

• 药材与资源 •

独行菜 LaF3H 基因克隆、序列分析及原核表达

赵 乐^{1,2}, 马利刚^{1,2}, 张金燕¹, 冯卫生^{1,2}, 匡海学³, 郑晓珂^{1,2*}

1. 河南中医药大学药学院, 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

3. 黑龙江中医药大学药学院, 黑龙江 哈尔滨 150040

摘要: **目的** 为了研究独行菜黄酮类化合物生物合成途径的关键基因, 从独行菜中克隆了黄烷酮-3-羟化酶 (flavanone-3-hydroxylase, F3H) 基因, 命名为 LaF3H, 并进行序列分析、原核表达和纯化。**方法** 根据独行菜转录组数据中 LaF3H 基因序列设计特异性引物, 克隆 LaF3H 基因的 cDNA 序列, 构建 pET-32a-LaF3H 原核表达载体并在大肠杆菌中诱导表达 LaF3H 重组蛋白。**结果** LaF3H 基因开放阅读框 (ORF) 长 1 080 bp, 编码 359 个氨基酸, 其蛋白质相对分子质量为 40 320。序列分析结果表明 LaF3H 具有 F3H 蛋白的 5 个保守基序, 系统进化分析结果显示 LaF3H 蛋白与拟南芥等十字花科植物 F3H 蛋白同源性较高。通过构建原核表达载体 pET-32a-LaF3H, 在大肠杆菌 BL21 (DE3) 菌株中成功表达 LaF3H 重组蛋白, 利用 Ni²⁺亲和色谱得到了纯化的 LaF3H 重组蛋白。**结论** 克隆了 LaF3H 基因, 获得 LaF3H 纯化蛋白, 为下一步制备 LaF3H 蛋白抗体、酶学活性检测, 研究 LaF3H 基因在独行菜黄酮类化合物生物合成途径中的功能奠定基础。

关键词: 独行菜; 黄烷酮-3-羟化酶; 基因克隆; 序列分析; 原核表达

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2018)23-5626-07

DOI: 10.7501/j.issn.0253-2670.2018.23.023

Cloning, sequence analysis, and prokaryotic expression of LaF3H gene from *Lepidium apetalum*ZHAO Le^{1,2}, MA Li-gang^{1,2}, ZHANG Jin-yan¹, FENG Wei-sheng^{1,2}, KUANG Hai-xue³, 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

3. College of Pharmacy, Heilongjiang University of Chinese Medicine, Harbin 150040, China

Abstract: **Objective** In order to study the key genes involved in flavonoid biosynthesis pathway, the flavanone-3-hydroxylase (F3H) gene was isolated from *Lepidium apetalum*, which is named as LaF3H. Meanwhile, the sequence analysis, prokaryotic expression, and purification were also performed. **Methods** Specific primers were designed according to LaF3H gene sequences in the transcriptome data of *L. apetalum*, and the cDNA sequence of LaF3H gene was isolated from *L. apetalum*. By construction the prokaryotic expression vector pET-32a-LaF3H, the recombinant LaF3H protein was expressed in *Escherichia coli* BL21 (DE3) cells under IPTG induction. **Results** The open reading frame (ORF) of LaF3H was 1 080 bp, which encoded a protein of 359 amino acid residues, with a predicted molecular mass of 40 320. Sequence analysis showed that LaF3H contains five conserved motifs of F3H protein. The phylogenetic analysis indicated that LaF3H protein showed the highest homology with F3H protein from cruciferous plants (such as AtF3H from *Arabidopsis thaliana*). The prokaryotic expression vector pET-32a-LaF3H was constructed and the recombinant LaF3H protein was successfully expressed in *E. coli* BL21 (DE3) cells. Furthermore, the recombinant LaF3H protein was purified through Ni²⁺ affinity chromatography. **Conclusion** The LaF3H gene was isolated from *L. apetalum* and the recombinant LaF3H protein was obtained. The results of this study provided the foundation for the further preparation of LaF3H antibody and detection of LaF3H enzyme activity, and were helpful for functional characterization of LaF3H gene involved in flavonoid biosynthesis pathway of *L. apetalum*.

Key words: *Lepidium apetalum* Willd.; flavanone-3-hydroxylase; gene cloning; sequence analysis; prokaryotic expression

收稿日期: 2018-04-06

基金项目: 国家重点基础研究发展计划“973”项目 (2013CB531802); 中央引导地方科技发展专项 (河南道地大宗药材种质评价及集约化种植与示范); 河南省科技攻关计划 (162102310468)

作者简介: 赵 乐 (1983—), 男, 博士, 副教授, 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

黄酮类化合物是一种在植物体内广泛存在的次生代谢产物,种类极为丰富,根据其结构及修饰基团的不同,黄酮类化合物主要分为黄酮、黄酮醇、黄烷酮、黄烷醇、异黄酮和花青素等 6 大类,具有多种生物学活性,而且在植物的生长发育、防御反应(抗逆、抗病等)以及植物花色形成中起着重要作用^[1-2]。黄酮类化合物的生物合成以香豆酰辅酶 A 和丙二酰辅酶 A 为底物,在查耳酮合酶(chalcone synthase, CHS)的催化下生成查耳酮,然后经查耳酮异构酶作用生成柚皮素,柚皮素进入黄酮类化合物的合成途径,经羟基化、糖基化、乙酰化及甲基化等多步反应,生成种类丰富的黄酮、黄酮醇和花青素等黄酮类化合物^[3],其中香豆酰辅酶 A 来源于苯丙烷代谢途径^[4],丙二酰辅酶 A 来源于乙酰辅酶 A。

