

熊 宏,陈海涛,宋 健,等. 麻风树油质蛋白 *JcOle16.6* 基因克隆及序列分析[J]. 江苏农业科学,2016,44(6):84-89.
doi:10.15889/j.issn.1002-1302.2016.06.021

麻风树油质蛋白 *JcOle16.6* 基因克隆及序列分析

熊 宏,陈海涛,宋 健,余进德,丁 勇

(西南林业大学生命科学学院,云南昆明 650224)

摘要:分离克隆麻风树油质蛋白基因家族成员 *JcOle16.6* 基因的 DNA 序列(GenBank 序列号为 JX073622.1)。*JcOle16.6* 基因全长 601 bp,没有内含子,转录的 *JcOle16.6* mRNA 具有 468 bp 的完整开放阅读框,编码含 155 个氨基酸残基、分子量为 16.6 ku、等电点为 9.99 的 *JcOle16.6* 蛋白(GenBank 序列号为 AFP19884),属于植物油体结合蛋白油质蛋白家族的新成员。*JcOle16.6* 属于稳定的两性蛋白质,具有 3 个结构域,即 N 端约 30 个氨基酸残基组成的含 1 个 α -螺旋的亲水性结构域,肽链中间约 78 个氨基酸残基组成的含 3 个 α -螺旋的高度疏水性跨膜结构域,C 端约 47 个氨基酸组成的含 1 个 α -螺旋的两亲性结构域。N 端亲水性结构域和 C 端两亲性结构域分布于油体朝向胞浆一面,中间疏水跨膜结构域分布于油体半单位膜内,存在于中间疏水跨膜结构域中的由 3 个脯氨酸和 1 个丝氨酸组成的 Pro-Knot 高度保守结构域在 *JcOle16.6* 蛋白与油体半单位膜准确定位和稳定维持油体结构方面起着重要作用。研究结果为 *JcOle16.6* 基因表达调控和生理功能等研究奠定了基础。

关键词:麻风树;种子;油体;油质蛋白;基因克隆;序列分析

中图分类号: Q785 **文献标志码:** A **文章编号:** 1002-1302(2016)06-0084-06

可作为现代生物质能源的植物油广泛存在于植物种子中^[1-2],植物油是植物种子最有效的能量储备形式,其主要以三酰甘油(triglycerides,TAGs)液态基质形式存在于植物亚细胞器颗粒-油体中,可为随后的代谢过程和其他生命活动提供能量和碳水化合物^[3-4]。研究认为,TAGs 位于油体内部,油体外部为由磷脂单分子层(phospholipids,PL)和油体结合蛋白组成的半单位膜结构^[5]。油体结合蛋白包括大量的油质蛋白和微量的油体钙蛋白、油体固醇蛋白 Steroleosin-A 和 Steroleosin-B^[6-9]。研究推测油质蛋白作为最丰富的油体结合蛋白在油体的发生到分解消失过程中起着重要的生物学功能^[1,10-11]。目前在油桐(*Verniciafordii*)^[12]、油菜(*Brassica campestris*)^[13-16]、芝麻(*Sesamumindicum*)^[17-18]和拟南芥(*Arabidopsis thaliana*)^[19]等物种中油质蛋白基因已被克隆,其生物学功能也得到了研究,同时发现在各物种中油质蛋白以多个异构体蛋白形式存在。

麻风树(*Jatropha curcas* L.)别称小桐子或膏桐,是大戟科(Euphorbiaceae)多年生落叶灌木或小乔木,其果实采摘期长达 50 年,种仁的含油率高达 60%~70%^[20-21],麻风树已经作为生物质能源植物得到了一定的开发与利用^[22-24]。研究麻风树油质蛋白及其基因将有助于揭示麻风树种子油体形成与稳定机制,并可为麻风树种子在新型生物能源及植物生物反应器方面深入开发与利用提供理论依据。本研究从

GenBank 数据库中搜索得到已知登录的麻风树油质蛋白家族基因 mRNA 全长序列(GenBank 序列号为 EU234462.2、EU234463.2、EU234464.2),但有关麻风树油质蛋白基因 DNA 序列克隆和序列分析的相关报道还未发现。笔者所在课题组根据 EU234464.2 序列设计特异性引物,克隆了麻风树油质蛋白家族基因 *JcOle14.3* 在染色体上的 DNA 序列(GenBank 序列号为 JX073623.1)。本研究根据 EU234463.2 序列设计特异性引物,克隆麻风树油质蛋白家族基因 *JcOle16.6* 在染色体上的 DNA 序列(GenBank 序列号为 JX073622.1),并对该基因序列及其推测编码的 *JcOle16.6* 蛋白进行生物信息学分析。

1 材料与方法

1.1 材料与试剂

试验材料为三年生麻风树植株,取其未成熟种子和幼叶于液氮中速冻后置于 -80℃ 冰箱中保存备用。

植物总 RNA 提取试剂盒(RNAiso Plus)、M-MLV 反转录试剂盒和 pMD18-T 载体均为 TaKaRa 产品,植物基因组 DNA 提取试剂盒为北京天根生物技术有限公司产品,胶回收试剂盒为 Omega 产品,大肠杆菌 DH5 α 为实验室保存菌种,其他化学试剂均为生工生物工程(上海)股份有限公司产品。

1.2 试验方法

1.2.1 总 RNA 提取与检测 麻风树幼叶基因组 DNA 提取按照植物基因组 DNA 提取试剂盒说明书完成,麻风树种子总 RNA 提取参照丁勇等方法^[25]完成;DNA 完整性和纯度分别采用 1.0% 琼脂糖凝胶电泳和紫外分光光度计检测,RNA 完整性和纯度分别采用 1.2% 琼脂糖凝胶电泳和紫外分光光度计检测。

1.2.2 麻风树 *JcOle16.6* 基因克隆 在已知麻风树油质蛋白基因 mRNA 全长序列(GenBank 序列号为 EU234463.2)的

收稿日期:2015-10-28

基金项目:国家自然科学基金(编号:31460076);云南省高校优势特色重点学科建设项目(编号:50097505);云南省高校林下生物资源保护及利用科技创新团队项目。

