛋白或具有 POG 酶活性



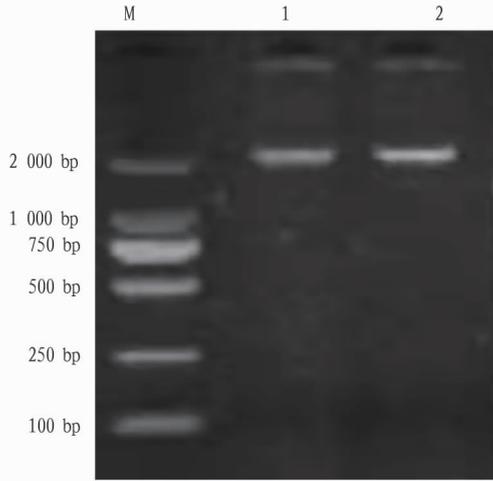
注: M.DL 2000 Marker; 1,2.基因组 DNA

Note: M.DL 2000 Marker; 1,2.Genomic DNA

图 3 麻疯树种子基因组 DNA 电泳

Fig.3 Genomic DNA electrophoresis of *Jatropa curcas* seeds

的 Caleosin 蛋白具有较高的同源性(表 3),与来自蓖麻(*Ricinus communis*)(XP\_002528367.1)、川桑(*Morus notabilis*)(XP\_010089211.1)、爱玉子(*Ficus pumila* var. *Auwekotsang*)(ABV72237.1)、异色山黄麻(*Trema orientalis*)(PON94834.1)等 17 个不同物种间的 Caleosin 蛋白同源相似性均达 70%以



注: M. DL2000 Marker; 1、2. DNA 扩增

Note: M. DL2000 Marker; 1, 2. DNA amplification

图 4 *JcCale26.8* 基因 DNA 序列的 PCR 产物电泳

Fig.4 Electrophoresis of genomic DNA PCR products of *JcCale26.8* gene

上,其中与蓖麻的同源相似度最高,达 87.07%。基于 MEGA-x 软件构建的系统进化树(图 7)显示,爱玉子(ABV72237.1)、杨梅(*Myrica rubra*) (KAB1219333.1)和川桑(XP\_010089211.1)聚为一个分支,且与 *JcCale26.8* 的亲缘关系最远;*JcCale26.8* 与蓖麻(XP\_002528367.1)、野大豆(*Glycine soja*) (KHN30721.1)的 Caleosin 蛋白聚为一个分支,表明 *JcCale26.8* 与蓖麻和野大豆中具有 POG 酶活性的 Caleosin 蛋白亲缘关系较近。

蛋白质理化性质分析结果显示,*JcCale26.8* 蛋白分子式为  $C_{1218}H_{1829}N_{317}O_{350}S_{10}$ , 相对分子量 26.83 kD, 等电点 5.50; *JcCale26.8* 蛋白由除天冬酰胺(Asn, N)外的 19 种常见氨基酸组成,亮氨酸(Leu, L)含量最高为 8.8%,半胱氨酸(Cys, C)含量最低为 0.8%,其他氨基酸含量在 2.1~8.0%,氨基酸组成差异较小;负电荷氨基酸残基(Asp+Glu)有 28 个,正电荷氨基酸残基(Arg+Lys)有 22 个;摩尔消光系数(extinction coefficients)为 45 505,不稳定系数(Instability index)为 52.01。蛋白信号肽分析显示,无信号肽酶切位点(图 8),推测 *JcCale26.8* 属于不稳定的非分泌性酸性蛋白质。