黄烷酮-3-羟化酶(flavanone-3-hydroxylase, F3H)是黄酮类化合物生物合成途径的关键酶之一,能够催化柚皮素生成二氢山柰酚(dihydrokaempferol, DHK),二氢山柰酚在 F3'H 的催化下生成二氢槲皮素(dihydroquercetin, DHQ),DHK 和 DHQ 在黄酮醇合成酶(flavonol synthase, FLS)和甲基转移酶 1(O-methyltransferase 1, OMT1)的作用下形成槲皮素、山柰酚、异鼠李素等黄酮醇。F3H 蛋白最早从矮牵牛中分离纯化获得,其功能是催化柚皮素 C 环第 3 位羟基化形成 DHK^[5],参与花青素的形成,后来又从矮牵牛的花瓣中克隆得到 F3H 基因,并进行原核表达和酶活性检测^[6]。金鱼草中 F3H 基因发生突变,则会导致花色变淡^[7]。在拟南芥突变体 tt6-5 中,AtF3H 基因发生突变,导致柚皮素含量增加, DHK 含量减少,同时总黄酮和花青素含量也明显降低,使突变体 tt6-5 抗氧化能力降低,对氧化胁迫更加敏感^[8]。所以 F3H 是黄酮类生物合成途径中一个调节点,在植物的次生代谢过程中发挥重要作用。

独行菜 *Lepidium apetalum* Willd. 是十字花科独行菜属植物,其干燥成熟的种子作为北葶苈子药材要用,具有泻肺平喘、利水消肿之功效^[9]。目前已从独行菜中分离得到强心苷、黄酮、香豆素、三萜、甾醇、生物碱等多种化学成分,其中黄酮类成分以槲皮素、异鼠李素、山柰酚为主^[10],黄酮类化合物具有抗氧化、抗肿瘤、消炎等药理作用,是独行菜重要的药效物质基础之一^[11]。目前关于独行菜的研究主要集中在化学成分的分离鉴定、药理作用等方面,但是关于独行菜黄酮类化合物的生物合成

途径及关键基因研究较少,赵乐等^[4]克隆了独行菜苯丙烷代谢途径 LaC4H 基因,进行了生物信息学分析、原核表达和荧光定量 PCR 分析,发现 LaC4H 的表达量与独行菜叶片中黄酮含量之间呈正相关。目前, F3H 基因已从茶树^[12]、鹤望兰^[13]、金花茶^[14]、桑树^[15]等多种植物中克隆,而独行菜 F3H 基因还未见报道。本研究从独行菜中克隆了黄酮类化合物生物合成途径中 F3H 基因的 cDNA 序列,命名为 LaF3H,进行了生物信息学分析、原核表达和纯化,为进一步研究 LaF3H 基因在独行菜黄酮类化合物生物合成中的功能奠定了基础。

1 材料与方法

1.1 材料与试剂

独行菜种子采自河南省伏牛山自然保护区,经河南中医药大学董诚明教授和朱昀昊博士鉴定为十字花科植物独行菜 *Lepidium apetalum* Willd. 干燥成熟的种子,将种子种植于人工气候箱中,培养条件为 16 h/23 °C 光照, 8 h/20 °C 黑暗。待种子萌发后,取独行菜幼苗的叶片,经液氮速冻后保存于 -80 °C 超低温冰箱中备用。

植物总 RNA 提取试剂盒、质粒小提试剂盒、普通 DNA 产物纯化试剂盒、琼脂糖凝胶 DNA 回收试剂盒购自北京天根公司;反转录试剂盒、大肠杆菌感受态细胞购自北京全式金公司;DNA Marker、DNA Taq 酶、限制性核酸内切酶、pMD19-T Vector、T4 DNA Ligase 购自 Takara 公司;Ni-Agarose His 标签蛋白纯化试剂盒购自北京康为世纪公司;原核表达载体 pET-32a 为本实验室保存;其他试剂均为国产分析纯。引物合成、样品测序由北京三博远志公司完成。

1.2 RNA 的提取与反转录

用植物总 RNA 提取试剂盒提取独行菜幼苗叶片总 RNA,经琼脂糖凝胶电泳和 NanoDrop 2000 检测 RNA 的完整性、纯度和浓度,使用反转录试剂盒,以提取得到总 RNA 为模板, Oligo(dT)₂₀ 为引物,反转录合成得到独行菜幼苗叶片的 cDNA。

1.3 LaF3H 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中 F3H 基因的序列信息,用 Oligo7 设计一对独行菜 LaF3H 基因的特异性引物, LaF3H-F: 5'-CTTTGTAAGGTTCTCAGCTAC-3', LaF3H-R: 5'-TAAAAGTAAAGGAGCAACAACG-3',以反转录得到的独行菜幼苗叶片 cDNA 为模板,扩增 LaF3H 基因,PCR 反应程序: 95 °C、2 min; 95 °C、10 s,

54.4 °C、15 s, 72 °C、80 s, 35 个循环; 72 °C 延伸 8 min, PCR 产物经琼脂糖凝胶电泳检测后, 切胶回收与预期大小一致的条带, 将其连接到 pMD19-T 载体上, 然后转化大肠杆菌 DH5 α 感受态细胞, 经菌落 PCR 检测后挑选阳性克隆进行测序。

