

独行菜 *C4H* 基因克隆与表达分析

赵 乐^{1,2}, 马利刚^{1,2}, 杨泽岸¹, 冯卫生^{1,2}, 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)

摘要: 以独行菜 (*Lepidium apetalum*) 为材料, 通过分析独行菜转录组数据, 设计特异性引物, 克隆了独行菜肉桂酸-4-羟化酶 (cinnamate-4-hydroxylase, C4H) 基因的 cDNA 序列, 命名为 *LaC4H*, GenBank 登录号为 KX064050, 并进行生物信息学分析, 原核表达、纯化, 组织特异性分析和诱导表达分析。结果表明: ① *LaC4H* 基因开放阅读框 (ORF) 长 1 518 bp, 编码 505 个氨基酸, 其蛋白质分子质量为 57.73 kD。② 生物信息学分析显示 *LaC4H* 蛋白包含细胞色素 P450 的保守基序和 5 个特征性底物结合位点, 是细胞色素 P450 超家族成员, 系统进化分析显示 *LaC4H* 蛋白与拟南芥等十字花科植物 C4H 蛋白同源性较高。③ 通过构建原核表达载体 pET-32a-*LaC4H* 在大肠杆菌 BL21 (DE3) 菌株中成功表达 *LaC4H* 重组蛋白, 利用 Ni²⁺亲和色谱得到了纯化的 *LaC4H* 重组蛋白。④ 荧光定量 PCR 结果表明 *LaC4H* 基因在茎中表达量最高, 叶和花中次之, 根中表达量最低。经茉莉酸甲酯 (methyl jasmonate, MeJA) 诱导后, 叶中 *LaC4H* 表达量明显上升, 诱导后 48 h 达到最高值, 叶中 *LaC4H* 的表达量与黄酮含量之间呈正相关。这为进一步研究 *LaC4H* 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; 肉桂酸-4-羟化酶; 基因克隆; 生物信息学分析; 表达分析

中图分类号: R931

文献标识码: A

文章编号: 0513-4870 (2017) 05-0821-11

Cloning and expression analysis of *C4H* gene from *Lepidium apetalum*

ZHAO Le^{1,2}, MA Li-gang^{1,2}, YANG Ze-an¹, 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 and Chinese Medicine Development of Henan Province, Zhengzhou 450046, China)

Abstract: *Lepidium apetalum* was used as an experimental material in this study. By analyzing the transcriptome data of *L. apetalum* and application of the specific primers, cDNA of cinnamate-4-hydroxylase (C4H) gene was isolated from *L. apetalum* and named as *LaC4H* (GenBank accession No. KX064050). Meanwhile, the bioinformatic analysis, prokaryotic expression, purification, tissue-specific expression analysis and expression analysis after methyl jasmonate (MeJA) treatment were carried out. The results indicated that: ① The open reading frame (ORF) of *LaC4H* was 1 518 bp, which encoded a protein of 505 amino acid residues, with a predicted molecular mass of 57.73 kD. ② Bioinformatic analysis showed that *LaC4H* protein contained the conserved sequences of cytochrome P450 (CYP450) and 5 substrate recognition sites (SRSs) of CYP73A1, therefore *LaC4H* protein was a member of CYP450 superfamily. The phylogenetic analysis indicated that *LaC4H* protein showed the highest homology with C4H protein from cruciferous plants (such as AtC4H from

收稿日期: 2017-02-07; 修回日期: 2017-02-20.

基金项目: 国家重点基础研究发展计划“973 计划”项目 (2013CB531802); 河南省科技攻关计划 (162102310468); 河南中医学院博士科研基金 (BSJJ2011-07).

*通讯作者 Tel / Fax: 86-371-65962746, E-mail: zhengxk.2006@163.com

DOI: 10.16438/j.0513-4870.2017-0115

Arabidopsis thaliana). ③ Through the construction of the prokaryotic expression vector pET-32a-LaC4H, the recombinant LaC4H protein was successfully expressed in *E. coli* BL21 (DE3) cells and the recombinant LaC4H protein was purified by Ni²⁺ affinity chromatography. ④ Real-time PCR analysis indicated that *LaC4H* was expressed in a high transcript level in stems, lower levels in leaves and flowers, the lowest level in roots. After MeJA treatment, the expression level of *LaC4H* in leaves was increased significantly to reach the highest level at 48 h. Furthermore, the expression levels of *LaC4H* were positively correlated with the flavonoids contents in leaves. The results of this study provide the fundamental information on *LaC4H* gene in the flavonoids biosynthesis pathway of *L. apetalum*.