作者简介:熊 宏(1994—),男,江西宜春人,硕士研究生,研究方向为代谢途径与分子生物学。E-mail:1570530221@qq.com。

通信作者:丁 勇,高级实验师,硕士生导师,研究方向为分子生物学。E-mail:dingyong@swfu.edu.cn。

ORF 区外侧设计引物,上游引物为 JcOle16.6F;5-CTTCTC-CACACTTATCATCAAC-3,下游引物为 JcOle16.6R;5-TA-AACCAACTCAACTACCTCAA-3,并由生工生物工程(上海)股份有限公司完成序列合成。以提取的总 RNA 为模板,采用反转录试剂盒合成第一链 cDNA,取第一链 cDNA 2 μ L,上下游引物各 1 μ L,2 \times Taq PCR Master Mix 12.5 μ L,ddH₂O 8.5 μ L,建立 25 μ L 的 RT-PCR 反应体系,克隆目的基因 cDNA 序列,反应条件为:94 $^{\circ}$ C 预变性 3 min;94 $^{\circ}$ C 变性 30 s,48 $^{\circ}$ C 退火 30 s,72 $^{\circ}$ C 延伸 1 min,35 个循环;72 $^{\circ}$ C 延伸 10 min。取 2 μ L DNA 为模板,上下游引物各 1 μ L,2 \times Taq PCR Master Mix 12.5 μ L,ddH₂O 8.5 μ L,建立 25 μ L 的 PCR 反应体系,克隆目的基因在染色体上的 DNA 序列,反应条件为:94 $^{\circ}$ C 预变性 4 min;94 $^{\circ}$ C 变性 30 s,48 $^{\circ}$ C 退火 30 s,72 $^{\circ}$ C 延伸 1 min,35 个循环;72 $^{\circ}$ C 延伸 10 min。

1.2.3 PCR 产物检测、回收、T/A 克隆与测序 应用 1.2% 琼脂糖凝胶电泳检测 PCR 产物,目的片段回收后连接到 pMD18-T 载体上,将重组子转化大肠杆菌 DH5 α 感受态细胞,转化复苏后的菌液涂布在含有 Amp、IPTG 和 X-Gal 的 LB 固体培养基上,过夜培养后随机筛选白色菌落,将阳性克隆菌液送往生工生物工程(上海)股份有限公司完成测序。

1.2.4 序列的生物信息学分析 生物信息学分析参照已知文献报道的方法^[15,26-27]进行,包括 mRNA 及推导的蛋白质氨基酸序列相似性搜索,mRNA 开放阅读框查找,蛋白理论分子量、等电点、氨基酸含量和稳定性分析,系统发生树构建,蛋白质亲水/疏水性特性、跨膜结构、信号肽及功能域预测,目的蛋白二级、三级结构预测。

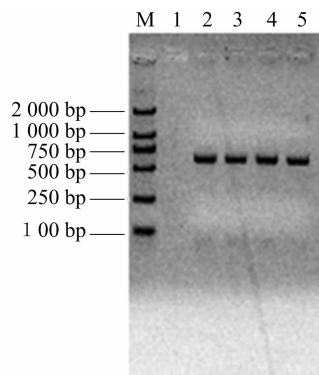
2 结果与分析

2.1 麻风树油质蛋白 JcOle16.6 基因克隆

本研究以提取的总 RNA 和基因组 DNA 为模板分别进行 RT-PCR 和普通 PCR 试验,均获得 1 条位于 500~750 bp 区间的特异性 cDNA 和 DNA 条带(图 1),将目的 cDNA 和 DNA 条带进行胶回收、TA 克隆和测序,结果表明,本研究克隆到的麻风树油质蛋白基因 cDNA 序列长 601 bp,序列与 GenBank 数据库中登录的麻风树油质蛋白基因 mRNA 序列 EU234463.2 具有 100% 的相似性,本研究克隆到的麻风树油质蛋白基因 DNA 序列也为 601 bp,序列与 601 bp 的 cDNA 序列完全一致。根据生物信息学分析结果,本研究将该目的基因命名为麻风树油质蛋白 JcOle16.6 基因,其 601 bp 的 DNA 序列登录 GenBank 的序列号为 JX073622。表明 JcOle16.6 基因只包含 1 个外显子,没有内含子。

2.2 序列分析

本研究克隆获得麻风树 JcOle16.6 基因 601 bp 的基因组 DNA 序列,其转录的 601 bp 的 mRNA 具有完整的 ORF,位于第 25 至第 492 bp 区域,第 1 至第 24 bp 区域为 5'UTR,第 493 至第 601 bp 区域为 3'UTR。ORF 推测编码的麻风树油质蛋白



M—DL2000 分子量标准;1—阴性空白对照;
2~3—JcOle16.6 基因 cDNA 产物;
4~5—JcOle16.6 基因组 DNA 产物

图1 麻风树 JcOle16.6 基因克隆产物的琼脂糖凝胶电泳

白 JcOle16.6 (GenBank 序列号为 AFP19884) 由 155 个氨基酸残基组成,相对分子质量为 16.6 ku,带负电荷的氨基酸残基(天冬氨酸+谷氨酸)为 9 个,带正电荷的氨基酸残基(精氨酸+赖氨酸)为 16 个,pI 为 9.99,推测 JcOle16.6 在生理条件下为碱性蛋白。除谷氨酰胺和半胱氨酸之外,JcOle16.6 含有其他 18 种常见的氨基酸残基,色氨酸含量最低,为 0.6%,天冬氨酸、酪氨酸含量也较低,均为 1.3%,亮氨酸含量最高,为 12.9%,其他氨基酸含量在 2.6%~10.3% 之间。因其具有 38.97 的不稳定系数值,推测 JcOle16.6 为稳定性蛋白质。