1.4 LaF3H 基因的生物信息学分析

将测序得到的 LaF3H 基因的序列信息提交到 NCBI, 通过 BLAST 与核酸和蛋白质数据库进行比对搜索, 使用 DNAMAN 软件将其翻译成氨基酸序列, 利用 NCBI ORF Finder 查找开放阅读框 (open reading frame, ORF)。使用 ExPASy Proteomics Server 提供的在线工具 ProtParam tool 预测 LaF3H 基因编码蛋白的理化性质, PredictProtein 预测蛋白质的二级结构, 使用 InterPro Scan 预测保守结构域, 采用 SWISS-MODEL 进行蛋白质的三维同源建模, 使用 SignalP 4.1 Server 进行信号肽的预测, 利用 TargetP 1.1 Server 预测蛋白亚细胞定位, 通过 TMHMM Server v.2.0 预测蛋白跨膜区。使用 DNAMAN 软件对不同植物 F3H 蛋白的氨基酸序列进行多序列比对, 使用 MEGA5 软件相邻连接法 (neighbor-joining) 构建系统进化树, bootstrap 检验的重复次数为 1 000 次。

1.5 pET-32a-LaF3H 原核表达载体的构建

根据 LaF3H 基因的测序结果和生物信息学分析结果, 设计一对 LaF3H 基因编码区的原核表达引物: LaF3H-Exp-F (5'-CGGGATCCATGGCTC-CAGGAAGTCTT-3', 下划线部分为 *Bam*H I 酶切位点), LaF3H-Exp-R (5'-CCGCTCGAGAGCAAAG-ATTGATCGAC-3', 下划线部分为 *Xho* I 酶切位点)。以测序正确的 pMD19-T-LaF3H 质粒为模板, 用 Prime STAR HS Taq 扩增 LaF3H 基因的编码区, PCR 程序: 95 °C、2 min; 95 °C、10 s, 62 °C、15 s, 72 °C、80 s, 30 个循环; 72 °C 延伸 8 min。PCR 产物经琼脂糖凝胶电泳检测后, 回收纯化目的基因条带, 然后用 *Bam*H I 和 *Xho* I 分别对纯化的 PCR 产物和原核表达载体 pET-32a 进行双酶切, 回收目的基因片段和载体片段, 将目的基因片段和载体片段用 T₄ DNA Ligase 在 16 °C 水浴中连接过夜, 然后将连接产物转化大肠杆菌 BL21 (DE3) 感受态细胞, 挑单克隆提取质粒, 用 *Bam*H I 和 *Xho* I 进行双酶切和单酶切鉴定, 将酶切鉴定正确的克隆送北京三博公司测序。

1.6 LaF3H 重组蛋白的诱导表达与纯化

将测序正确的单克隆 (含 pET-32a-LaF3H 质粒)

接种到 LB 液体培养基中, 过夜培养, 第 2 天按照 1:100 比例接种到 LB 液体培养基中, 培养至对数生长期 (A_{600} 至 0.6~0.8), 然后加入 IPTG, 在 28 °C、150 r/min、0.4 mmol/L IPTG 条件下培养 8 h, 诱导 LaF3H 重组蛋白的表达, 诱导完成后用 SDS-PAGE 检测重组蛋白的表达。

根据诱导条件, 大量培养含 pET-32a-LaF3H 质粒的大肠杆菌 BL21 (DE3) 菌株, 诱导完成后在 4 °C、8 000 $\times g$ 条件下离心 5 min 收集菌体, 在冰浴中超声破碎大肠杆菌细胞, 直到溶液变清。大肠杆菌裂解液在 4 °C、12 000 $\times g$ 离心 20 min, 取上清用 0.45 μ m 滤膜滤过, 根据 Ni²⁺ 亲和色谱, 使用 Ni-Agarose His 标签蛋白纯化试剂盒纯化目的蛋白, 用不同浓度 (50、100、200、300、500 mmol/L) 咪唑进行洗脱, 洗脱液经 SDS-PAGE 检测后确定最佳咪唑洗脱浓度, 收集只含有单一条带的目的蛋白样品进行透析和超滤, 最后获得纯化的 LaF3H 重组蛋白。

2 结果与分析

2.1 LaF3H 基因克隆

通过对本实验室前期得到的独行菜转录组数据进行分析, 发现 1 条注释为 F3H 基因的转录本, 长度为 1 199 bp, 使用 NCBI ORF Finder 检测发现该序列包含一个完整的 ORF, 根据该基因的序列信息, 设计一对特异性引物, 扩增 LaF3H 基因, PCR 产物约为 1 200 bp, 电泳结果如图 1 所示, 经测序得到独行菜 LaF3H 基因的 cDNA 序列大小为 1 199 bp, 包含独行菜 LaF3H 基因完整的 ORF 序列大小为 1 080 bp, 位于 94~1 173 bp 区域, 编码 359 个氨基酸, LaF3H 基因的 cDNA 序列信息已经提交到 NCBI GenBank, 登录号为 KY651378, LaF3H 基因 cDNA 的核苷酸序列及对应氨基酸序列如图 2 所示。

