

独行菜磷酸甲羟戊酸激酶 LaPMK 基因克隆、生物信息学分析及原核表达

赵 乐^{1,2}, 马利刚^{1,2}, 杨方方¹, 冯卫生^{1,2}, 郑晓珂^{1,2*}

1. 河南中医学院药学院, 河南 郑州 450046

2. 呼吸疾病诊疗与新药研发河南省协同创新中心, 河南 郑州 450046

摘要: 目的 从独行菜 *Lepidium apetalum* 中克隆萜类合成途径关键酶磷酸甲羟戊酸激酶(phosphomevalonate kinase, PMK)基因, 进行序列分析和原核表达。方法 根据独行菜转录组数据中 LaPMK 基因序列设计特异性引物, 克隆得到 LaPMK 基因的开放阅读框(open reading frame, ORF), 构建 pET32a-LaPMK 原核表达载体, 在大肠杆菌 BL21 (DE3) 中用 IPTG 诱导表达 LaPMK 融合蛋白。结果 LaPMK 基因 ORF 全长为 1 518 bp (Genbank 注册号 KT004541), 编码 505 个氨基酸。生物信息学分析结果显示 LaPMK 蛋白没有跨膜区, 不含信号肽, 位于细胞质中, 含有 GHMP 激酶家族特异的 N 末端和 C 末端的保守结构域。系统进化树结果显示 LaPMK 蛋白与芜菁的 PMK 蛋白具有 92% 的序列相似性, 亲缘关系较近。构建 pET32a-LaPMK 原核表达载体, 在大肠杆菌中成功表达 LaPMK 融合蛋白。结论 克隆了 LaPMK 基因, 建立其稳定的原核表达体系, 为 LaPMK 蛋白纯化、抗体的制备, 进一步研究 LaPMK 基因在独行菜萜类合成途径中的作用奠定了基础。

关键词: 独行菜; 磷酸甲羟戊酸激酶; 萜类合成; 基因克隆; 原核表达

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2016)17-3087-07

DOI: 10.7501/j.issn.0253-2670.2016.17.021

Cloning, bioinformatic analysis, and prokaryotic expression of LaPMK gene in *Lepidium apetalum*ZHAO Le^{1,2}, MA Li-gang^{1,2}, YANG Fang-fang¹, FENG Wei-sheng^{1,2}, ZHENG Xiao-ke^{1,2}

1. School of Pharmacy, Henan University of Traditional Chinese Medicine, Zhengzhou 450046, China

2. Collaborative Innovation Center for Respiratory Disease Diagnosis and Treatment & Chinese Medicine Development of Henan Province, Zhengzhou 450046, China

Abstract: Objective To obtain the key enzyme gene involved in terpenoid biosynthesis pathway, phosphomevalonate kinase (PMK) gene was cloned from *Lepidium apetalum*, and sequence analysis and prokaryotic expression were performed. **Methods** Based on the transcriptome data of *L. apetalum*, by designing specific primers of LaPMK gene, an open reading frame (ORF) of LaPMK gene was isolated from *L. apetalum*. *Escherichia coli* BL21 (DE3) cells were transformed with the prokaryotic expression vector pET32a-LaPMK and used for prokaryotic expression under IPTG induction. **Results** LaPMK gene has ORF of 1 518 bp (GenBank accession number KT004541), which encoded a protein of 505 amino acid residues. Bioinformatic analysis indicated that LaPMK protein which located in cytoplasm had no transmembrane domain and signal peptide, and exhibited the specific N-terminal domain and C-terminal domain of GHMP kinase super family. Phylogenetic analysis indicated that LaPMK protein showed the highest homology, 92% similarity, with PMK protein from *Brassica rapa*. The recombinant LaPMK protein was successfully expressed in *E. coli* BL21 (DE3) cells. **Conclusion** The LaPMK gene is cloned from *L. apetalum*, and the stable prokaryotic expression system of pET32a-LaPMK is constructed. This study will provide the fundamental information for the further purification and the antibody preparation of LaPMK protein be helpful for the functional researches of LaPMK gene in terpenoid biosynthesis pathway.