Key words: *Lepidium apetalum*; cinnamate-4-hydroxylase; gene cloning; bioinformatic analysis; expression analysis

苯丙烷代谢途径可以合成包括黄酮、木质素和香豆素在内的多种植物次生代谢产物,是植物一条重要的次生代谢途径,在植物的生长发育、机械支持、防御反应(抗逆、抗病等)以及花色形成中起着重要作用^[1,2]。苯丙烷代谢途径起源于苯丙氨酸,苯丙氨酸在苯丙氨酸解氨酶(phenylalanine ammonia-lyase, PAL)的作用下生成反式肉桂酸,反式肉桂酸在肉桂酸-4-羟化酶(cinnamate-4-hydroxylase, C4H)催化下生成香豆酸,香豆酸在4-香豆酸辅酶A连接酶(4-coumaroyl CoA ligase, 4CL)的作用下生产香豆酰辅酶A,然后进入黄酮、木质素、芥子酸酯等化合物的下游合成途径(图1)^[3]。

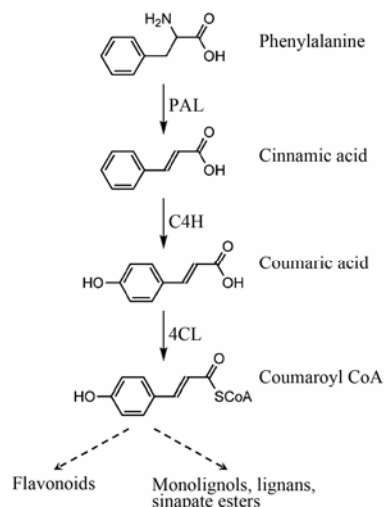


Figure 1 The general phenylpropanoid pathway^[3]

C4H 催化苯丙烷代谢途径的第二步反应,植物中 *C4H* 基因最早从豌豆苗的顶芽中发现,具有催化反式肉桂酸苯环第4位碳原子上形成羟基的功能^[4],属于细胞色素 P450 氧化酶(cytochrome P450, CYP450)家族,是 CYP73 亚家族成员之一^[5]。随后 *C4H* 基因从绿豆^[6]、菊苣^[5]、拟南芥^[7]、烟草^[8]和番茄^[9]等植物

中被克隆,在拟南芥中 *AtC4H* 蛋白由单拷贝基因编码,在拟南芥生长发育以及应对环境信号中起作用^[10],拟南芥 *AtC4H* 基因与 *AtPAL1*、*At4CL* 基因在 mRNA 水平具有类似表达模式,不仅组织表达特异性类似,都是在茎中表达量最高,根和长角果次之,叶和花中表达量最低,而且在应对光照、机械损伤和诱导子处理时表达模式也类似,此外 *AtC4H* 基因的启动子中也包含了 *AtPAL* 和 *At4CL* 基因的3个顺式作用元件^[7]。在番茄中同源正义超表达 *C4H* 基因,降低了番茄茎中木质素的含量却增加了番茄果实中黄酮类化合物的含量^[9]。在逆境胁迫下,苦苣、矮牵牛和茶树等植物中 *C4H* 基因的表达量和黄酮含量之间具有显著的相关性^[11-13]。所以, *C4H* 是苯丙烷代谢途径中一个调节点,在植物的次生代谢过程中起重要作用。

独行菜(*Lepidium apetalum* Willd.)是十字花科独行菜属植物,其干燥成熟的种子称为“北葶苈子”,是中医临床常用的泻肺平喘、利水消肿的传统中药,对心血管、高血脂等疾病有显著的药理作用^[14]。通过对独行菜药效物质基础及作用机制进行研究,目前已从独行菜中分离得到硫苷类、强心苷类、生物碱类、黄酮类、香豆素类等多种成分^[15],其中黄酮类成分以槲皮素、异鼠李素、山柰酚为主^[14],黄酮类成分具有抗心脑血管疾病、抗肿瘤、抗炎、抗氧化等药理作用,是独行菜重要的药效物质基础之一^[16]。目前独行菜的研究多集中在化学成分分离、药理作用以及低温胁迫等方面,且已有良好的研究基础,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,Zhou 等^[17]利用转录组高通量测序分析了可能参与独行菜次生代谢产物合成和代谢的基因,其中参与苯丙氨酸代谢途径的 unigene 有 92 个。目前已见报道独行菜基因克隆有种子黏液质基因 *MUM4*、*TTG1* 基因^[18,19],以及植物低温胁迫相关的 *CBF* (C-repeat binding factor) 基因^[20]等,独行菜次生代谢产