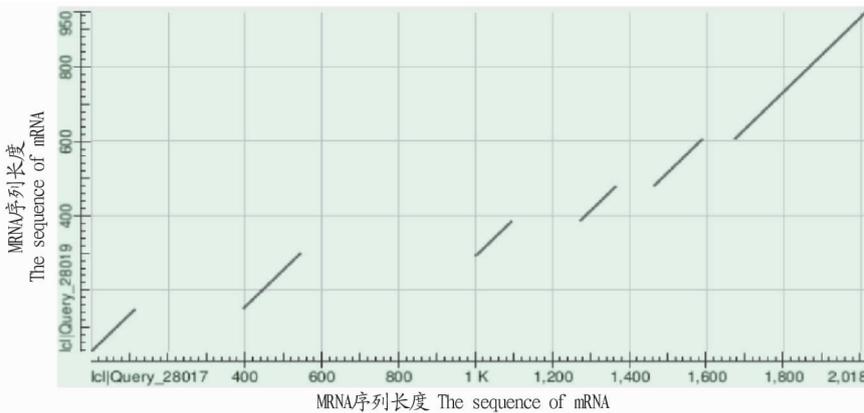


图 5 *JcCale26.8* 基因 DNA 与 mRNA 序列比对分析

Fig.5 Alignment analysis of DNA and mRNA sequence of *JcCale 26.8* gene

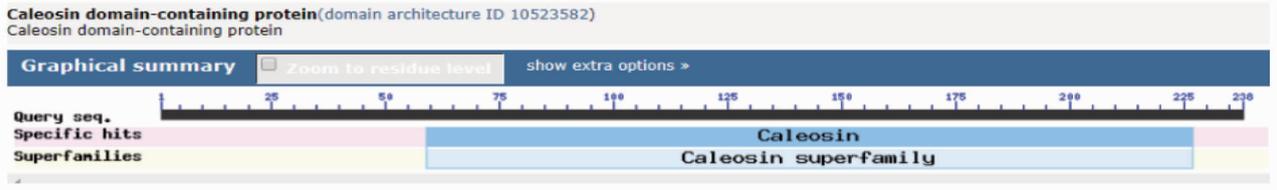


图 6 *JcCale26.8* 蛋白在 CD-Search 服务器上的分析结果

Fig.6 Analysis of *JcCale26.8* protein on CD-Search Server

表 3 麻疯树 *JcCale26.8* 蛋白与其他植物已知 Caleosin 蛋白同源性比较

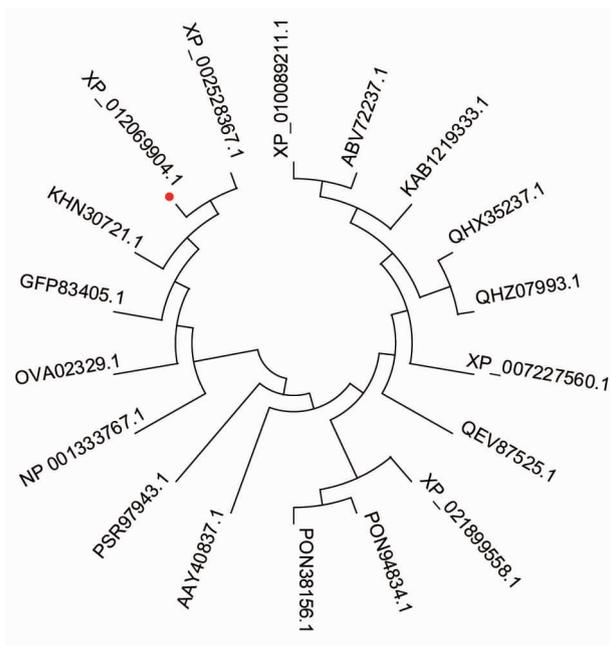
Table 3 Similarity comparison of *JcCale26.8* with Caleosins from other plants known in GenBank

序列号 Sequence ID	同源蛋白 Homologous protein	物种 Species	长度 Length//aa	S 值 Total score	E 值 E value	相似度 Similarity %	覆盖率 Coverage %
XP_002528367.1	Peroxygenase	蓖麻 <i>Ricinus communis</i>	233	436	1e-153	87.07	97
XP_010089211.1	Peroxygenase	川桑 <i>Morus notabilis</i>	239	405	2e-144	79.66	99
ABV72237.1	Caleosin	爱玉子 <i>Ficus pumila</i> var. <i>aukeotsang</i>	239	394	4e-137	79.31	97
PON94834.1	Caleosin-related	异色山黄麻 <i>Trema orientalis</i>	240	399	6e-139	78.11	97
PON38156.1	Caleosin-related	山黄麻 <i>Parasponia andersonii</i>	240	399	8e-139	78.11	97
PSR97943.1	Peroxygenase	中华猕猴桃 <i>Actinidia chinensis</i> var. <i>chinensis</i>	268	396	2e-137	77.64	99