Key words: *Lepidium apetalum* Willd.; phosphomevalonate kinase; terpenoid biosynthesis; gene cloning; prokaryotic expression

独行菜 *Lepidium apetalum* Willd. 为十字花科 子为传统中药北葶苈子, 具有泻肺平喘、利水消肿、
(Cruciferae) 独行菜属 *Lepidium* L. 植物, 其成熟种 强心之功效, 含有强心苷、黄酮、三萜、甾醇、香

收稿日期: 2016-02-13

基金项目: 国家重点基础研究发展计划“973 计划”项目(2013CB531802); 河南中医学院博士科研基金(BSJJ2011-07)

作者简介: 赵 乐(1983—), 男, 博士, 讲师, 主要从事药用植物分子生物学研究。E-mail: zhaole1983@126.com

*通信作者 郑晓珂, 女, 博士, 教授, 博士生导师, 主要从事中药活性成分及作用机制研究。E-mail: zhengxk.2006@163.com

豆素、生物碱、酚酸、挥发油等多种成分,对心血管、高血脂等疾病具有显著功效,具有降压、抗癌、细胞毒、抗氧化等药理活性^[1]。独行菜强心苷类化学成分伊夫单苷^[2]等,对其进行结构分析可知属于植物甾醇(五环内酯),目前,对植物甾醇和三萜类皂苷的生物合成途径已有初步认识,从生物合成来源分析,植物甾醇与三萜类似,在生物体内也是由鲨烯以不同方式环化而成,即通过甲羟戊酸途径(MVA)衍生而来^[3]。

MVA 途径存在于细胞质中,是植物萜类物质合成的主要途径。磷酸甲羟戊酸激酶(phosphomevalonate kinase, PMK)是 GHMP 激酶超家族中的一员^[4],是萜类化合物生物合成 MVA 途径中的一个重要限速酶,在这个途径中有 3 个连续依赖 ATP 的限速反应,其中 PMK 催化第 2 个限速反应,其作用是将 ATP 上的 γ 磷酸基团转移到甲羟戊酸-5-磷酸(MVAP)上生成甲羟戊酸-5-二磷酸(MVAPP),是控制整个代谢途径的关键酶之一^[5-6]。目前,关于植物中 PMK 基因的研究较少,杨恩泽等^[7]首次从药用植物阳春砂中克隆了 AvPMK 基因,并使用生物信息学的方法初步分析了酶结构特点和功能。对于哺乳类动物的 PMK 基因研究的较多,如果 PMK 基因突变将可能致使哺乳动物产生甲羟戊酸缺乏症(导致动物夭折),或产生半乳糖激酶缺乏症(导致动物产生白内障)^[8]。

在对本实验室前期获得的大量独行菜转录组数据的分析中,发现多个参与 MVA 途径的基因,本研究首次从独行菜中克隆了 PMK 基因(LaPMK)的 cDNA 序列,进行生物信息学分析和原核表达,为进一步研究 PMK 基因在独行菜萜类合成途径中的作用奠定了基础。

1 材料与方法

1.1 材料

样品采自河南省伏牛山区,由河南中医学院董诚明教授和谢小龙博士鉴定独行菜 *Lepidium apetalum* Willd. 种子,种子采集后,种植于人工气候培养箱中,光照强度为 3 000 lx,在 16 h、23 °C 光照,8 h、20 °C 黑暗,70%相对湿度条件下培养。种子萌发,长出 4 片真叶后,移至培养罐,每个培养罐放置 4 颗幼苗,每隔 3 d 更换 1 次 Hoagland 营养液,培养 60 d 的独行菜幼苗,取根、茎、叶,采集后样品经液氮速冻,保存于-80 °C 超低温冰箱中备用。本研究所用的实验材料为独行菜幼苗叶片。

1.2 试剂

植物总 RNA 提取试剂盒、质粒小提试剂盒、PCR 产物纯化试剂盒购自北京天根生化科技公司;TransScript II Reverse Transcriptase、大肠杆菌 *Escherichia coli* DH5 α 感受态细胞和 BL21 感受态细胞购自北京全式金生物技术公司;Ex Taq 酶、限制性核酸内切酶、T₄ DNA 连接酶、pMD19-T vector 购自 Takara 公司;引物由北京三博远志生物技术公司合成。一般试剂均参照《分子克隆实验指南》配制,所使用化学试剂均为分析纯试剂。

1.3 RNA 提取及 cDNA 合成

取 100 mg 独行菜幼苗的叶片,在液氮中研磨,根据天根公司植物总 RNA 提取试剂盒说明书提取独行菜叶片的总 RNA,用 NanoDrop 2000 和 1% 琼脂糖凝胶电泳检测 RNA 的量和完整性。用全式金公司的 TransScript II Reverse Transcriptase 试剂盒反转录合成 cDNA。