蛋白质相似性分析结果见(图 2)显示,麻风树 JcOle16.6 含有“油质蛋白”保守结构域,属于油质蛋白超家族蛋白之一。JcOle16.6 蛋白与油桐(*Vernicia fordii*)的油质蛋白 II、胡杨(*Populus euphratica*)的油质蛋白 18.2 ku-like、毛果杨(*P. trichocarpa*)油质蛋白家族蛋白、蓖麻(*Ricinus communis*)的油质蛋白 1、川桑(*Morus notabilis*)的油质蛋白 5、爱玉子(*Ficus pumila* var.)的油质蛋白高分子量异构体、醉蝶花(*Tarenaya hassleriana*)的油质蛋白 5 异构体、亚麻(*Linum usitatissimum*)的油质蛋白高分子量异构体、胡萝卜(*Daucus carota*)的脂质体膜蛋白、油茶(*Camellia oleifera*)的 Ole IV 蛋白、烟草(*Nicotiana glauca*)的油质蛋白 5-like、甜橙(*Citrus sinensis*)的油质蛋白 18.2 ku-like 等都具有较高的相似性,E 值从 9×10^{-75} ~ 3×10^{-56} ,一致性 60%~83%,相似度 74%~91%(表 1)。但 JcOle16.6 蛋白与该课题组克隆的另一个麻风树油质蛋白家族基因 JcOle14.3 (GenBank 序列号为 JX073623)编码的异构体 JcOle14.3 蛋白(GenBank 登录号为 AFP19885)之间的相似度较低,以致分析结果未能显示出。表明本研究克隆的 JcOle16.6 基因编码的 JcOle16.6 蛋白属于麻风树油质蛋白异构体之一。

将 13 种不同植物的油质蛋白进行序列比对分析,结果见图 3,显示麻风树 JcOle16.6 蛋白与其他的油质蛋白在肽链中间序列具有较高的一致性,而在肽链两端序列却存在较大差异性。1 个丝氨酸和 3 个脯氨酸组成的脯氨酸结(Pro-Knot)

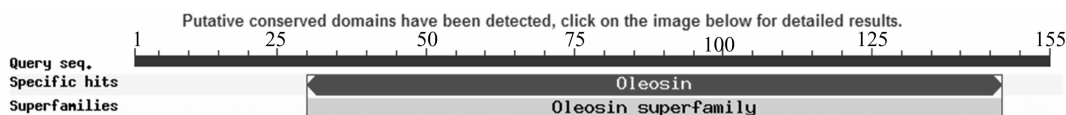


图2 JcOle14.3 蛋白序列的 Blastp 分析结果

表 1 JcOle16.6 蛋白与其他已知功能蛋白同源性比较

GenBank 序列号	同源蛋白	拉丁名	长度 (个)	S 值	E 值	一致性 (%)	相似度 (%)	间隔 (%)
ADB03185.1	油质蛋白 II	<i>Vernicia fordii</i>	154	232 bits (592)	9×10^{-75}	83	91	1
XP_011005135.1	油质蛋白 18.2 ku - like	<i>Populus euphratica</i>	149	231 bits (588)	2×10^{-74}	77	85	8
XP_002323804.1	油质蛋白 family protein	<i>Populus trichocarpa</i>	149	226 bits (575)	3×10^{-72}	73	83	8
XP_002521458.1	油质蛋白 1	<i>Ricinus communis</i>	153	218 bits (556)	2×10^{-69}	69	84	6
XP_010108818.1	油质蛋白 5	<i>Morus notabilis</i>	157	210 bits (535)	3×10^{-66}	66	80	4
ABQ57396.1	油质蛋白 H - isoform	<i>Ficus pumila</i> var.	155	205 bits (521)	4×10^{-64}	63	80	2
XP_010524296.1	油质蛋白 5 isoform X1	<i>VTarenaya hassleriana</i> V	161	196 bits (497)	2×10^{-60}	62	79	3
ABB01616.1	油质蛋白 high molecular weight isoform	<i>Linum usitatissimum</i>	180	196 bits (499)	2×10^{-60}	63	74	21
AAB24078.1	Lipid body membrane protein	<i>Daucus carota</i>	180	191 bits (485)	2×10^{-58}	63	75	5
ABF57561.1	OleIV	<i>Camellia oleifera</i>	154	190 bits (483)	3×10^{-58}	63	77	2
XP_009772501.1	油质蛋白 5 - like	<i>Nicotiana glauca</i>	166	187 bits (476)	3×10^{-57}	60	74	8
XP_006470260.1	油质蛋白 18.2 ku - like	<i>Citrus sinensis</i>	166	185 bits (470)	3×10^{-56}	62	75	7

为已知植物油质蛋白序列的标志性结构^[12,28],该保守结构域也存在于麻风树 JcOle16.6 和其他参比的油质蛋白序列中。

油质蛋白进化树分析结果见图 4, 大戟科麻风树 *JcOle*16.6 和大戟科油桐油质蛋白 II (ADB03185.1) 亲缘关系最近, 聚为一支, 和大戟科蓖麻油质蛋白 I (EEF40948.1) 聚为一个大的支; 同样来自麻风树的 *JcOle*14.3 (AFP19885.1) 和油质蛋白 I-like (NP_001295613.1) 聚为一支, 二者与爱玉子

(*Ficus pumila* var.) 的油质蛋白高分子量异构体 (ABQ57396.1)、醉蝶花 (*Tarenaya hassleriana*) 的油质蛋白 5 异构体 (XP_010524296.1) 属于一个大分支, 但与 JcOle16.6 的系统发育进化距离较远。推测来自不同植物的同一种油质蛋白异构体基因具有较近的进化关系, 而来自同一植物的不同油质蛋白异构体基因具有较远的进化关系。