接下表

续表 3

序列号 Sequence ID	同源蛋白 Homologous protein	物种 Species	长度 Length//aa	S 值 Total score	E 值 E value	相似度 Similarity %	覆盖率 Coverage %
QHX35237.1	Caleosin I	欧榛 <i>Corylus avellana</i>	239	393	1e-136	77.31	99
KAB1219333.1	Peroxygenase	杨梅 <i>Myrica rubra</i>	239	399	5e-139	77.31	99
XP_021899558.1	Peroxygenase 1	番木瓜 <i>Carica papaya</i>	240	385	3e-133	77.16	97
QHZ07993.1	Caleosin H1	欧榛 <i>Corylus avellana</i>	239	392	4e-136	76.89	99
AAV40837.1	Caleosin	甘蓝型油菜 <i>Brassica napus</i>	245	378	2e-130	75.00	97
OVA02329.1	Caleosin	博落回 <i>Macleaya cordata</i>	289	382	2e-131	74.57	97
KHN30721.1	Peroxygenase	野大豆 <i>Glycine soja</i>	235	379	6e-131	74.46	97
GFP83405.1	Peroxygenase	松蒿 <i>Phtheirospermum japonicum</i>	230	379	4e-131	74.24	96
XP_007227560.1	Peroxygenase	桃 <i>Amygdalus persica</i>	239	382	4e-132	74.15	99
NP_001333767.1	Caleosin	番茄 <i>Lycopersicon esculentum</i>	245	376	1e-129	73.36	96
QEV87525.1	Caleosin	芍药 <i>Paeonia lactiflora</i>	239	372	3e-128	72.27	99



注:XP\_002528367.1:蓖麻;KHN30721.1:野大豆;GFP83405.1:松蒿;OVA02329.1:博落回;NP\_001333767.1:番茄;PSR97943.1:中华猕猴桃;AAY40837.1:甘蓝型油菜;PON38156.1:*Parasponia andersonii*;PON94834.1:异色山黄麻;XP\_021899558.1:番木瓜;QEV87525.1:芍药;XP\_007227560.1:桃;QHZ07993.1:欧榛;QHX35237.1:欧榛;KAB1219333.1:杨梅;ABV72237.1:爱玉子;XP\_010089211.1:川桑

Note:XP\_002528367.1:*Ricinus communis*;KHN30721.1:*Glycine soja*;GFP83405.1:*Phtheirospermum japonicum*;OVA02329.1:*Macleaya cordata*;NP\_001333767.1:*Lycopersicon esculentum*;PSR97943.1:*Actinidia chinensis* var. *Chinensis*;AAY40837.1:*Brassica napus*;PON38156.1:*Parasponia andersonii*;PON94834.1:*Trema orientalis*;XP\_021899558.1:*Carica papaya*;QEV87525.1:*Paeonia lactiflora*;XP\_007227560.1:*Amygdalus persica*;QHZ07993.1:*Corylus avellana*;QHX35237.1:*Corylus avellana*;KAB1219333.1:*Myrica rubra*;ABV72237.1:*Ficus pumila* var. *Auwekeotsang*;XP\_010089211.1:*Morus notabilis*

图7 麻疯树 *J. curcas* JcCale26.8 蛋白与其他植物 Caleosin 蛋白的进化树  
Fig.7 Evolutionary tree analysis of *J. curcas* JcCale26.8 and caleosins from other plants