1.4 LaPMK 基因克隆

根据已获得的独行菜转录组数据,用 Primer 5 设计 LaPMK 基因的特异性引物,正向引物:LaPMK-Exp-F(5'-CGGAATTCATGGCTGTAGTTGCT-3',下划线部分为 *Eco*RI 酶切位点);反向引物:LaPMK-Exp-R(5'-CCCAAGCTTTTACTCGAGACGAAT3',下划线部分为 *Hind* III 酶切位点)。PCR 反应体系为 cDNA 模板 1 μ L, LaPMK-Exp-F 1 μ L, LaPMK-Exp-R 1 μ L, 10 $\times$ Ex Taq buffer 5 μ L, 2.5 mmol/L dNTP 4 μ L, Ex Taq 酶 0.25 μ L,加水至 50 μ L。反应程序为 95 °C 预变性 2 min,然后进行 30 个循环:95 °C 变性 10 s,55 °C 退火 15 s,72 °C 延伸 1 min 40 s,最后 72 °C 延伸 8 min。PCR 产物用 1% 琼脂糖凝胶电泳检测分析,用 PCR 产物纯化试剂盒纯化后与 pMD19-T vector 连接,转化大肠杆菌 DH5 α 感受态细胞,在氨苄抗性平板上进行筛选,经菌液 PCR 和电泳检测后,挑选阳性克隆送北京三博公司测序。

1.5 LaPMK 基因的生物信息学分析

LaPMK 基因编码的蛋白质使用 ExPASy Proteomics Server 提供的在线工具 ProtParam (<http://web.expasy.org/protparam/>) 预测其相对分子质量与理论等电点(PI),使用 InterPro Scan (<http://www.ebi.ac.uk/interpro/>) 进行蛋白的保守结构域预测,使用 SWISS-MODEL (<http://swissmodel.expasy.org/>) 进行三维同源建模,使用 TargetP

1.1Server (<http://www.cbs.dtu.dk/services/TargetP/>) 分析蛋白质亚细胞定位, SignalP 4.1 server (<http://www.cbs.dtu.dk/services/SignalP/>) 分析蛋白质信号肽, 使用 TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) 进行跨膜区预测。

将 LaPMK 基因编码蛋白质的氨基酸序列提交到美国国立生物技术信息中心 NCBI 的蛋白质序列数据库进行 BLASTP 搜索相似序列 (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), 从搜索结果中选取与 LaPMK 基因编码的蛋白质相似性较高的其他植物的蛋白质序列作为参考 (包括模式植物拟南芥、水稻、玉米, 药用植物丹参、三七、长春花等植物), 使用 MEGA4 软件的相邻连接法 (neighbor-joining) 构建系统进化树^[9]。

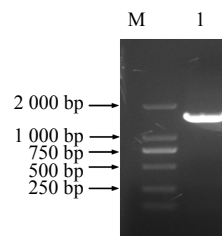
1.6 LaPMK 原核表达载体的构建与诱导表达

用限制性内切酶 *Eco*RI 和 *Hind*III 分别对原核表达载体 pET-32a (+) 和测序正确的 pMD19-T-LaPMK 质粒进行双酶切, 琼脂糖凝胶电泳检测后, 切胶回收载体片段和目的基因片段, *T*₄DNA 连接酶 16 °C 连接过夜。将连接产物转化 *E. coli* BL21 (DE3) 感受态细胞, 在氨苄抗性平板上进行筛选, 挑取单克隆, 经菌液 PCR 和双酶切鉴定后送北京三博公司测序。将测序正确的单克隆, 接种于含氨苄青霉素的 LB 液体培养基中, 37 °C 培养过夜后, 按照 1:100 比例稀释到含氨苄青霉素的 LB 液体培养基中, 220 r/min 振荡培养至对数生长期 [(600 nm 吸光度 (A_{600}) 至 0.6], 在 0.4 mmol/L IPTG, 28 °C, 150 r/min 条件下培养 8 h, 诱导独行菜 LaPMK 重组蛋白的表达。诱导完成后, 用 SDS-PAGE 检测重组蛋白的表达。

2 结果与分析

2.1 LaPMK 的基因克隆

通过分析本实验室前期获得的独行菜转录组数据, 发现多个参与 MVA 途径的基因, 其中有一个注释为 PMK 的转录本, 长度为 1 925 bp, 利用 NCBI ORF Finder 分析其含有一个完整的 ORF, 根据该基因开放阅读框 (ORF) 序列设计特异性引物 LaPMK-Exp-F 和 LaPMK-Exp-R, 以独行菜叶片总 RNA 反转录得到的 cDNA 为模板, PCR 扩增得到 1 条 1 500 bp 左右的片段, 电泳结果如图 1 所示, 与预期大小相符, 进行测序后得到独行菜 LaPMK 基因的编码序列, 大小为 1 518 bp, 编码 505 个氨基酸 (图 2), 序列信息已经提交到 NCBI Genbank, 登录号为 KT004541。