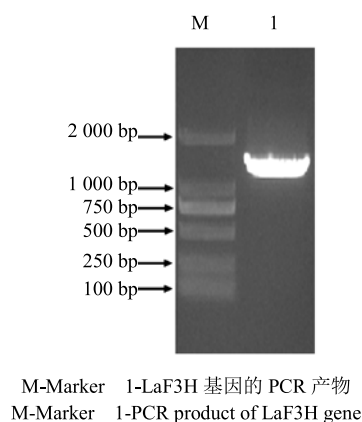


图 1 PCR 扩增 LaF3H 基因
Fig. 1 PCR amplification of LaF3H gene

1	ctttgtgaaggtttctcagctac	CTTTCTCTCTCTTTATTTATTATTCATTCCTTCTCTCTCTTCGTAATTTAAAA	
76	CTCTAATAAAACATATAAATGGTCCGAGGACTTTACTGAAGCTCGCCGGAGAGTCTAAGCTCAACTCTAAATTC		
	M A P G P T L T E L A G E S K L N S K F		19
151	GTCCGGGAGCAGAGTGAACGCCCTAAAGTCGCTTACAACAAGTTCAGTGACGAGATCCCTGTGATTCTCTCTCT		
	V R D E D E R E E A A Y N K F S D E I P V I S L S		44
226	GGAATCGACGATGTTGATGGAAAAGAGGAGATATTTGCTGCAGATTGTGGAGGCTGTGAGAAATGGGGAAATC		
	G I D D V D G K R G D I C R Q I V E A C E N W G I		69
301	TTCCAAGTTGTTGATCATGGCGTTGACACTAGTTTGGTGACTGACATGCTCGTCTAGCTCGTGATTTCTTTGGT		
	F Q V V D H G V D T S L V T D M M T R L A R D F F A		94
376	TTACCTCCGGAAGAAGCTCCGTTTGTATGTGAGGTGGTAAAAAAGGAGGATTCATCGTCTCTAGTCAACCTC		
	L P P E E K L R F D M S G G K K G G F I V S S H L		119
451	CAGGGAAGAGGCTGTACAAGATTGGAGAGAGATAGTAACTGATTTTTCGTATCCGGTGAGAAAACAGAGACTACTCA		
	Q G E A V Q D W R E I V T Y F S Y P V R N R D Y S		144
526	CGGTGGCCGGGATAAGCCAGAAGATGGATGAAAGTGACTGAAACTTACAGTGAGAAGCTGATGAATTTAGCTTGT		
	R W P D K P E G W M K V T E T Y S E K L M N L A C		169
601	AAGCTTCTTGAGGTTTGTCTGAAGCTATGGGACTTGAGAAAGAGGCATTAAACCAATGCATGTGTTGATATGGAC		
	K L L E V L S E A M G L E K E A L T N A C V D M T		194
676	CAAAAGATAGTTGTTAATTATACCCAAATGCTCTGAGCTGATCTCCTCTTGAGACTCAAAAGTCACATGAT		
	Q K G I V V N Y Y P K C P Q P D L T L G L K R H T D		219
751	CCTGGAACCATCACTTTGCTGCTTCAAGACCCAGGTGGTGGTTTACAAGGCCACAAAAGATAATGGCAAGACATGG		
	P G T I T L L Q D Q V G G L Q A T K D N G K T W		244
826	ATTACGGTTGAGCTGTGTAAGAGAGCTTTGTGGTCAATTCGGCGACACCGGTCACTTTCTGAGCAACGGGGAG		
	I T V Q P V E G A F V V N L G D H G H F L S N G R		269
901	TTCAAAAACGCTGATCATCAGGCGGTGGTGAATCTAACTCAAGCAGACTATCTATAGCCACGTTTCAGAATCCG		
	F K N A D H Q A V N S S R L S I A T F Q N P		294
976	GCACCGGAAGCCATGTCATTCGCTTAAAGTGAGAAAAGGAGAAAGCGGATTTTGAAGAGGCCAATTACGTTT		
	A P E A T V Y P L K V R K G E K A I L E E P I T F		319
1051	GCAGAGATGTATAAGAGAAAGTGGCTAAAGATTTGGAGCTGGCTCTTAAGAAGCTGGCCAAAGAGGAACAA		
	A E M Y K R K M A A G D L T R L K K L A K E E Q		344
1126	ACCAACAAGGAATTTGTGCAAGCGCTGTGCATCAAACTTTGGCTTGAAGAAATcggtgtgtgctcctttactttta		
	T N K E F V A K P V D Q I F A *		359

终止密码子用*标记，左侧数字表示核苷酸位置，右侧数字表示氨基酸位置，小写加粗斜体为 LaF3H 基因的克隆引物序列

The marker “*” represents termination codon, the left number indicates nucleotide position, the right number indicates amino acid position, the bold italic lowercase letters represent cloning primer sequences of LaF3H gene

图 2 LaF3H 基因 cDNA 的核苷酸序列及对应的氨基酸序列

Fig. 2 Nucleotide sequence and amino acid sequence of LaF3H gene cDNA

2.2 LaF3H 基因编码蛋白的生物信息学分析

2.2.1 理化性质分析 由 ExPASy Proteomics Server 提供的 ProtParam tool 预测独行菜 LaF3H 蛋白的相对分子质量 40 320, pI 值 5.69, 分子式为 $C_{1793}H_{2834}N_{486}O_{542}S_{14}$, 带负电荷的残基(Asp+Glu): 53, 带正电荷的残基 (Arg+Lys): 47, 不稳定系数是 35.44, 说明该蛋白属于稳定蛋白。