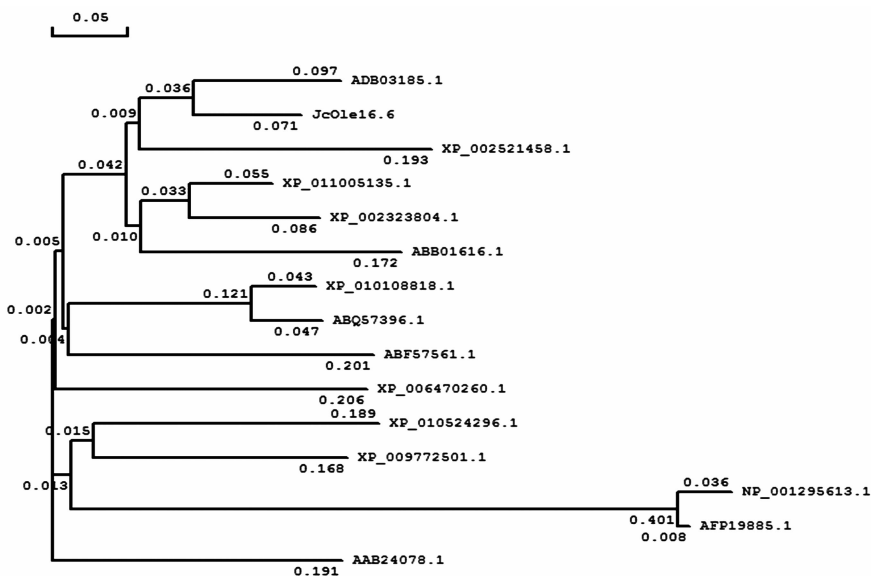
ADB03185.1	MADR...LQHQCVCVHPQLR.....YD.A...YK.GQCKRSTSKVLAVITLLEGGGLLALSGITIVG	56
XP_011005135.1	MADR...MQHQCVCV.GLK.....GH.....CQHGFSAKRALAVITLLEGGGLLALSGITIVG	51
XP_002323804.1	MADR...KQHQCVCV.GLR.....GQ.....CQCGFSASKLAVLIMPEVGGGLLALSGITIVG	51
XP_002521458.1	MADR...PQHQCVCV.RYD.....FTTGG...YK.GQCKGFSASKVLAVITLLEGGGLLALSGITITIN	56
XP_010108818.1	MADR...PQHQCVCVHPQVG.....AGKGFK...QR.GEQCGFSAGKILAVITLLEGGTLLFLAGLITIG	59
ABQ57396.1	MADR...PQHQCVCVHPQY.....GKGFK...QR.GEQCGFSAGKILAVITLLEGGTLLFLAGLITIG	57
XP_010524296.1	MADTVR.TQHQCVCVHPQR.....CHEGGVK...TL.LPGRGFSATQLVAVITLLEGGTLLFLAGLITIG	60
ABB01616.1	MADR.TTQHQCVCVTHGHYFTGGAGFRYEGGGLKGGPHHGGGSGFSASKLAVMTALHEGTTLLALACITILAG	74
AAB24078.1	MAERG.TYAHQCVCVHPQTA.....NCPGGVK...SL.LPKNESTSQLAVVITLLEGGTLLALACITIVG	62
ABF57561.1	MADR....FHQCVCVHPHHR.....YDGGKL...TR....CDGSEFSKLVAMVALHEGTTLLALACITIVG	55
XP_009772501.1	MADTVR.RHQCVCVHPQHF.....RYEGGVK...TL.LPARGFSATQLAVITLLEGGTLLALACITILG	60
XP_006470260.1	MAERDRPQHQCVCVHPQGH.....GKSLMG...QRGAAGGFSASKLAVLALHEGTTLLALACITIVG	62
JcOle16.6	MAER...SQHQCVCVHPQHR.....YEAA...FK.GQCKGFSACKVLAVITLLEGGGLLALSGITIVG	57
Consensus	ma hq p Knot ps la p g g	
ADB03185.1	TLIGLATITLHFLFFSFVLPFARFVLGLAVMTFLASGAGLTGLTSLSVWLKYYVQEATR...NIPEQLCAHRR	127
XP_011005135.1	TLIGLATITLHFLIFSFVLPFARFVLGLAVTSFLASGAGLTGLTSSWLSVWGXYQEATQ...TMFESLDEAHR	122
XP_002323804.1	TLIGTVTLHFLFFLSFVLPFARFLIGFAVTSFLASGALGTGLTGRFSVWVRYQEATR...TMFENLDCAKRC	122
XP_002521458.1	TLIGLATITLHFLIFGTLIFPAAFVVLGLAMAFEMVAGLGTGLTSGSWALKYFREGT...AMPESLDCAKRR	126
XP_010108818.1	TLIGLATITLHFLVLLSFVLPFARITLIGLIAAGLTGLGAGFVIGLSSLSWVFNFKR.T...TVFEQLFAHRR	128
ABQ57396.1	TLIGLATITLHFLVLLSFVLPFARITLIGLIAAGLTGLGAGFVIGLSSLSWVFNFKR.T...TVFEQLYAHRR	126
XP_010524296.1	SVIGLMVTLHFLVIFSFVLPFARLAVGLAVMTFLASGAGLTGLSSLSWVFNLYNLSRAKE...AVFEQILYAHRR	131
ABB01616.1	TMIGLATITLHFLVIFCSFVLPFARITLIGLAVSAFLASGAGLTGLTSLFWARYLQCGAGGVGVGVDFSEFCARH	149
AAB24078.1	TLIGLATITLHFLVLLSFVLPFARITLIGLATVTLGFLGAGLTGLTSSLSWVYFQRASQ...RVFQDILAKHR	133
ABF57561.1	TLIGLATITLHFLIICSFVLPFARITATGLAVAGILTSGAGLTGLSWLIRVVMYLRGVGA...SVEQQLDEAHR	126
XP_009772501.1	TLIGLATITLHFLVLLSFVLPFARITLAVLAVGTGLTSGAGLTGLSSLSWVIMYFKQG.R...TVFEQLYHRR	130
XP_006470260.1	TLIGCVTLHFLIIFSFVLPFARITLAVLAVGTGLTSGAGLTGLSSLSWVLSIRQGT...SVFEMADCAHRR	133
JcOle16.6	TLIGLATITLHFLVIFSFVLPFARFVLGLSVMAFLASGAGLTGLSSLSWVLKYYLQEVTR...RMFEQLIHAKRR	128
Consensus	ig p f p paa g g	
ADB03185.1	MQD...MAGYVGQKTRVGEICRGAHEGK.....	154
XP_011005135.1	MQD...MAGYVGQKTRVGEICRGAHEGK.....	149
XP_002323804.1	MQD...MAGYVGQRAKEVGEICRGAHEGK.....	149
XP_002521458.1	MQD...MAGYVGKTRVGEICDRCGCEGK.....	153
XP_010108818.1	VQD...LQGVTKRTKAECEICGSGTCEGK...RT.....	157
ABQ57396.1	VQD...MAGYVGQRAKEVGEICVSGGCEGK...RT.....	155
XP_010524296.1	MAG...VQGVGSKTRVGEICSIIEAHEVQIST.....	161
ABB01616.1	MQD...MAGYVGQKTRVGEICRQASQVRASEK.....	160
AAB24078.1	AGE...MAGYVGQKTRVGEITIGSAQAAGQTITATTGDRTRSTARDTSR.....	179
ABF57561.1	MQD...VQVCGKTRVGEIAICMAEAG.K.....	154
XP_009772501.1	MQERMADAACVQKTRVGEICACISACGEGKEGGR.....	166
XP_006470260.1	VAG...MADYVGQKTRVGEICDRCGSGVHEAGGKTGR.....	166
JcOle16.6	MQD...MAGYVGQKTRVGEICRGAHEGK.....	155
Consensus	a g k g k	