疏水作用在维持蛋白质高级结构稳定中发挥着重要作用,较高正值的氨基酸具有较强的疏水性,而较低负值的氨基酸则具有较强的亲水性。通过在线软件 ExPASy 对 JcCale26.8 进行蛋白质亲疏水性分析,结果(图 9)显示,大部分氨基酸为负值,表现为亲水性,尤其是 N-末端、127~150 位氨基酸残基和 C-末端表现为较明显的亲水区,第 46 位氨基酸残基处亲水性最强(-2.144);87~107 位和 182~201 位氨基酸残基处表现出较明显的疏水性,99 和 100 位氨基酸残基疏水性最强(3.067),推测 JcCale26.8 为两性蛋白。基于 TMHMM 在线软件分析结果显示,JcCale26.8 蛋白无明显的跨膜结构域(图 10),推测该蛋白可能以中间疏水区域锚定在油体膜表面而发挥作用。

二级结构预测结果显示,JcCale26.8 主要由无规卷曲(46.64%)和  $\alpha$ -螺旋(35.29%)结构构成,还含有少量的延伸链(10.92%)和  $\beta$ -转角(7.14%)等二级结构。无规卷曲主要位于第 3~7、13~27、31~38、42~59、80~83、107~122、128~138、146~149、160~168、179~184、207~210 位氨基酸残基区域, $\alpha$ -螺旋结构主要位于第 8~12、60~67、84~89、94~105、150~159、169~177、185~207、211~218、234~238 位氨基酸残基区域,延伸链主要位于第 68~70、77~79、103~104、123~127、139~141、198~202、222~225 位氨基酸残基区域, $\beta$ -转角主要位于第 72~75、90~91、143~145、205~206、219~220、231~232 位氨基酸残基区域(图 11)。

蛋白质三级结构预测(图 12)显示,JcCale26.8 主要由 9 个  $\alpha$ -螺旋和无规卷曲结构构成,模型置信度为 97.7%,符合二级结构预测结果。根据三维结构预测结果,并结合亲疏水性、跨膜结构和二级结构预测结果,预测麻疯树 JcCale26.8 蛋白符合油体钙蛋白典型结构特征,包括可结合钙离子的 N 端亲水区,中间含有脯氨酸的疏水区和 C 端的磷酸化位点<sup>[28]</sup>,可能在油体的形成及调动中传递信号,在种子萌发过程中通过信息转运来启动分解油体内中性脂肪产生能量。

### 3 结论与讨论

Caleosin 是一种与植物油体发育发生密切相关的油体结合蛋白之一,其编码基因已从水稻<sup>[7]</sup>、芝麻<sup>[9]</sup>和拟南芥等<sup>[21]</sup>多种植物中克隆得到,不同植物的 Caleosin 蛋白序列具有较