M-Marker 1-LaPMK 基因 PCR 扩增产物

M-Marker 1-PCR product of LaPMK gene

图 1 PCR 扩增 LaPMK 基因

Fig. 1 PCR amplification of LaPMK gene

2.2 LaPMK 基因编码蛋白的生物信息学分析

LaPMK 基因预测编码 505 个氨基酸 (图 2), 利用 ExPASy Proteomics Server 的在线工具 ProtParam 对其编码蛋白的理化性质进行预测分析。推测分子式为 $C_{2411}H_{3840}N_{650}O_{742}S_{20}$, 相对分子质量 54 450, pI: 5.49; 正电残基 (Arg+Lys): 44, 负电残基 (Asp+Glu): 56。不稳定指数为 33.13, 说明 LaPMK 蛋白较稳定, 亲水性总平均系数为 -0.073, 说明 LaPMK 蛋白为亲水性蛋白。

利用 InterPro Protein 进行 LaPMK 蛋白的保守结构域预测 (图 3), 结果显示 LaPMK 蛋白含有 5 种保守结构域, 分别在氨基酸序列的 1~495 位是磷酸甲羟戊酸激酶 (IPR016005)、2~259 位是核糖体蛋白质-5S 区域 (2 类型折叠, IPR020568)、2~109 和 157~257 位是核糖体蛋白质-5S 区域 (2 类型折叠小组) (IPR014721)、179~250 位是 GHMP 激酶 N 端区域 (IPR006204)、397~467、295~350 与 391~484 位是 GHMP 激酶 C 端区域 (IPR003750), 含有 GHMP 激酶家族特异的 N 末端和 C 末端的保守结构域。

SignalP 4.1 server 预测 (图 4), LaPMK 蛋白不含信号肽, 为非分泌蛋白, TargetP 1.1 server 预测, 显示 LaPMK 蛋白可能定位于细胞质, TMHMM server 预测该蛋白不含跨膜域 (图 5)。用 PredictProtein 对 LaPMK 蛋白的二级结构进行预测, 结果显示在该蛋白中 α -螺旋结构占 37.82%, β -折叠占 14.65%, 无规则卷曲占 47.52%。用 SWISS-MODEL 以哺乳动物甲羟戊酸激酶蛋白 (PDB ID: 1kvk) 为模板, 预测 LaPMK 蛋白的三维结构, 序列相似度为 20.4%, 在第 3~495 位氨基酸建模, 模型覆盖率 72%。从预测的三维模型中可以看到 LaPMK 可能以二聚体形式发挥作用, ATP 结合位点位于 LaPMK 蛋白三维结构所形成的一个口袋状结构中 (图 6), 推测 LaPMK 催化的底物磷酸化反应可能在这个口袋状结构中进行^[7]。

```

1 ATGGCTGTAGTTGCTTCTGCCCTGGGAAGGTGTTAATGACTGGAGGCTACCTCATACTCGAGAGGCCAAATACA
1 M A V V A S A P G K V L M T G G Y L I L E R P N T
76 GGGCTTGTGTTGAGTACAAATGCACGGTTTATGCGATCGTGAAGCCATAAATGAAGAACTCAAGCCTGATAGT
7 G L V L S T N A R F Y A I V K P I N E E V K P D S
26 TGGGCGTGGCAATGGACGGATGTCAAATTAACATCACCTCAGCTTTCGAGAGAAAGCATGTATAAACTGTCACCT
51 W A W Q W T D V K L T S P Q L S R E S M Y K L S L
226 AAACATCTGACTCTCAGTGCGTGTCTCAAGTGACTCAAAAAACCCCTTGTAGAGCATGTATTTCAGTATGCT
76 K H L T L Q C V S S S D S K N P F V E H A I Q Y A
301 ATAGCTGCTGCTCATTGGCATCTGGCAAGGGCAAGGAAGCACTACAAAACTCTCTTGAAGGACTTGATATC
101 I A A A H L A S G K G K E A L Q K L F L Q G L D I
376 ACAATCTTAGGCTCTAAGGACTTTTACTCATATCGGAATCAGATAGAATCGCTTGGGCTACCATTAAACACAGAA