ADB03185.1, 油桐(*Vernicia fordii*)的脂质蛋白 II; XP_011005135.1, 胡杨(*P. euphratica*)的脂质蛋白 18.2 ku-like; XP_002323804.1, 毛果杨(*P. trichocarpa*)脂质蛋白家族; XP_002521458.1, 蓖麻(*Ricinus communis*)的脂质蛋白 I; XP_010108818.1, 川桑(*Morus notabilis*)的脂质蛋白 5; ABQ57396.1, 爱玉子(*Ficus pumila* var.)的脂质蛋白 H-isoform; XP_010524296.1, 醉蝶花(*Tarenaya hassleriana*)的脂质蛋白 5 isoform X1; ABB01616.1, 亚麻(*Linum usitatissimum*)的脂质蛋白 high molecular weight isoform; AAB24078.1, 胡萝卜(*Daucus carota*)Lipid body membrane蛋白; ABF57561.1, 油茶(*Camellia oleifera*)的Ole IV蛋白; XP_009772501.1, 烟草(*Nicotiana sylvestris*)的脂质蛋白 5-like; XP_006470260.1, 甜橙(*Citrus sinensis*)的脂质蛋白 18.2 ku-like

图3 JcOle16.6 蛋白与其他植物油质蛋白的序列比对

信号肽分析结果未发现可能的信号肽序列存在于 JcOle16.6 蛋白中,表明该蛋白为分泌性蛋白。疏水性分析结果见图 5,显示 JcOle16.6 蛋白质肽链在 N 端 1~31 位和 C

端 109~155 位氨基酸残基区域都表现为高度亲水性,而中间 32~108 位氨基酸残基区域表现为高度疏水性,表明麻风树 JcOle16.6 蛋白为两亲性蛋白质,即蛋白的两末端区域为亲水



ADB03185.1, 油桐 (*Vernicia fordii*) 的油质蛋白 II; XP_002521458.1, 蓖麻 (*Ricinus communis*) 的油质蛋白 I; XP_011005135.1, 胡杨的油质蛋白 18.2 ku-like; XP_002323804.1, 毛果杨油质蛋白家族; ABB01616.1, 亚麻的油质蛋白 high molecular weight isoform; XP_010108818.1, 川桑的油质蛋白 5; ABQ57396.1, 爱玉子的油质蛋白 H-isoform; XP_010524296.1, 醉蝶花的油质蛋白 5 isoform X1; XP_009772501.1, 烟草的油质蛋白 5-like; AAB24078.1, 胡萝卜 Lipid body membrane; ABF57561.1, 油茶的 Ole IV 蛋白; XP_006470260.1, 甜橙的油质蛋白 18.2 ku-like; AFP19885.1, 麻风树 14.3 ku 油质蛋白; NP_001295613.1, 麻风树油质蛋白 1-like

图4 JcOle16.6 与其他油质蛋白类似蛋白的进化树分析

性,中间区域为疏水性。蛋白质跨膜结构分析结果见图 6,显示 JcOle16.6 蛋白在中间疏水区域存在 2 个潜在的跨膜螺旋结构域,第 49 ~ 72 位的 24 个氨基酸残基形成第 1 个跨膜螺旋结构域,方向为从膜外到膜内螺旋;第 80 ~ 101 位的 22 个氨基酸残基形成第二个跨膜螺旋结构域,方向为从膜内到膜外螺旋。表明麻风树 JcOle16.6 蛋白质为含有 2 个跨膜螺旋结构域的两亲性蛋白质,中间高度疏水区域位于油体半单位膜内,N - 末端和 C - 末端 2 个高度亲水区域位于油体半单位膜外,即位于胞质内。