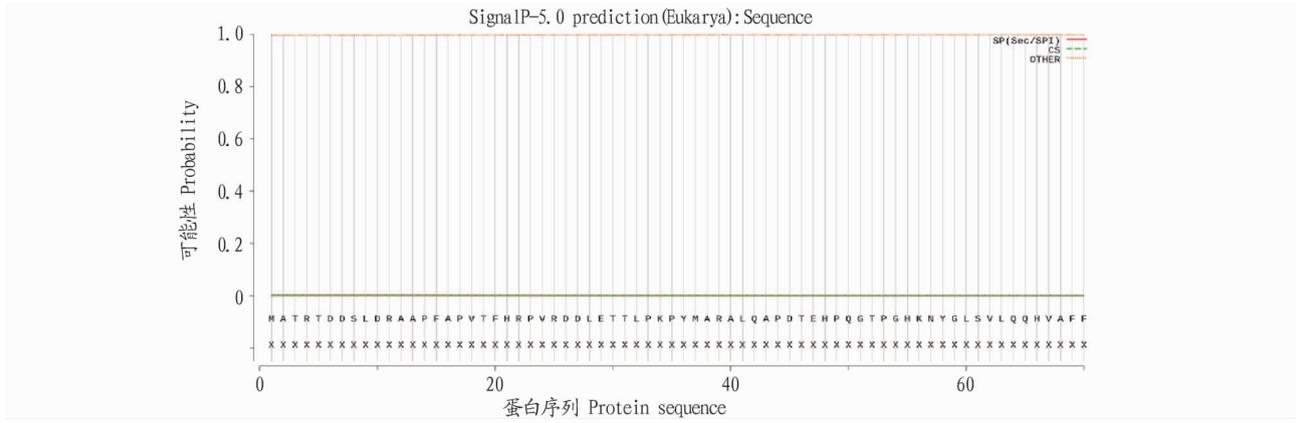


图 8 *JcCale26.8* 蛋白质信号肽分析

Fig.8 Analysis of *JcCale26.8* protein signaling peptides

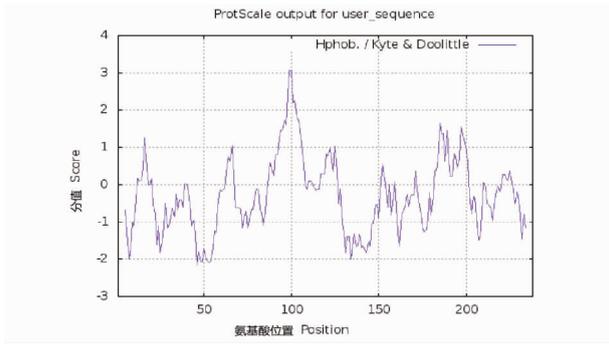


图 9 *JcCale26.8* 蛋白质亲水性和疏水性分析

Fig.9 Prediction of *JcCale26.8* protein hydrophobicity and hydrophilicity profile

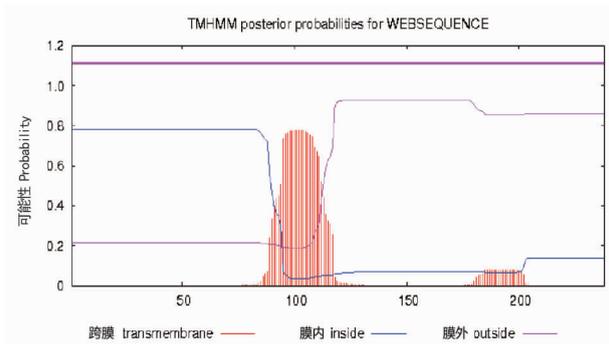
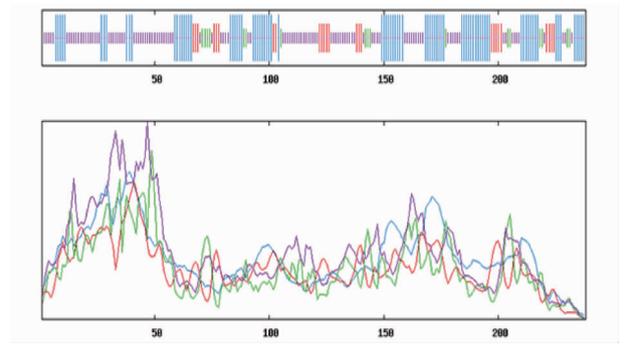


图 10 *JcCale26.8* 蛋白跨膜结构域预测

Fig.10 Prediction of transmembrane regions of *JcCale26.8*

高的同源性,尤其是具有结合单一钙离子的 EF-hand 保守域,并类似主要油体结合蛋白 Oleosin 能与油体相结合等特性<sup>[29]</sup>。该试验通过 RNA-seq 和 PCR 技术从麻疯树成熟种子中成功克隆了编码 26.8 kD Caleosin 蛋白的 *JcCale26.8* 基因 mRNA 和对应基因组 DNA 序列,推测编码的 *JcCale26.8* 蛋白明显具有 Caleosin 蛋白结构特征,并与已知多种植物 Caleosin 蛋白具有较高的序列相似性。

对 *JcCale26.8* 基因序列进行生物信息学分析发现,该序列早期在 Zhang 等<sup>[30]</sup> 对盐胁迫下麻疯树幼苗叶片的基因组测序数据库中 (PRJNA63485) 分析拼接获得,并被注释为未知蛋白 JCGZ\_02382 (KDP40384.1),而 Jalali 等<sup>[31]</sup> 对麻疯树基因组进行测序获得的数据库 (PRJNA673911) 进行分析拼