126 T I L G S N D F Y S Y R N Q I E S L G L P L T P E
451 TCACTGGGTACCCCTTGCACCTTTTGCATCAATCATTCAATGCTGCCGAGTCAAATGGTGCTAATGCAAGCCT
151 S L G T L A P F A S I T F N A A E S N G A N C K P
526 GAAGTAGCAAAAACCTGGCTTAGGTTCTTCTGCGCAATGACACAGCTGTGGTTGCAGCTTTGTTACATTATCTT
176 E V A K T G L G S S A A M T T A V V A A L L H Y L
601 GGAGTGGTTGACCTATCTGATCCATGTAAAGAAAGGAAATTTTGGTTGTTCTGATCTAGATGTTATCCATGATGA
201 G V V D L S D P C K E G N F G C S D L D V I H M V
676 GCACAAACATCTCATTGTCTTGACACAGGAAGGTGCGGAAGTGGGTTTGATGTGCTGTCTATGGGAGT
A Q T S H C L A Q G K V G S G F D V S C A V Y G S
226 CAACGTATGTTCTGCTTCTCTCCAGAAAGTCTTGCTATTGCTCAGGTTGAAGTGAAGGCTGCGCTTACATGAA
751 Q R Y V R F S P E V L S F A Q V E V K G L P L H E
251 GTTACGGGCAGAAATTTGAAGGGAATTTGGGACAATGAGAGAACCAGTTCTCTTTACCGCCACTGATGAATCTT
826 V T G R I L K G N W D N E R T E F S L P P L M N L
276 GTTCTTGGGGAACCTGGAAGTGGTGGATCCTCCACCCATCGATGGTAGGTGCTGTAAAAAAGTGGCAAAATGTCT
901 V L G E P G S G G S S T P S M V G A V K K W Q M S
301 GATCCAGAGAAGGCACGAGAAAATTTGGCAGAAGTTGTGAGGGCAAACTTAGAGCTTGAACTAAGCTAAACATT
976 D P E K A R E N W Q K L S E A N L E L E T K L N I
326 CTGAGCAAAATAGCTAAAGACCATTTGGGATGCTTATCTAGGAGTCATTAAGTCATGTAGTGAGCTTACTTCTGAA
1 051 L S K L A K D H W D A Y L G V I K S C S E L T S E
351 AAGTGGGTGCTACATGCTACTGAACCAACCAAGCCATTATTAAGAACTCTTAGGGGCAAGGAAAGCTATG
1 126 K W V L H A T E P N N E A I I K E L L G A R E A M
376 TTGAGAATCAGAACTCTTATGCGTGAGATGGGTGAGGCTGCTAGTGTCTATAGAGCCTGAATCCCAACTCGA
1 201 L R I R T L M R E M G E A A S V P I E P E S Q T R
401 CTTTGGATTCTACGATGAATGCGGAAGGGGTTTACTTGTGCTGTTCTGAGCTGGTGGATTGATGCCATT
1 276 L L D S T M N A E G V L L A G V P G A G G F D A I
426 TTTGCGATTACTTTAGGGGATTCTGGCAGCAAACTGACCCAGGAATGGAGTTCACACAATGTTTGGCCATTATG
1 351 F A I T L G D S G S K L T Q E W S S H N V L A L L
451 GTGAGAGAAGATCCACACGGCGTTTGTCTTAGAAAGTGGTGATCCACGCACCACAAATATTACTTCAGGCGTTTCA
1 426 V R E D P H G V C L E S G D P R T T N I T S G V S
1 501 TCAATTCGTCTCGAGTAA
501 S I R L E *