物生物合成途径中关键基因的克隆报道较少, Ma 等^[21]克隆了独行菜二萜类化合物生物合成途径的关键酶牻牛儿基牻牛儿基焦磷酸合成酶 (geranylgeranyl pyrophosphate synthase, GGPS) 基因, 并对 *LaGGPS* 基因进行了生物信息学分析和原核表达。目前, *C4H* 基因已从拟南芥等多种植物中克隆, 而独行菜中 *C4H* 基因还未见报道, *LaC4H* 基因的表达量与独行菜黄酮含量的相关性也有待研究。本研究从独行菜中克隆了参与黄酮类化合物生物合成途径 *C4H* 基因的 cDNA 序列, 命名为 *LaC4H*, 进行了生物信息学分析, 原核表达、纯化和组织特性分析, 还分析了独行菜在茉莉酸甲酯 (methyl jasmonate, MeJA) 诱导下 *LaC4H* 的表达水平与总黄酮含量变化趋势的相关性, 这为进一步研究 *LaC4H* 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。

材料与方法

材料 独行菜种子采自河南省伏牛山自然保护区, 由河南中医药大学董诚明教授鉴定为十字花科独行菜属植物独行菜干燥成熟的种子, 将种子在人工气候箱中培养, 培养条件为 16 h/23 °C 光照, 8 h/20 °C 黑暗。种子萌发培养 45 天待独行菜开花后, 采集独行菜的根、茎、叶和花, 液氮速冻, 保存于 -80 °C 冰箱, 用于组织特异性分析。种子萌发培养 30 天后, 喷洒 MeJA 诱导, MeJA 工作浓度为 50 $\mu\text{mol}\cdot\text{L}^{-1}$, 分别于诱导后 4、12、24、48 和 72 h 取独行菜叶片, 样品经液氮速冻后保存于 -80 °C, 用于诱导表达分析。

试剂 植物总 RNA 提取试剂盒、质粒小提试剂盒、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生化公司。反转录试剂盒、大肠杆菌感受态细胞购自北京全式金生物公司。DNA *Taq* 酶、限制性核酸内切酶、pMD19-T Vector、T4 DNA ligase 购自 TaKaRa 公司。引物合成、样品测序由北京三博远志生物公司完成。原核表达载体 pET-32a 由本实验室保存, 其他试剂均为国产分析纯。

RNA 提取及 cDNA 的合成 使用植物总 RNA 提取试剂盒提取独行菜根、茎、叶、花的总 RNA, 以及 MeJA 诱导后不同时间点独行菜叶片的总 RNA, 用 1% 琼脂糖凝胶电泳检测 RNA 的完整性。使用反转录试剂盒, 以得到的独行菜各组织的总 RNA 为模板, oligo (dT)₂₀ 为引物, 反转录合成得到独行菜各组织的 cDNA 和 MeJA 诱导后不同时间点独行菜叶片的 cDNA。

***LaC4H* 基因克隆** 通过分析实验室前期获得的

独行菜转录组数据中 *C4H* 基因序列信息, 用 Primer 5 设计一对 *LaC4H* 基因的特异性引物: *LaC4H-F* (5'-TTCTGCTTGTATTTCATTCGG-3') 和 *LaC4H-R* (5'-AGTCATTTACTTTCTTGGCTT-3'), 以反转录得到的叶片 cDNA 为模板, 进行 PCR 扩增, PCR 程序为: 95 °C 2 min; 95 °C 10 s, 54.9 °C 15 s, 72 °C 1 min 42 s, 35 个循环; 72 °C 延伸 8 min。PCR 产物经 1% 琼脂糖凝胶电泳检测后, 切胶回收与预期大小一致的条带, 通过 T-A 克隆将其连接到 pMD19-T 载体上, 转化大肠杆菌 DH5 α 感受态细胞, 均匀涂布在 LB 平板 (含氨苄青霉素) 上, 37 °C 倒置培养 12 h, 经菌落 PCR 检测后挑选阳性克隆进行测序。