2.2.2 保守结构域预测 用 InterPro Scan 预测 LaF3H 蛋白的保守结构域: 非血红素双加氧酶 *N* 端保守结构域 (non-haem dioxygenase *N*-terminal domain, IPR026992), 2-酮戊二酸铁依赖的双加氧酶 (2-oxoglutarate/iron-dependent dioxygenase, 2-ODD, IPR005123), 属于 2-ODD 家族成员, Fe²⁺ 是其辅助因子, 此外对 LaF3H 蛋白进行 GO 功能预测, 结果显示 LaF3H 参与的生物学过程 (biological process) 是氧化还原过程, 其分子功能 (molecular function) 是氧化还原酶活性。

2.2.3 信号肽、亚细胞定位及跨膜区预测 根据 SignalP 预测 LaF3H 蛋白不含信号肽, TargetP 预测结果表明该蛋白可能位于细胞质中, 根据 TMHMM 预测 LaF3H 蛋白不含跨膜区, 为非膜蛋白。

2.2.4 二级结构和三维结构预测

用 PredictProtein 预测 LaF3H 蛋白二级结构, 结果显示 α -螺旋占

31.48%, β -折叠占 17.27%, 无规卷曲占 51.25%, α -螺旋和无规卷曲是 LaF3H 蛋白二级结构的主要元件。用 SWISS-MODEL 预测 LaF3H 蛋白的三维结构, 见图 3。以拟南芥花青素合成酶 (anthocyanidin synthase, ANS, PDB ID: 1gp4) 的晶体三维结构为模板, 在第 5~332 位氨基酸残基建模, 三维模型覆盖率为 91.4%, 序列的一致性为 28.66%, 根据预测结果 LaF3H 蛋白可能以单体的形式发挥作用。

2.2.5 多序列比对分析 将 LaF3H 蛋白的氨基酸序列提交到 NCBI 用 Blastp 与 Nr 非冗余蛋白数据库进行比对, 比对结果显示 LaF3H 蛋白与拟南芥 *Arabidopsis thaliana* L. (AtF3H, AAC49176)、油菜 *Brassica napus* L. (BnF3H, ABB91635)、