```

黑色下划线: ATP 结合位点

ATP binding sites are underlined

图 2 LaPMK 基因 cDNA 编码区序列及氨基酸序列

Fig. 2 Coding sequence and amino acid sequence of LaPMK gene cDNA

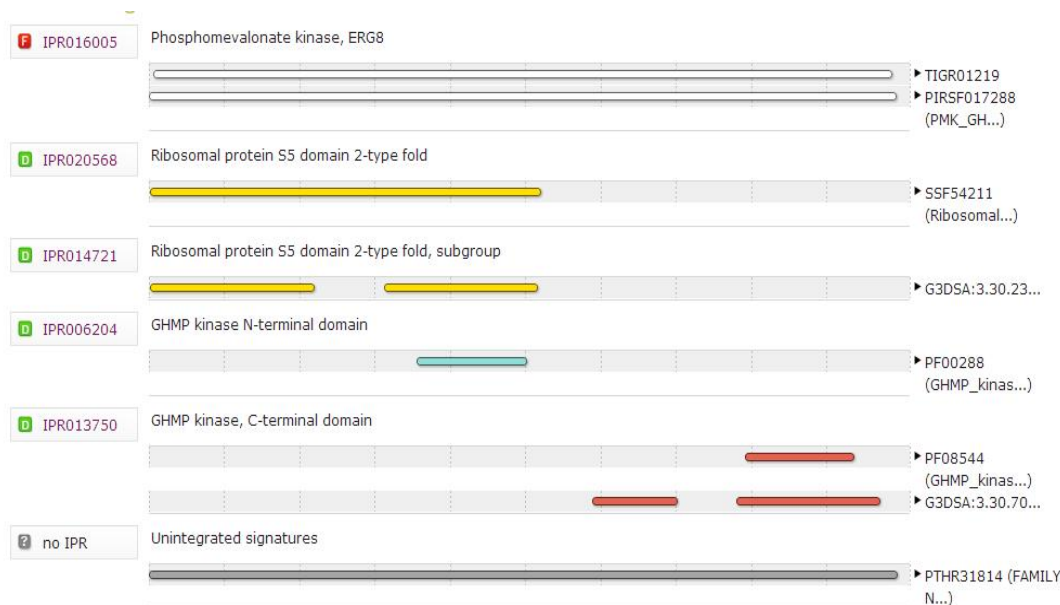


图 3 LaPMK 基因的编码蛋白质保守结构域预测

Fig. 3 Prediction of conserved domain of LaPMK protein encoding by LaPMK gene

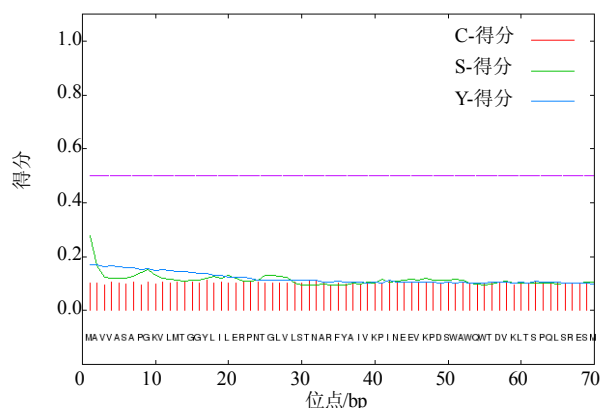


图 4 LaPMK 基因的编码蛋白质信号肽预测分析 (SignalP 预测)

Fig. 4 Prediction of signal peptide of LaPMK protein encoding by LaPMK gene (SignalP prediction)

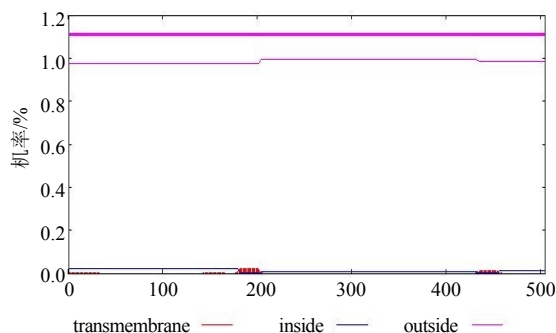


图 5 LaPMK 基因的编码蛋白质的跨膜区预测结果 (TMHMM 预测)

Fig. 5 Prediction of transmembrane domains of LaPMK protein encoding by LaPMK gene (TMHMM prediction)

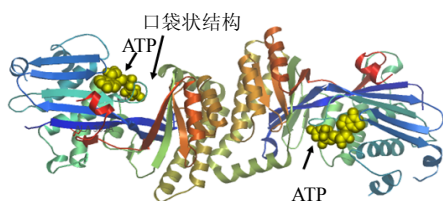


图 6 LaPMK 蛋白三维结构预测

Fig. 6 Prediction of 3D structure of LaPMK protein

将 LaPMK 基因编码蛋白质的氨基酸序列提交到 NCBI 的蛋白质序列数据库进行 BLASTP 搜索, 结果表明: LaPMK 与芜菁的 PMK 蛋白序列相似性最高, 达 92%。从 BLASTP 搜索结果中选取与 LaPMK 基因编码的蛋白质相似性较高的 13 条蛋白质序列, 绘制系统进化树 (图 7)。可以看

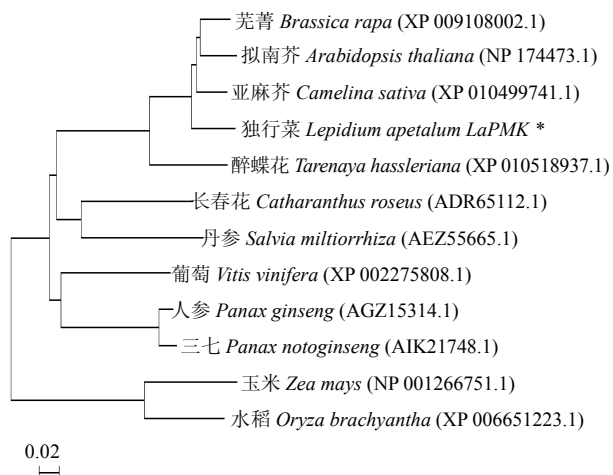
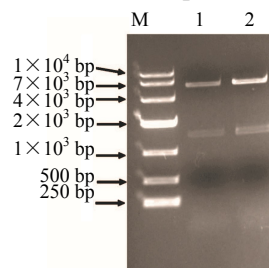


图 7 LaPMK 蛋白与其他物种 PMK 蛋白的系统进化树分析
Fig. 7 Phylogenetic tree of LaPMK protein and PMK proteins from other species