***LaC4H* 基因的生物信息学分析** 将测序得到的 *LaC4H* 基因的序列信息提交到 NCBI 网站, 通过 blastp 与蛋白质非冗余数据库进行序列比对, 然后通过 NCBI ORF Finder 查找开放阅读框 (open reading frame, ORF)。使用 ExPASy 提供的 ProtParam tool 预测 *LaC4H* 蛋白的理化性质, 使用 InterPro Scan 预测保守结构域, PredictProtein 预测蛋白质的二级结构, SWISS-MODEL 进行蛋白质的三维同源建模, 使用 SignalP 4.0 Server 进行蛋白质信号肽预测, TargetP 1.1 Server 进行蛋白亚细胞定位预测, 通过 TMHMM Server v.2.0 预测蛋白跨膜区。使用 DNAMAN 软件对不同植物 *C4H* 蛋白的氨基酸序列进行多序列比对分析, 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 1000 次。

pET-32a-*LaC4H* 原核表达载体的构建 根据 *LaC4H* 基因的测序结果和生物信息学分析结果, 设计一对 *LaC4H* 基因的原核表达引物: *LaC4H-Exp-F* (5'-CGGAATTCATGGACCTTCTCTTGTT-3', 下划线部分为 *EcoR* I 切位点), *LaC4H-Exp-R* (5'-CCGCTCGAGTTAAACGGCTCTTGTT-3', 下划线部分为 *Xho* I 酶切位点), 以测序正确的 pMD19-T-*LaC4H* 质粒为模板, 用 PrimeSTAR HS *Taq* 扩增 *LaC4H* 基因的 ORF 序列, PCR 程序为: 95 °C 2 min; 95 °C 10 s, 60 °C 15 s, 72 °C 1 min 36 s, 30 个循环; 72 °C 延伸 8 min。PCR 产物经 1% 琼脂糖凝胶电泳检测后, 用 DNA 产物纯化试剂盒回收纯化目的基因, 然后用 *EcoR* I 和 *Xho* I 分别对原核表达载体 pET-32a 和 PCR 纯化产物进行双酶切, 回收载体片段和目的基因片段, 用 T4 DNA ligase 将双酶切后的载体片段和目的基因片段在 16 °C 过夜连接, 然后将连接产物转化大肠杆菌 BL21 (DE3) 感受态细胞, 均匀涂布在 LB 平板 (含氨苄青霉素) 上, 37 °C 倒置培养 12 h, 挑单克隆摇菌提取质粒, 用

EcoR I 和 *Xho* I 对得到的质粒进行双酶切鉴定, 酶切鉴定正确的克隆送公司测序。

LaC4H 重组蛋白的诱导表达与纯化 将测序正确的单克隆 (含 pET-32a-LaC4H 质粒) 接种于 LB 液体培养基 (含氨苄青霉素) 中, 37 °C 培养过夜后, 第 2 天按照 1:100 比例接种到含氨苄青霉素的 LB 液体培养基中, 在 37 °C、220 r·min⁻¹ 条件下培养至对数生长期 (OD₆₀₀ 至 0.6~0.8), 加入 IPTG 使其终浓度为 0.4 mmol·L⁻¹, 然后继续在 28 °C、150 r·min⁻¹ 条件下培养 8 h, 诱导独行菜 LaC4H 重组蛋白的表达, 诱导完成后, 用 SDS-PAGE (5% 浓缩胶和 12% 分离胶) 检测重组蛋白的表达。

根据 LaC4H 重组蛋白的诱导条件, 大量培养大肠杆菌 BL21 (DE3) 菌株 (含 pET-32a-LaC4H 质粒), 诱导完成后在 4 °C、8 000×g 条件下离心 5 min 收集菌体, 经 PBS 缓冲液洗涤后, 在冰浴中超声破碎大肠杆菌细胞, 大肠杆菌裂解液在 4 °C、10 000×g 条件下离心 10 min, 分别收集上清液和沉淀, 将沉淀重悬于含 8 mol·L⁻¹ 尿素的结合缓冲液 (20 mmol·L⁻¹ Tris-HCl、5 mmol·L⁻¹ 咪唑、0.5 mol·L⁻¹ NaCl、8 mol·L⁻¹ 尿素) 中, 使包涵体充分溶解, 10 000×g 条件下离心 20 min, 收集上清液, 经 0.45 μm 滤膜过滤后, 利用 Ni²⁺ 亲和色谱, 使用 Ni-Agarose His 标签蛋白纯化试剂盒纯化 LaC4H 重组蛋白, 用不同咪唑浓度 (50、100、200、300 和 500 mmol·L⁻¹) 洗脱目的蛋白, 经 SDS-PAGE 检测后确定最佳咪唑洗脱浓度, 收集单一条带的目的蛋白样品, 经过透析和超滤, 得到纯化的 LaC4H 重组蛋白。