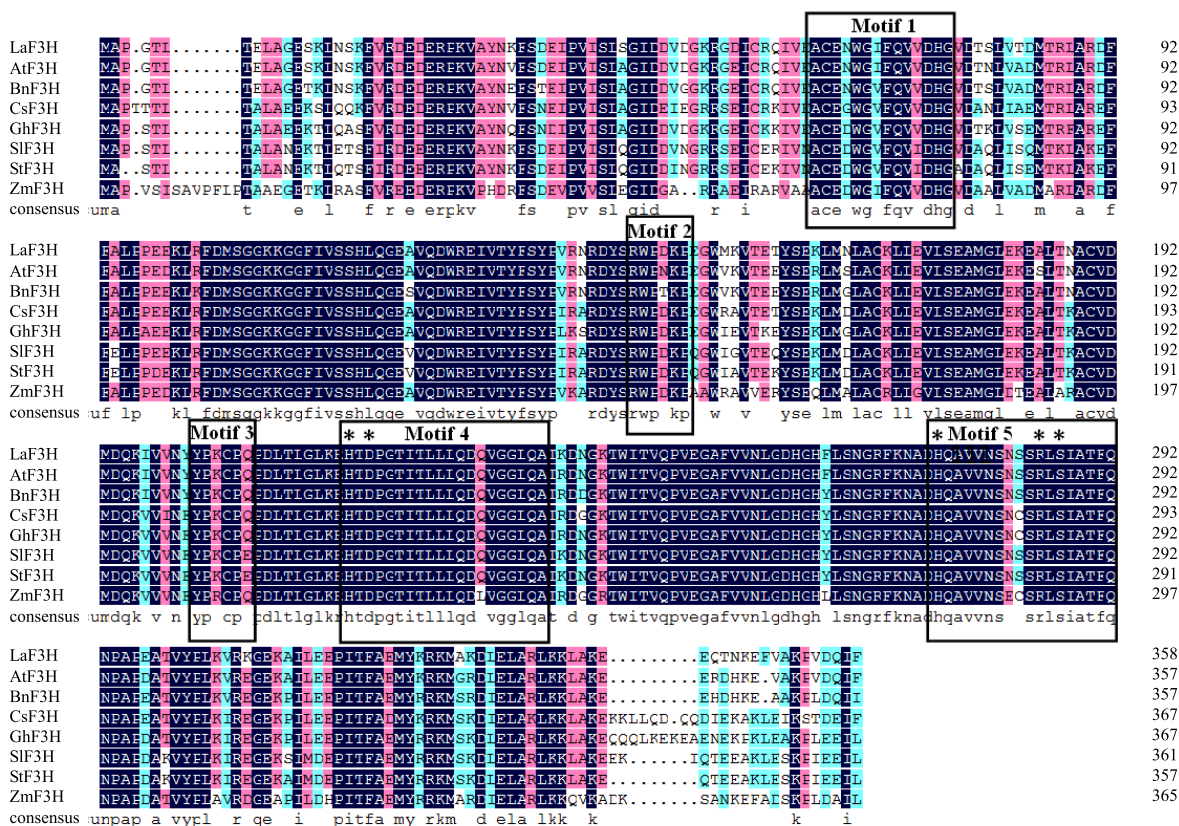


图 3 LaF3H 蛋白质三级结构预测

Fig. 3 Prediction of three-dimensional structure of LaF3H protein

茶 *Camellia sinensis* (L.) O. Ktze. (CsF3H, AAT68774)、陆地棉 *Gossypium hirsutum* L. (GhF3H, ABM64799)、番茄 *Solanum lycopersicum* L. (SIF3H, NP_001316412)、马铃薯 *Solanum tuberosum* L. (StF3H, NP_001274930)、玉米 *Zea mays* L. (ZmF3H, NP_001130275) 的 F3H 蛋白序列一致性较高, 分别为 94%、91%、87%、83%、82%、81%、

76%, 说明 LaF3H 是 F3H 家族成员。用 DNAMAN 软件对这 8 种植物 F3H 蛋白的氨基酸序列进行多序列比对, 结果见图 4。根据 InterPro Scan 预测结果 LaF3H 蛋白有 5 个保守基序, 其中 His217、Asp219 和 His275 参与 Fe^{2+} 的结合, Arg285 和 Ser287 参与 2-酮戊二酸的结合 (RXS 序列), 这些保守基序也存在于其他植物 F3H 蛋白的氨基酸序列中^[16]。



F3H 蛋白的 5 个保守基序用黑框标出, 氨基酸序列一致性为 100%用黑色标记, 一致性 $\geq 75\%$ 用粉色标记, 一致性 $\geq 50\%$ 用蓝色标记

The five conserved motifs of F3H are labeled with black boxes, the identity of amino acid sequences which were 100% was highlighted with black, the identity $\geq 75\%$ was highlighted with pink, and the identity $\geq 50\%$ was highlighted with blue

图 4 LaF3H 蛋白与其他植物中 F3H 蛋白多序列比对分析

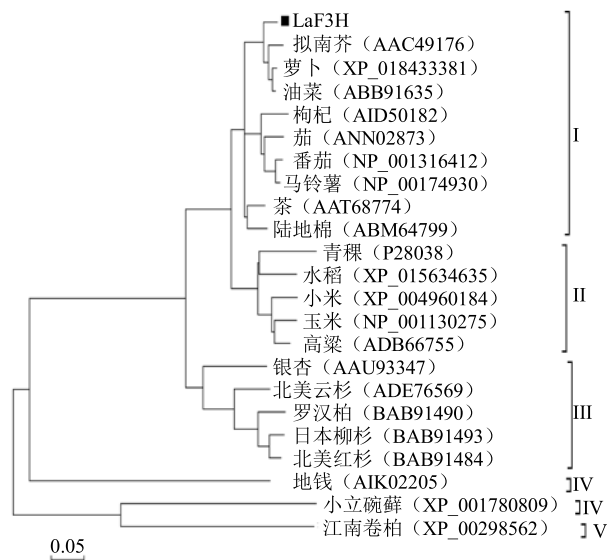
Fig. 4 Multiple sequence alignment of LaF3H and F3H protein from other plant species

2.2.6 系统进化分析 从 NCBI Blastp 的比对结果中, 选取来源于其他 22 种植物的 F3H 蛋白与独行菜 LaF3H 蛋白, 使用 MEGA5 软件相邻连接法 (neighbor-joining) 构建系统进化树 (图 5)。从图 5 中可以看出, 不同物种来源的 F3H 蛋白在进化上归属不同的分支, LaF3H 与来源于十字花科植物的拟南芥、油菜和萝卜的 F3H 亲缘关系较近, 处于同一分支上, 然后与陆地棉、番茄、马铃薯、枸杞等植物的 F3H 聚为一类, 都归属被子植物的双子叶植物 (图 5-I); 水稻、玉米、青稞和高粱等禾本科植物的 F3H 聚为一类, 都归属被

子植物的单子叶植物 (图 5-II); 银杏、北美云杉和罗汉柏等植物的 F3H 聚为一类, 属于裸子植物 (图 5-III); 地钱、小立碗藓和江南卷柏等植物在进化上较为原始, 分别属于苔藓植物和蕨类植物 (图 5-IV 和 V)。

2.3 LaF3H 原核表达载体的构建

用 *Bam*H I 和 *Xho* I 分别对原核表达载体 pET-32a 和带有酶切位点的 PCR 纯化产物进行双酶切, 利用重组 DNA 技术, 构建原核表达载体 pET-32a-LaF3H, 然后转化大肠杆菌 BL21 (DE3) 感受态细胞, 挑单克隆经菌落 PCR 鉴定后, 提取质粒用 *Bam*H I 和



I-双子叶植物 II-单子叶植物 III-裸子植物 IV-苔藓 V-蕨类
登录号在物种名称后, 底部标尺表示每个位点发生 0.05 次氨基酸
替代的进化距离
I-Dicots II-Monocots III-Gymnosperms IV-Mosses V-Ferns
The accession numbers are showed after the name of species,
bootstrap values of the nodes represent the percentage drawn from
the bootstrap test, and the scale bar corresponds to an evolutionary
distance of 0.05 amino acid substitution per site