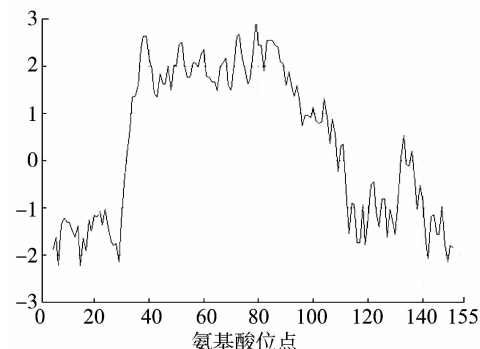


图5 JcOle16.6 蛋白质疏水性和亲水性分析结果

功能蛋白质分子在生物体内都会执行特定的生命活动,其多肽链翻译后通常折叠和盘曲成比较稳定的二级结构,并进一步形成高级结构,完成活性功能域构象的构建。二级结构分析结果见图 7,显示麻风树 JcOle16.6 蛋白可能含有 3 种二级结构类型,分别为 α - 螺旋占 72.26%、随机卷曲占 23.87%,少量延伸链占 3.87%。可见 α - 螺旋为 JcOle16.6 蛋白主要二级结构,约占整个肽链的 3/4,分别位于 18 ~ 22 位 ($\alpha 1$)、31 ~ 42 位 ($\alpha 2$)、45 ~ 68 位 ($\alpha 3$)、75 ~ 93 位 ($\alpha 4$)、

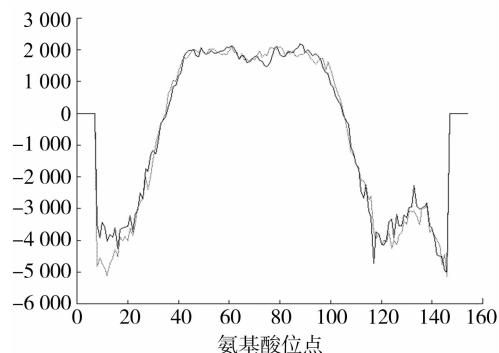
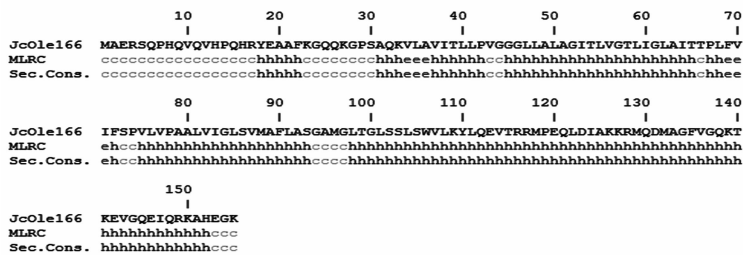


图6 JcOle16.6 蛋白质跨膜结构分析结果

98 ~ 152 位 ($\alpha 5$) 氨基酸残基区域。推测 JcOle16.6 蛋白 α - helix 的形成可能与预测的 2 个跨膜螺旋结构域的形成及与油体半单位膜的定位结合有着密切的关系。JcOle16.6 蛋白二级结构,尤其是 α - 螺旋结构在应用 I - TASSER 程序预测的高级结构(图 8)中可清晰显示出。GO 功能分类结果显示 JcOle16.6 与磷脂单分子层膜的油脂储藏体 (GO:0012511) 相关,并能整合到单分子层膜上 (GO:0016021)。因此,根据蛋白质疏水性、跨膜结构、二级结构、高级结构和功能分析结果,推测 JcOle16.6 蛋白在结构上可分为 3 个部分:(1)N 端约 30 个氨基酸残基组成的亲水性区域,分布于油体朝向胞浆一面, $\alpha 1$ 位于此结构域;(2)C 端约 47 个氨基酸组成的 α - 螺旋两亲性结构域,也分布于油体朝向胞浆一面, $\alpha 5$ 位于此结构域;(3)肽链中间约 78 个氨基酸残基组成的高度疏水性跨膜区域,分布于油体半单位膜内, $\alpha 2$ 、 $\alpha 3$ 、 $\alpha 4$ 位于此结构域, $\alpha 3$ 、 $\alpha 4$ 之间存在着由 3 个脯氨酸和 1 个丝氨酸组成的 Pro - Knot 高度保守结构域(图 8),这与 JcOle16.6 蛋白向油体半单位膜的正确定位和油体结构的稳定维持密切相关。



h— α -螺旋；e—延伸链；c—随机卷曲
图7 JcOle16.6 蛋白二级结构预测

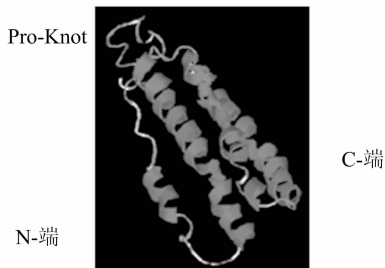


图8 麻风树 JcOle16.6 蛋白质三维结构预测

3 结论与讨论

油质蛋白是与植物油体特异性结合的主要蛋白质,在油料植物种子中普遍存在^[29],在植物种子成熟过程中油质蛋白几乎覆盖整个油体表面,使得油体虽以互相挤压形式充满整个细胞但从不相互聚合,推测油质蛋白与植物油体的形成和稳定有着密切的关系^[30-32]。油质蛋白及其基因已在多种植物中得到了一定的研究^[33-37]。本研究分离克隆了麻风树油质蛋白基因家族成员 *JcOle16.6* 基因,长 601 bp 的 DNA 序列 (GenBank 序列号为 JX073622.1), *JcOle16.6* 基因在染色体水平上没有内含子,转录的 *JcOle16.6* mRNA 具有 468 bp 的完整开放阅读框,编码含 155 个氨基酸残基、分子量为 16.6 ku 的 *JcOle16.6* 蛋白 (GenBank 序列号为 AFP19884),属于植物油体结合蛋白油质蛋白家族的新成员。