出 LaPMK 蛋白与芜菁 *Brassica rapa* Linn.、拟南芥 *Arabidopsis thaliana* (Linn.) Heynh. 和 亚 麻 芥 *Camelina sativa* (Linn.) Crantz 等十字花科植物的 PMK 蛋白处于同一分支上, 亲缘关系较近, 都属于双子叶植物。

2.3 LaPMK 基因原核表达载体的构建

用限制性内切酶 *EcoRI* 和 *HindIII* 分别对原核表达载体 pET-32a (+) 和测序正确的 pMD19-T-LaPMK 质粒进行双酶切, 连接得到重组质粒 pET32a-LaPMK, 然后转化 *E. coli* BL21(DE3), 挑取单克隆, 进行双酶切鉴定, 结果显示有 1 500 bp 左右的目的条带 (图 8)。将酶切正确的单克隆送测序公司测序, 测序结果显示与目的基因 LaPMK 序列一致, 未发生移码突变, 表明已成功将 LaPMK 基因克隆到原核表达载体 pET-32a 上。

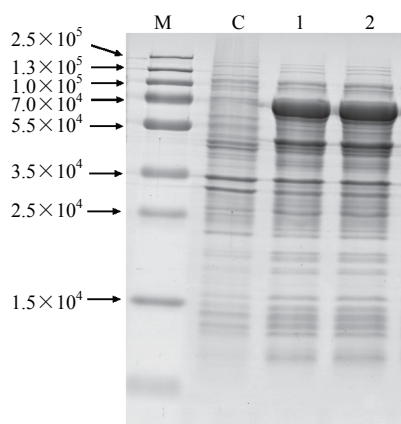


M-Marker 1、2-*EcoRI* 和 *HindIII* 双酶切
M-Marker 1、2-double digestion results by *EcoRI* and *HindIII*

图 8 原核表达载体 pET32a-LaPMK 的双酶切鉴定
Fig. 8 Identification of prokaryotic expression vector pET32a-LaPMK with double digestion

2.4 LaPMK 蛋白的表达

将测序正确的重组质粒 pET32a-LaPMK 转化 *E. coli* BL21 (DE3) 后, 用 IPTG 诱导 LaPMK 蛋白表达。在 28 °C、IPTG 终浓度为 0.4 mmol/L、150 r/min 的条件下, 诱导表达 8 h 后, 提取总蛋白进行 SDS-PAGE 分析 (图 9), 蛋白相对分子质量与预测结果相一致, 结果表明 LaPMK 基因编码的蛋白质原核表达成功。下一步研究需要检测原核表达 LaPMK 蛋白的可溶性, 通过优化表达条件 (温度、诱导时间、IPTG 浓度等), 提高原核表达 LaPMK 蛋白的可溶性, 为纯化 LaPMK 蛋白和检测目的蛋白活性奠定基础, 避免因诱导时间过长 (24 h 以上)、温度过高 (35~37 °C) 等因素降低目的蛋白的可溶性, 形成包涵体^[10]。



M-Marker C-含 pET32a 空载体的 *E. coli* 菌株 1、2-IPTG 诱导的含 pET32a-LaPMK 质粒的 *E. coli* 菌株

M-Marker C-*E. coli* containing pET32a used as control 1, 2-*E. coli* containing pET32a-LaPMK under IPTG induction

图 9 重组 LaPMK 蛋白 SDS-PAGE 电泳

Fig. 9 SDS-PAGE analysis on recombinant LaPMK protein

3 讨论

丹参酮、紫杉醇、人参皂苷等萜类化合物具有重要的药用和经济价值, 而关于独行菜中萜类物质生物合成途径的报道较少。独行菜以干燥成熟的种子入药, 本研究从独行菜种子萌发后, 长成幼苗的叶片中克隆得到了 PMK 基因, 该基因编码蛋白的氨基酸序列与其他植物中 MVA 途径 PMK 蛋白的序列同源性高, 表明 LaPMK 为 MVA 途径的第 2 个限速酶 PMK, 也是控制整个 MVA 代谢途径的关键酶之一^[11], 独行菜 PMK 蛋白作为其萜类物质生物合成途径的关键酶, 具有重要的研究价值。有研究表明在植物体内的 MVA 途径中的 MK 和 PMK 基因表

达量较低, 提高它们的表达量可以提高萜类在植物中的产量, 同时若 PMK 在植物体内过量表达, 可以增加类异戊二烯衍生物的量^[12-13]。目前 PMK 基因在植物研究中的报道较少, 仅在药用植物阳春砂中克隆了 PMK 基因和检测不同组织中 PMK 基因表达差异, 结果显示在叶中 PMK 基因表达量较高, 茎、根中表达量相对较低, 推测叶片可能是阳春砂中萜类的最主要合成部位^[7]。