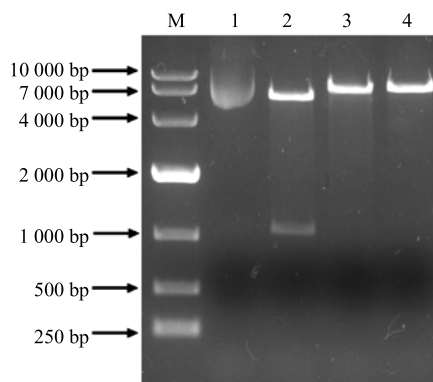
图 5 LaF3H 与其他植物 F3H 蛋白的系统进化树分析

Fig. 5 Phylogenetic tree analysis of LaF3H protein with other plant F3H proteins

Xho I 进行双酶切、*Bam*H I 单酶切、*Xho* I 单酶切鉴定, 结果见图 6, 电泳结果显示 pET-32a-LaF3H 原核表达载体经 *Bam*H I 和 *Xho* I 双酶切后, 有 6 000 bp 左右的载体条带和 1 100 bp 左右的目的基因条带, 而经 *Bam*H I 和 *Xho* I 分别单酶切后, pET-32a-LaF3H 原核表达载体只有 7 000 bp 左右的一条带, pET-32a 空载体呈超螺旋构象, 大小在 6 000 bp 左右。将酶切鉴定正确的单克隆送测序公司测序, 测序结果显示重组质粒 pET-32a-LaF3H 中 LaF3H 的序列与目的基因 LaF3H 的 ORF 序列一致, 未发生碱基突变、插入和缺失突变, 说明原核表达载体 pET-32a-LaF3H 构建成功。

2.4 LaF3H 重组蛋白的诱导表达与纯化

将测序正确的重组质粒 pET-32a-LaF3H 转化大肠杆菌 BL21 (DE3) 后, 当大肠杆菌生长至 A_{600} 为 0.6~0.8 时, 用 IPTG 诱导 LaF3H 蛋白表达。在 28 °C、150 r/min、0.4 mmol/L IPTG 条件下, 诱导表达 8 h 后, 提取大肠杆菌总蛋白进行 SDS-PAGE 分析(图 7)。含 pET-32a-LaF3H 质粒的表达菌株经 IPTG 诱导后, 在约 60 000 处出现目的蛋白条带(图 7), 因

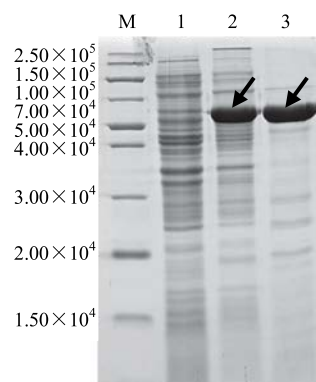


M-Marker 1-pET-32a 空载体 2-*Bam*H I 和 *Xho* I 双酶切
3-*Bam*H I 单酶切 4-*Xho* I 单酶切

M-Marker 1-pET-32a empty vector 2-double digestion result with
*Bam*H I and *Xho* I 3-single digestion result with *Bam*H I 4-single
digestion result with *Xho* I

图 6 原核表达载体 pET-32a-LaF3H 的酶切分析

Fig. 6 Analysis of prokaryotic expression vector pET-32a-LaF3H with restriction endonuclease digestion



M-蛋白标准分子量 1-未诱导的含 pET-32a-LaF3H 质粒的 *E. coli*
菌株 2-诱导的含 pET-32a-LaF3H 质粒的 *E. coli* 菌株 3-纯化的
LaF3H 重组蛋白, 箭头显示为 LaF3H 重组蛋白

M-protein marker 1-uninduced *E. coli* containing pET-32a-LaF3H
2-Induced *E. coli* containing pET-32a-LaF3H 3-The purified
recombinant LaF3H protein The arrows show the recombinant
LaF3H proteins

图 7 LaF3H 重组蛋白的 SDS-PAGE 分析

Fig. 7 SDS-PAGE analysis of recombinant LaF3H protein

为 LaF3H 蛋白在进行原核表达时, 其 *N*-末端融合了 pET-32a 载体的标签序列 (Trx-Tag、His-Tag 和 S-Tag), 总相对分子质量为 60 320。根据 LaF3H 重组蛋白的诱导条件, 扩大培养体积, 超声裂解菌体, 大肠杆菌裂解液经离心取上清, 利用 Ni^{2+} 亲和色谱, 采用 Ni-Agarose His 标签蛋白纯化试剂盒纯化 LaF3H 重组蛋白, 最终得到纯化的 LaF3H 重组蛋白(图 7)。

3 讨论

黄酮类化合物是植物次生代谢的重要产物之一,具有抗氧化、抗肿瘤、抗心血管疾病等功能^[17]。F3H 是植物体内黄酮类化合物生物合成途径的关键酶之一,拟南芥 F3H 基因发生突变,导致拟南芥中总黄酮和花青素含量降低^[8]。通过构建烟草 F3H 基因反义表达载体,抑制烟草 F3H 基因的表达,使代谢流的方向转向异黄酮和黄酮合成方向,增加了染料木黄酮(属于异黄酮类化合物)的含量^[18]。这些研究结果说明 F3H 是黄酮类化合物生物合成途径中一个调节点,具有重要的研究价值。黄酮类化合物作为独行菜重要的药效物质基础之一,具有显著的药理活性,但是在独行菜中黄酮类化合物的含量较低,所以可通过基因工程技术调控黄酮生物合成途径关键酶的活性,比如可构建 LaF3H 基因的植物表达载体,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,有希望提高独行菜中黄酮类化合物的含量,进一步培育高黄酮含量的独行菜新品种,具有重要的商业价值和现实意义。