不同植物来源的油质蛋白在分子量上具有较大的变化。研究表明,特异性存在于植物种子中的油质蛋白分子量在 15~26 ku 之间变化^[2]。本研究从麻风树种子中分离克隆的 *JcOle16.6* 基因表达的 *JcOle16.6* 蛋白分子量也在此范围之内,为 16.6 ku,符合 Anthony 和 Huang 的分析结果。在蛋白质结构上 *JcOle16.6* 可分为 3 个结构区域,即 N 端约 30 个氨基酸残基组成的含 1 个 α -螺旋的亲水性结构域,肽链中间约 78 个氨基酸残基组成的含 3 个 α -螺旋的高度疏水性跨膜结构域,C 端约 47 个氨基酸组成的含 1 个 α -螺旋的两亲性结构域。N 端亲水性结构域和 C 端两亲性结构域均分布于油体朝向胞浆一面,中间疏水跨膜结构域分布于油体半单位膜内,存在于中间疏水跨膜结构域中的由 3 个脯氨酸和 1 个丝氨酸组成的 Pro-Knot 高度保守结构域在 *JcOle16.6* 蛋白与油体半单位膜准确定位和稳定维持油体结构方面起着重要作用。与相关研究关于油质蛋白结构和功能的论述^[35]一致,表明油质蛋白基因在不同植物中表达的油质蛋白能形成类似的结构特性并发挥相同的生物学功能。

油质蛋白基因常以为为数不多的基因家族形式存在。在裸子植物中只发现 1 种油质蛋白^[36], 在已研究的被子植物中发现在同一植物中存在氨基酸序列和分子量大小都明显不同的多种油质蛋白异构体^[37~38]。刘玉君等对麻风树种子油体蛋白进行电泳分析, 发现分子量分别为 20、24 ku 的 2 种油质蛋白异构体可能存在于麻风树种子油体中^[39]。本研究从麻风树种子中分离克隆的 *JcOle16.6* 基因产物推测为 16.6 ku 的油质蛋白, 同时笔者从麻风树种子中成功克隆的 *JcOle14.3* 基因产物推测为 14.4 ku 的油质蛋白。推测麻风树油质蛋白基因也以基因家族形式存在, 其表达产物可能以不同分子量油质蛋白异构体形式同时存在于同一油体中。

油质蛋白作为植物油体中发现最早、含量最丰富的油体结合蛋白被认为是需要长期贮藏的油体所必需的^[40],推测从油体的发生到分解消失过程中发挥重要的生物学功能。研究推测油质蛋白在植物种子发育及成熟过程中直接影响油体的形成、稳定性以及大小^[41],在种子萌发、油体动员与解体过程中可作为酯酶识别并结合油体的特殊信号^[42-43]。但JcOle16.6及其他异构体蛋白在麻风树种子油体的发生到分解消失过程中的具体生物学功能还有待于进一步研究阐明。

本研究分离克隆了麻风树油质蛋白基因家族成员 *JcOle16.6* 基因长 601 bp 的 DNA 序列 (GenBank 序列号为 JX073622.1), *JcOle16.6* 基因没有内含子, 转录的 *JcOle16.6* mRNA 具有 468 bp 的完整开放阅读框, 编码含 155 个氨基酸残基、分子量为 16.6 ku 的 *JcOle16.6* 蛋白 (GenBank 序列号为 AFP19884), 属于植物油体结合蛋白油质蛋白家族的新成员。*JcOle16.6* 具有 3 个结构域, 即 N 端约 30 个氨基酸残基组成的含 1 个 α -螺旋的亲水性结构域, 肽链中间约 78 个氨基酸残基组成的含 3 个 α -螺旋的高度疏水性跨膜结构域, C 端约 47 个氨基酸组成的含 1 个 α -螺旋的两亲性结构域。N 端亲水性结构域和 C 端两亲性结构域分布于油体朝向胞浆一面, 中间疏水跨膜结构域分布于油体半单位膜内, 存在于中间疏水跨膜结构域中的由 3 个脯氨酸和 1 个丝氨酸组成的 Pro-Knot 高度保守结构域在 *JcOle16.6* 蛋白与油体半单位膜准确定位和稳定维持油体结构方面起着重要作用。

参考文献:

- [1]Huang A H. Oil bodies and oleosins in seeds[J]. Annual Review of Plant Biology,1992,43(1):177-200.
- [2]Huang A H. Oleosins and oil bodies in seeds and other organs[J]. Plant Physiology,1996,110(4):1055-1061.
- [3]Yatsu L Y,Jacks T J. Spherosome membranes:half unit-membranes