注:蓝色:α-螺旋;红色:延伸链;紫色:无规卷曲;绿色:β-转角  
Note: Blue. α-Spiral; Red. Extension chain; Purple. Random curl; Green.β-corner

图 11 *JcCale26.8* 蛋白二级结构预测

Fig.11 Prediction of secondary structure of *JcCale26.8* protein

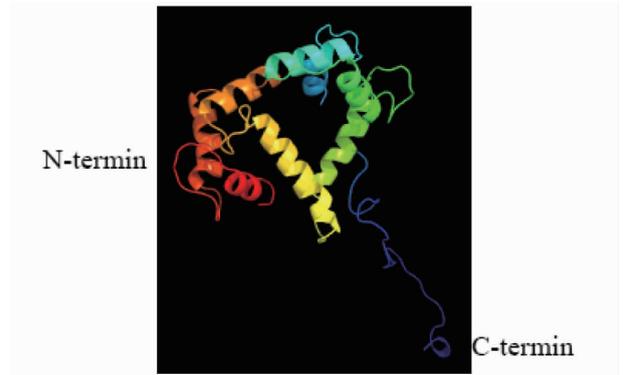


图 12 *JcCale26.8* 蛋白三级结构预测结果

Fig.12 Prediction of tertiary structure of *JcCale26.8* protein

接获得了该序列,并根据同源性自动被注释为 POG (XP\_012069904.1) 基因蛋白,但并未得到研究。有研究发现,Caleosin 在非种子组织中表现出 POG 活性<sup>[32]</sup>,因而目前在 NCBI 数据库中具有 Caleosin 功能结构域的 Caleosin 蛋白序列多数被注释为 POG 蛋白。为了探究麻疯树油体钙蛋白及其基因的结构与功能,笔者从麻疯树种子中克隆分离了 *JcCale26.8* 基因 mRNA 和 DNA 序列,序列特征分析结果表明 *JcCale26.8* 基因表达产物为植物 Caleosin 蛋白家族成员,而在麻疯树中 *JcCale26.8* 基因表达产物 *JcCale26.8* 蛋白是否具

有 POG 酶活性仍需要深入研究。

Caleosin 在植物中通常由多基因家族编码,分子量一般在 25~35 kD。目前已至少在 15 种植物中鉴定出了 84 个 *Caleosin* 基因<sup>[33]</sup>,如拟南芥基因组中鉴定出了 8 个 *Caleosin* 基因<sup>[34]</sup>,榛子 (*Corylus avellana* L.) 基因组中鉴定出 5 个 *Caleosin* 基因<sup>[35]</sup>,油菜 (*Brassica napus*) 中至少包含 25.0、27.0、28.1 kD 的 3 个 *Caleosin* 蛋白成员<sup>[29]</sup>。推测在麻疯树中至少存在 4 个 *Caleosin* 家族异构体蛋白。该试验克隆分离了 26.8 kD 的 *Caleosin* 基因序列,同时该试验麻疯树种子转录组数据库结果中还显示有另一具有 *Caleosin* 功能结构域的 119 个氨基酸残基组成的推测 13.7 kD 的 *Caleosin* 基因序列;根据对 NCBI 登陆的麻疯树基因组数据库 PRJNA673911 显示的另外 2 个 POG 酶蛋白序列 XP\_020534220.1 和 XP\_012072291.2 分析显示,XP\_020534220.1 (NCBI 注释为 peroxxygenase isoform X2) 序列蛋白为 206 个氨基酸残基组成的 22.9 kD 的 *Caleosin*,XP\_012072291.2 (NCBI 注释为 probable peroxygenase 4) 序列蛋白为 249 个氨基酸残基组成的 28.7 kD 的 *Caleosin*。因此,推测以上 4 个序列对应的蛋白均属于 *Caleosin* 超家族成员。这可为后期开展麻疯树 *Caleosin* 家族蛋白的深入研究提供参考。