LaC4H 组织特异性表达、诱导表达分析和独行菜叶片总黄酮含量测定 独行菜各组织的 cDNA 和 MeJA 诱导后不同时间点独行菜叶片的 cDNA, 稀释 10 倍作为荧光定量 PCR (Real-time PCR) 反应模板, 以 *EF-1α* 基因^[22, 23] 为内参检测 *LaC4H* 基因在独行菜各组织的表达量 (正向引物: 5'-CAAGAGGCCATCA GACAAGC-3'; 反向引物: 5'-TACCTGTCTCAACAC GTCCC-3'), *LaC4H* 基因 RT-PCR 正向引物 (5'-GAGCTGTCCCGGGATCATA-3')、反向引物 (5'-AGCTCGAAGTTCTGGACCAA-3'), 每个样品重复 3 次, RT-PCR 反应体系为: 10 μL 2×QuantiNova SYBR Green PCR Master Mix (QIAGEN)、正反向引物各 0.4 μL、cDNA 模板 2 μL、加 ddH₂O 至 20 μL, 反应程序为: 95 °C 预变性 3 min, 95 °C 变性 10 s, 56 °C 退火 20 s, 72 °C 延伸 30 s, 35 个循环, 根据熔解曲线判断 RT-PCR 产物的特异性, 相对定量分析采用 2^{-ΔΔCt} 方

法^[24], 实验数据用统计分析软件 SPSS 16.0 进行单因素方差分析, 分析 *LaC4H* 基因组织特性表达时以独行菜叶片 Ct 值的平均数设为 1, 分析 *LaC4H* 基因诱导表达时以 0 h 叶片 Ct 值平均数设为 1。

参照 Zheng 等^[25] 的方法, 采用紫外分光光度法测定 MeJA 诱导后 48 h 和 72 h 独行菜叶片总黄酮含量, 以芦丁为标准品绘制标准曲线并计算回归方程, 测定样品中总黄酮含量 (mg·g⁻¹ DW), 每个样品重复 3 次, 实验数据用 SPSS 16.0 进行单因素方差分析。

结果与分析

1 *LaC4H* 基因的克隆

通过分析本实验室前期获得的独行菜转录组数据, 发现一个注释为肉桂酸-4-羟化酶 (C4H) 的转录本, 长度为 1 639 bp, 利用 NCBI ORF Finder 分析显示其包含一个完整的 ORF, 根据该基因的序列信息, 设计一对特异性引物, 以独行菜叶片 cDNA 为模板, 进行 PCR 扩增, PCR 产物约为 1 600 bp, 与预期大小相符, 电泳结果如图 2 所示。将 PCR 产物连接到 pMD19-T 载体上, 测序后获得独行菜 *LaC4H* 基因的 cDNA 序列, 大小为 1 639 bp, 包含 *LaC4H* 基因完整的 ORF 序列, 大小为 1 518 bp, 位于 90~1 607 bp 区域, 编码 505 个氨基酸, *LaC4H* 基因 cDNA 的序列信息已提交到 NCBI Genbank, 登录号为 KX064050, *LaC4H* 基因 cDNA 的核苷酸序列及对应氨基酸序列如图 3 所示。

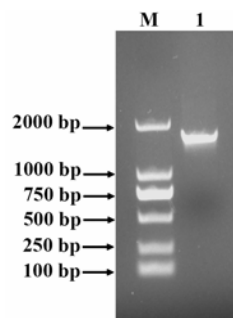


Figure 2 PCR amplification of *LaC4H* gene. M: DL2000 DNA marker; 1: PCR product of *LaC4H* gene