萜类化合物是植物次生代谢产物中最大的一个家族, 在生物体内有着重要的生理作用, 而且广泛应用于工农业生产和医药卫生中, 但是萜类化合物在生物体内的量很低, 在植物体内的量往往只达到百万分之一的水平, 如何提高植物萜类产量一直是个热门研究方向^[14]。近年来随着萜类化合物生物合成途径及其关键酶研究的深入, 以分子生物学技术为基础的合成生物学和代谢工程, 成为提高植物体内萜类化合物产量最有潜力的途径之一。所以, 通过构建 PMK 基因的植物表达载体并通过农杆菌介导的遗传转化体系转化独行菜, 对其进行遗传改良, 有希望提高转基因独行菜中萜类物质的量, 进一步培育高品质的独行菜新品种, 具有重要的商业价值和现实意义。本实验首次从独行菜中克隆了 PMK 基因, 利用 pET32a 原核表达载体和大肠杆菌 BL21 (DE3) 进行原核表达, 大肠杆菌 BL21 (DE3) 菌株可以增加细胞质中蛋白质二硫键的形成, 使得蛋白的可溶性更好, 同时, 利用 pET32a 载体表达的重组蛋白 N 端带有 6 个 His Tag 标签, 方便后续重组蛋白的纯化^[15]。通过构建原核表达载体 pET32a-LaPMK, 转入大肠杆菌 BL21 (DE3), 成功诱导表达出 LaPMK 蛋白, 并且目的蛋白表达量较高, 为下一步纯化 LaPMK 蛋白进行蛋白质结构和生物学功能研究奠定了基础。该研究将有助于深入认识独行菜萜类化合物生物合成途径及其调控机制, 也为独行菜萜类代谢工程提供候选基因。

参考文献

- [1] 李红伟, 郑晓珂, 弓建红, 等. 独行菜和播娘蒿化学成分及药理作用研究进展 [J]. 药物评价研究, 2013, 36(3): 235-240.
- [2] Hyun J W, Shin J E, Lim K H, et al. Evomonoside: the cytotoxic cardiac glycoside from *Lepidium apetalum* [J]. *Planta Med*, 1995, 61(3): 294-295.
- [3] 刘 强, 丛丽娜, 张宗申. 植物甾醇与三萜类皂苷生物合成基因调控的研究进展 [J]. 安徽农业科学, 2006, 34(19): 4844-4846.

- [4] 常卿, 刘俊峰, 梁栋材. 磷酸甲羟戊酸激酶 PMK 的重组表达和晶体生长 [A] // 第十次中国生物物理学术大会论文摘要集 [C]. 北京: 中国生物物理学大会, 2006.
- [5] Garcia D E, Keasling J D. Kinetics of phosphomevalonate kinase from *Saccharomyces cerevisiae* [J]. *PLoS One*, 2014, 9(1): e87112.
- [6] 王宝莲, 樊庆琦, 李永波, 等. 甲羟戊酸激酶基因研究进展 [J]. 中国农业科技导报, 2011, 13(3): 17-25.
- [7] 杨恩泽, 陆颖锬, 吴秋红, 等. 阳春砂磷酸甲羟戊酸激酶的基因克隆与表达谱分析 [J]. 暨南大学学报: 自然科学与医学版, 2015, 36(2): 1-7.
- [8] Andreassi J L, Leyh T S. Molecular functions of conserved aspects of the GHMP kinase family [J]. *Biochemistry*, 2004, 43(46): 14594-14601.
- [9] Tamura K, Dudley J, Nei M, *et al.* MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0 [J]. *Mol Biol Evol*, 2007, 24(8): 1596-1599.
- [10] Galloway C A, Sowden M P, Smith H C. Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase [J]. *Biotechniques*, 2003, 34(3): 524-526.
- [11] 张长波, 孙红霞, 巩中军, 等. 植物萜类化合物的天然合成途径及其相关合酶 [J]. 植物生理学通讯, 2007, 43(4): 779-786.
- [12] Redding-Johanson A M, Batth T S, Chan R, *et al.* Targeted proteomics for metabolic pathway optimization: Application to terpene production [J]. *Metab Eng*, 2011, 13(2): 194-203.
- [13] Singh P, Batth T S, Juminaga D, *et al.* Application of targeted proteomics to metabolically engineered *Escherichia coli* [J]. *Proteomics*, 2012, 12(8): 1289-1299.
- [14] Ajikumar P K, Tyo K, Carlsen S, *et al.* Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms [J]. *Mol Pharm*, 2008, 5(2): 167-190.
- [15] Rosano G L, Ceccarelli E A. Recombinant protein expression in *Escherichia coli*: advances and challenges [J]. *Front Microbiol*, 2014, 5: 1-17.