#### 参考文献

- [1] SHIMADA T L, HAYASHI M, HARA-NISHIMURA I. Membrane dynamics and multiple functions of oil bodies in seeds and leaves[J]. *Plant physiology*, 2018, 176(1): 199–207.
- [2] CHAPMAN K D, DYER J M, MULLEN R T. Biogenesis and functions of lipid droplets in plants[J]. *Journal of lipid research*, 2012, 53(2): 215–226.
- [3] LAIBACH N, POST J, TWYMAN R M, et al. The characteristics and potential applications of structural lipid droplet proteins in plants[J]. *Journal of biotechnology*, 2015, 201: 15–27.
- [4] ITABE H. Intracellular lipid droplet-associated proteins: Unique members and their biological functions[J]. *Biological and pharmaceutical bulletin*, 2010, 33(3): 341.
- [5] CAPUANO F, BEAUDOIN F, NAPIER J A, et al. Properties and exploitation of oleosins[J]. *Biotechnology advances*, 2007, 25(2): 203–206.
- [6] 孙静, 姜宇, 陶俊. 植物油质蛋白的结构、功能及应用[J]. *植物生理学学报*, 2018, 54(3): 363–369.
- [7] FRANSEN G, MÜLLER-URI F, NIELSEN M, et al. Novel plant  $Ca^{2+}$ -binding protein expressed in response to abscisic acid and osmotic stress[J]. *The journal of biological chemistry*, 1996, 271(1): 343–348.
- [8] POXLEITNER M, ROGERS S W, LACEY SAMUEL A, et al. A role for caleosin in degradation of oil-body storage lipid during seed germination[J]. *The plant journal*, 2006, 47(6): 917–933.
- [9] CHEN J C F, TSAI C C Y, TZEN J T C. Cloning and secondary structure analysis of caleosin, a unique calcium-binding protein in oil bodies of plant seeds[J]. *Plant and cell physiology*, 1999, 40(10): 1079–1086.
- [10] JIANG P L, TZEN J T C. Caleosin serves as the major structural protein as efficient as oleosin on the surface of seed oil bodies[J]. *Plant signaling behavior*, 2010, 5(4): 447–449.
- [11] PARTRIDGE M, MURPHY D J. Roles of a membrane-bound caleosin and putative peroxygenase in biotic and abiotic stress responses in *Arabidopsis* [J]. *Plant physiology and biochemistry*, 2009, 47(9): 796–806.
- [12] BLÉE E, BOACHON B, BURCKLEN M, et al. The reductase activity of the *Arabidopsis* caleosin RESPONSIVE TO DESSICATION 20 mediates gib-