- [J]. *Plant Physiology*, 1972, 49(6): 937–943.
- [4] Murphy D J. Structure, function and biogenesis of storage lipid bodies and oleosins in plants[J]. *Progress in Lipid Research*, 1993, 32(3): 247–280.
- [5] Tzen J, Cao Y, Laurent P, et al. Lipids, proteins, and structure of seed oil bodies from diverse species[J]. *Plant Physiology*, 1993, 101(1): 267–276.
- [6] Tzen J T, Lai Y K, Chan K L, et al. Oleosin isoforms of high and low molecular weights are present in the oil bodies of diverse seed species[J]. *Plant Physiology*, 1990, 94(3): 1282–1289.
- [7] Chen E C, Tai S S, Peng C C, et al. Identification of three novel unique proteins in seed oil bodies of sesame[J]. *Plant & Cell Physiology*, 1998, 39(9): 935–941.
- [8] Lin L J, Tzen J T. Two distinct steroleosins are present in seed oil bodies[J]. *Plant Physiology and Biochemistry*, 2004, 42(7/8): 601–608.
- [9] 丁勇, 徐春雷, 甘莉. 植物油体及其相关蛋白的研究进展[J]. *华中农业大学学报*, 2008, 27(4): 558–563.
- [10] Frandsen G I, Mundy J, Tzen J T. Oil bodies and their associated proteins, oleosin and caleosin[J]. *Physiologia Plantarum*, 2001, 112(3): 301–307.
- [11] Mayfield J A, Fiebig A, Johnstone S E, et al. Gene families from the *Arabidopsis thaliana* pollen coat proteome[J]. *Science*, 2001, 292(5526): 2482–2485.
- [12] 龙洪旭, 谭晓风, 陈洪, 等. 油桐油体蛋白基因的克隆及序列分析[J]. *中南林业科技大学学报*, 2010, 30(4): 31–38.
- [13] Murphy D J, Cummins I, Kang A S. Synthesis of the major oil-body membrane protein in developing rapeseed embryos[J]. *Biochemical Journal*, 1989, 258(1): 285–293.
- [14] Keddie J S, Hübner G, Slocumbe S P, et al. Cloning and characterisation of an oleosin gene from *Brassica napus*[J]. *Plant Molecular Biology*, 1992, 19(3): 443–453.
- [15] 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.
- [16] 丁勇, 常玮, 刘小焯. 甘蓝型油菜 *BnClo1* 基因克隆、表达载体的构建及原核表达[J]. *中国农业科学*, 2010, 43(2): 252–258.
- [17] Chen J C, Tsai C C, Tzen J T. Secondary structure analysis of caleosin. A unique calcium-binding protein in oil bodies of plant seed[J]. *Plant & Cell Physiology*, 1999, 40(10): 1079–1086.
- [18] Lin L J, Tai S S, Peng C C, et al. Steroleosin, a sterol-binding dehydrogenase in seed oil bodies[J]. *Plant Physiology*, 2002, 128(4): 1200–1211.
- [19] Naested H, Frandsen 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] Openshaw K. A review of *Jatropha curcas*: an oil plant of unfulfilled promise[J]. *Biomass & Bioenergy*, 2000, 19(1): 1–15.
- [21] 王曦, 龙春林. 云南小桐子资源调查与评价[J]. *云南植物研究*, 2009, 31(5): 455–460.
- [22] Augustus G D, Jayabalan M, Seiler G J. Evaluation and bioinduction of energy components of *Jatropha curcas*[J]. *Biomass & Bioenergy*, 2002, 23(3): 161–164.
- [23] Berchmans H J, Hirata S. Biodiesel production from crude *Jatropha curcas* L. seed oil with a high content of free fatty acids[J]. *Bioresource Technology*, 2008, 99(6): 1716–1721.
- [24] Fairless D. Biofuel: the little shrub that could – maybe[J]. *Nature*, 2007, 449(7163): 652–655.
- [25] 丁勇, 范红波, 张高磊, 等. 麻风树种子总 RNA 提取方法研究[J]. *中南林业科技大学学报*, 2012, 32(3): 158–161.
- [26] 杜维, 丁勇, 朱东阳, 等. 樟叶越橘熊果苷合成酶基因 *VdASI* 的克隆及序列分析[J]. *植物分类与资源学报*, 2015, 37(1): 71–77.
- [27] 宋健, 熊宏, 朱东阳, 等. 樟叶越橘糖基转移酶 *VdUGT1* 基因克隆及序列分析[J]. *中南林业科技大学学报*, 2015, 35(6): 80–86.
- [28] Vance V B, Huang A H. The major protein from lipid bodies of maize, characterization and structure based on cDNA cloning[J]. *Journal of Biological Chemistry*, 1987, 262(23): 11275–11279.
- [29] Bowman V B, Huang V, Huang A H. Expression of lipid body protein gene during maize seed development, spatial, temporal, and hormonal regulation[J]. *The Journal of Biological Chemistry*, 1988, 263(3): 1476–1481.
- [30] Tzen J T, Huang A H. Surface-structure and proper-ties of plant seed oil bodies[J]. *Journal of Cell Biology*, 1992, 117(2): 327–335.
- [31] Leprince O, van Aelst A C, Pritchard H W, et al. Oleosins prevent oil-body coalescence during seed imbibition as suggested by a low-temperature scanning electron microscope study of desiccation-tolerant and -sensitive oilseeds[J]. *Planta*, 1997, 204(1): 109–119.
- [32] Jiang P L, Wang C S, Hsu C M, et al. Stable oil bodies sheltered by a unique oleosin in lily pollen[J]. *Plant & Cell Physiology*, 2007, 48(6): 812–821.
- [33] Lee K, Huang A H. Genomic nucleotide sequence of a brassica napus 20-kilodalton oleosin gene[J]. *Plant Physiology*, 1991, 96(4): 1395–1397.
- [34] Li M, Smith L J, Clark D C, et al. Secondary structures of a new class of lipid body proteins from oilseeds[J]. *The Journal of Biological Chemistry*, 1992, 267(12): 8245–8253.
- [35] Lacey D J, Wellner N, Beaudoin F, et al. Secondary structure of oleosins in oil bodies isolated from seeds of safflower (*Carthamus tinctorius* L.) and sunflower (*Helianthus annuus* L.)[J]. *Biochemical Journal*, 1998, 334(Pt 2): 469–477.
- [36] Wu L S, Hong G H, Hou R F, et al. Classification of the single oleosin isoform and characterization of seed oil bodies in gymnosperms[J]. *Plant & Cell Physiology*, 1999, 3(40): 326–334.
- [37] Tzen J T, Wang M M, Tai S S, et al. The abundant proteins in sesame seed; storage proteins in protein bodies and oleosins in oil bodies[J]. *Adv Plant Physiology*, 2003, 6: 93–105.
- [38] Tzen J T, Chuang R L, Chen J C, et al. Coexistence of both oleosin isoforms on the surface of seed oil bodies and their individual stabilization to the organelles[J]. *Journal of Biochemistry*, 1998, 123(2): 318–323.
- [39] 刘玉君, 沈世华. 小桐子种子油体蛋白的提取及其电泳分析[J]. *林业科学*, 2008, 44(8): 37–41.
- [40] Murphy D J, Vance J. Mechanisms of lipid-body formation[J]. *Trends in Biochemical Sciences*, 1999, 24(3): 109–115.
- [41] Siloto R M, Findlay K, Lopez-Villalobos A, et al. The accumulation of oleosins determines the size of seed oilbodies in *Arabidopsis*[J]. *Plant Cell*, 2006, 18(8): 1961–1974.
- [42] Tzen J T, Peng C C, Cheng D J, et al. A new method for seed oil body purification and examination of oil body integrity following germination[J]. *Journal of Biochemistry*, 1997, 121(4): 762–768.
- [43] Thompson J E, Froese C D, Madey E, et al. Lipid metabolism during plant senescence[J]. *Progress in Lipid Research*, 1998, 37(2/3): 119–141.