目前关于独行菜的研究多集中于化学成分的分离和鉴定,以及药理作用等方面,对于独行菜中黄酮类化合物的生物合成途径研究较少,本研究首次从独行菜幼苗的叶片中克隆得到了黄酮类化合物生物合成途径的 LaF3H 基因,并对 LaF3H 基因及其编码蛋白进行了生物信息学分析,LaF3H 基因 ORF 为 1 080 bp,编码 359 个氨基酸。氨基酸序列比对结果表明,LaF3H 蛋白与拟南芥和油菜的 F3H 蛋白序列一致性较高,分别为 94%和 91%,属于 2-酮戊二酸双加氧酶(2-ODD)家族成员,具有 Fe²⁺和 2-酮戊二酸结合位点。系统进化树分析表明,独行菜 LaF3H 蛋白与十字花科植物 F3H 蛋白亲缘关系较近,然后与双子叶植物的 F3H 蛋白聚为一支。通过构建原核表达载体 pET-32a-LaF3H,在大肠杆菌中成功表达 LaF3H 重组蛋白,而且表达的 LaF3H 重组蛋白在上清中,通过 Ni²⁺亲和色谱获得了纯化的 LaF3H 重组蛋白,为下一步进行体外酶促反应,研究 LaF3H 酶学活性以及底物特异性奠定基础,将有助于深入认识独行菜黄酮类化合物生物合成途径及其调控机制,为研究 LaF3H 基因在独行菜黄酮类化合物生物合成中的功能奠定基础,也为通过基因工程技术遗传改良独行菜提供候选基因。

参考文献

[1] Saito K, Yonekura-Sakakibara K, Nakabayashi R, et al.

- The flavonoid biosynthetic pathway in Arabidopsis: Structural and genetic diversity [J]. *Plant Physiol Biochem*, 2013, 72: 21-34.
- [2] Winkel-Shirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology [J]. *Plant Physiol*, 2001, 126(2): 485-493.
- [3] Tanaka Y, Sasaki N, Ohmiya A. Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids [J]. *Plant J*, 2008, 54(4): 733-749.
- [4] 赵乐, 马利刚, 杨泽岸, 等. 独行菜 C4H 基因克隆与表达分析 [J]. *药学学报*, 2017, 52(5): 821-831.
- [5] Britsch L, Grisebach H. Purification and characterization of (2S)-flavanone 3-hydroxylase from *Petunia hybrida* [J]. *Eur J Biochem*, 1986, 156(3): 569-577.
- [6] Britsch L, Ruhnau-Brich B, Forkmann G. Molecular cloning, sequence analysis, and *in vitro* expression of flavanone 3 beta-hydroxylase from *Petunia hybrida* [J]. *J Biol Chem*, 1992, 267(8): 5380-5387.
- [7] Martin C, Prescott A G, Mackay S, et al. Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus* [J]. *Plant J*, 1991, 1(1): 37-49.
- [8] 孙鹏. 拟南芥黄烷酮-3-羟化酶调控非生物胁迫抗性机理研究 [D]. 济南: 山东农业大学, 2014.
- [9] 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.
- [10] 赵海誉, 范妙璇, 石晋丽, 等. 北葶苈子化学成分研究 [J]. *中草药*, 2010, 41(1): 14-18.
- [11] 乔小燕, 马春雷, 陈亮, 等. 植物类黄酮生物合成途径及重要基因的调控 [J]. *天然产物研究与开发*, 2009, 21(2): 354-360.
- [12] 胡晓婧, 许玉娇, 高丽萍, 等. 茶树黄烷酮 3-羟化酶基因 (F3H) 的克隆及功能分析 [J]. *农业生物技术学报*, 2014, 22(3): 309-316.
- [13] 樊荣辉, 黄敏玲, 钟淮钦, 等. 鹤望兰黄烷酮 3-羟化酶基因 *SrF3H* 的克隆及表达分析 [J]. *分子植物育种*, 2014, 12(1): 168-172.
- [14] 周兴文, 殷恒福, 朱宇林, 等. 金花茶黄烷酮 3-羟化酶基因 *CnF3H* 的克隆及表达分析 [J]. *分子植物育种*, 2015, 13(9): 2051-2057.
- [15] 尚亚君, 钱娇, 倪红梅, 等. 桑树黄烷酮-3-羟化酶基因的克隆及在 35 份桑种质资源植株叶片中的表达差异 [J]. *蚕业科学*, 2016, 42(5): 775-782.
- [16] Shen G, Pang Y, Wu W, et al. Cloning and characterization of a flavanone 3-hydroxylase gene from *Ginkgo biloba* [J]. *Bioscience Rep*, 2006, 26(1): 19-29.
- [17] 张华玲, 黄元射, 杨春贤, 等. 苦荞黄烷酮 3-羟化酶基因 F3H 的克隆及序列分析 [J]. *西北植物学报*, 2010, 30(3): 447-452.
- [18] Liu R, Hu Y, Li J, et al. Production of soybean isoflavone genistein in non-legume plants via genetically modified secondary metabolism pathway [J]. *Metab Eng*, 2007, 9(1): 1-7.