- berellin-dependent flowering time, abscisic acid sensitivity, and tolerance to oxidative stress[J]. *Plant physiology*, 2014, 166(1): 109–124.
- [13] 韦剑锋, 韦冬萍, 吴炫柯, 等. 麻疯树种子特性及其影响因素研究进展[J]. *种子*, 2013, 32(2): 51–55.
- [14] LIU H, WANG C P, KOMATSU S, et al. Proteomic analysis of the seed development in *Jatropha curcas*: From carbon flux to the lipid accumulation[J]. *Journal of proteomics*, 2013, 91: 23–40.
- [15] 邓志军, 程红焱, 宋松泉. 麻疯树种子的研究进展[J]. *云南植物研究*, 2005, 27(6): 605–612.
- [16] 郑元, 罗娅娜, 罗玛妮娜, 等. 牛樟芝羊毛甾醇合成酶基因 *AcLSS* 的克隆和表达分析[J]. *西南林业大学学报(自然科学)*, 2021, 41(3): 94–99.
- [17] 阳江华, 张希财, 邹智. 橡胶树捕光叶绿素 a/b 结合蛋白基因 *CAB2* 的克隆与分析[J]. *西南林业大学学报(自然科学)*, 2019, 39(1): 88–94.
- [18] 徐志文, 任雪敏, 赵满, 等. 黄粉甲储存蛋白 *hexamerin* 基因的克隆及表达分析[J]. *西南林业大学学报(自然科学)*, 2019, 39(4): 96–102.
- [19] NæSTED H, FRANSEN G I, JAUH G Y, et al. Caleosins:  $Ca^{2+}$ -binding proteins associated with lipid bodies[J]. *Plant molecular biology*, 2000, 44(4): 463–476.
- [20] JOLIVET P, ROUX E, D' ANDREA S, et al. Protein composition of oil bodies in *Arabidopsis thaliana* ecotype WS[J]. *Plant physiology and biochemistry*, 2004, 42(6): 501–509.
- [21] HYUN T K, KUMAR D, CHO Y Y, et al. Computational identification and phylogenetic analysis of the oil-body structural proteins, oleosin and caleosin, in castor bean and flax[J]. *Gene*, 2013, 515(2): 454–460.
- [22] 熊宏, 陈海涛, 宋健, 等. 麻疯树油质蛋白 *JcOle16.6* 基因克隆及序列分析[J]. *江苏农业科学*, 2016, 44(6): 84–89.
- [23] 宋健, 熊宏, 余进德, 等. 麻疯树油质蛋白 *JcOle14.3* 基因克隆及序列分析[J]. *中南林业科技大学学报*, 2016, 36(6): 15–22.
- [24] 余进德, 熊宏, 宋健, 等. 麻疯树种子发育过程中 *JcOle14.3* 和 *JcOle16.6* 基因的表达模式研究[J]. *广西植物*, 2017, 37(9): 1096–1100.
- [25] 徐秋红, 章镇, 佟兆国, 等. 山梨醇对李果肉组织总 RNA 提取的影响[J]. *江苏农业学报*, 2010, 26(2): 390–394.
- [26] 张君堂, 陶承光, 王志刚, 等. *Trizol-A<sup>+</sup>* 试剂法提取百合总 RNA[J]. *江苏农业科学*, 2009(2): 35–36.
- [27] 丁勇, 李涛, 龚秀会, 等. 3 种杜仲枝叶总 RNA 提取方法的比较[J]. *江苏农业科学*, 2011, 39(5): 29–31.
- [28] 向兰舟, 胡娅晴, 谢涵, 等. 拟南芥油体钙蛋白基因 *CALEOSIN 3* 对胁迫环境及 ABA 诱导的响应[J/OL]. *分子植物育种*, 2021-01-22 [2021-04-27]. <http://kns.cnki.net/kcms/detail/46.1068.S.2021.20210122.1218.004.html>.
- [29] 丁勇, 陈庆波, 徐春雷, 等. 油菜油体钙蛋白基因 *BnClo1* 的克隆和表达[J]. *作物学报*, 2008, 34(11): 1921–1928.
- [30] ZHANG L, ZHANG C, WU P Z, et al. Global analysis of gene expression profiles in physic nut (*Jatropha curcas* L.) seedlings exposed to salt stress[J]. *PLoS One*, 2014, 9(5): 1–9.
- [31] JALALI K, KANCHARLA N, YEPURI V, et al. Exploitation of Hi-C sequencing for improvement of genome assembly and *in-vitro* validation of differentially expressing genes in *Jatropha curcas* L. [J]. *3 Biotech*, 2020, 10(5): 307–321.
- [32] HANANO A, ALKARA M, ALMOUSALLY I, et al. The peroxygenase activity of the *Aspergillus flavus* caleosin, *AfPXG*, modulates the biosynthesis of aflatoxins and their trafficking and extracellular secretion via lipid droplets[J]. *Frontiers in microbiology*, 2018, 9: 1–19.
- [33] SONG W L, QIN Y J, ZHU Y, et al. Delineation of plant caleosin residues critical for functional divergence, positive selection and coevolution[J]. *BMC evolutionary biology*, 2014, 14(1): 1–14.
- [34] SHEN Y, XIE J, LIU R D, et al. Genomic analysis and expression investigation of caleosin gene family in *Arabidopsis* [J]. *Biochemical and biophysical research communications*, 2014, 448(4): 365–371.
- [35] LAMBERTI C, NEBBIA S, BALESTRINI R, et al. Identification of a caleosin associated with hazelnut (*Corylus avellana* L.) oil bodies[J]. *Plant biology*, 2020, 22(3): 404–409.