2 *LaC4H* 基因编码蛋白的生物信息学分析

2.1 理化性质分析 使用 ExPASy Proteomics Server 的在线软件 ProtParam tool 预测 *LaC4H* 蛋白的分子质量为 57.73 kD, 等电点是 9.04, 分子式为 C₂₆₂₂H₄₁₅₀N₇₁₄O₇₂₁S₁₆, 带负电荷的残基 (Asp+Glu): 61, 带正电荷的残基 (Arg+Lys): 68, 不稳定系数为 43.13, 说明该蛋白属于不稳定蛋白。

```

1      ctgottgttattatcattoggggACGAGAATTTCTCACCGACTATAATCAATCTCACTGTGAATTACTAAATTCGAGAG
79     CGAACAGTTATGGACCTTCTCTTGTGGAGAAGTCTCTAATTGCCGCTTCGCGGGCGGTGGTTCTCGCTACGGTGATT
      M D L L L L E K S L I A V F A A V V L A T V I 23
157    TCCAAGCTCCGCGGCAAAAACCTGAAGCTACCTCTGGTCTATGCGGATTCCTATCTTCGGAATTTGGCTACAAGTC
      S K L R G K K L K L P P G P M P I P I F G N W L Q V 49
235    GGAGATGATCTTAACCACCGTAATCTCGTCGATTACCGTAAAAAATTTGGTGACCTTCTCCTTCTCAGGATGGGTCAA
      G D D L N H R R N L V D Y A K K F G D L F L L R M G Q 75
313    AGAAACCTAGTGGTGTCTCTTCACCGGATCTGACTAAGGAAGTACTCCACACACAGGGTGTGTAGTTTGGATCGAGA
      R N L V V V S S P D L T K E V L H T Q G V E F G S R 101
391    ACCAGAAACGTTGTTCGATATCTTCACTGGGAAAGGACAAGATATGGTGTTCACAGTTTACGGTGAGCACTGGAGG
      T R N V V F D I F T G K G Q D M V F T V Y G E H W R 127
469    AAGATCGAAGGATTATGACGGTTCCGTTCTTCAACAAGGTTGTGCAAGGAATTCGTGAAGGATGGGAATTTGAA
      K M R R I M T V P F F T N K V V Q R N R E G W E F E 153
547    GCTGCGAGTGTGTGAAGATGTGAAGAAGAATCCAGATTCTGCAACTAAAGGAATGTGTGGAGGAGCGTTTGCAA
      A A S V V E D V K K N P D S A T K G I V L R K R L Q 179
625    TTGATGATGACAACNAATGTTCCGTTATGTTGATAGAAGGTTTCGATAGCGAGGATGATCCCTTTTCGTTAGA
      L M M N M F R I M F D R R F D S E D D P L F V R 205
703    CTCAAGCTTTGAACGGAGAAAGAGTCGATTGGCTCAGAGCTTTGAGTACAACATATGGCGATTTTCATTCTATCCTT
      L K A L N G E R S R L A Q S F E Y N Y G D F I P I L 231
781    AGACCATTCCTTAGAGGCTATTGAAAGATTTGCCAAGATGTGAAGGATCGAAGAATCGCACTCTTCAAAAAGTACTTT
      R P F L R G Y L K I C Q D V K D R R I A L F K K Y F 257
859    GTTGACGAGAGGACAAATTCGAGTACTAAGCCTGCAGGAAGTGACGGATTGAAATGCGCCATCGATCACATTCTT
      V D E R K Q I A S T K P A G S D G L K C A I D H I L 283
937    GAAGCTCAGGATAAGGGTGAATCAACGAAGATAACCTCCTGTACATTGTGCGAGAACATCAACGTCGCTGCTATTGAG
      E A Q D K G E I N E D N V L Y I V E N I N V A A I E 309
1015   ACAACATTGTGGTCAATCGAGTGGGAATTCAGAGCTAGTGAACCATCTGAAATCCAGAATAAGCTAAGGAACGAA
      T T L W S I E W G I A E L V N H P E I Q N K L R N E 335
1093   ATCGACACAGTCTTGGACCGAGGAGTGAAGTACAGAGCCTGTGCTTCACAACTTCATACCTCCAGCCGATGATC
      I D T V L G P G V Q V T E P V L H K L P Y L Q A V I 361
1171   AAGGAAACACTTCGCTAAGAATGGCTATCCACTTCTCGTGCACATGAACCTCCACGACGCTAAGCTCCTGSC
      K E T L R L R M A I P L L V P H M N L H D A K L S G 387
1249   TAGCAGATCCCAGCAGAAAGCAAAATCCTTGTCAATGCCTGGTGGCTAGCAAACAACCTGACAGCTGGAAGAACCTT
      Y D I P A E S K I L V N A W W L A N N P D S W K N P 413
1327   GAAGAGTTTAGACCCGAGAGGTTCTTCGAAGAAGAAGCACATGTGAAGCAAACGGTAATGACTTCAGGTATGTGCCA
      E E F R P E R F F E E E A H V E A N G N D F R Y V P 439
1405   TTTGGTGTGGGCGTAGAAGCTGTCGCGGATCATATTGGCATTGCCCATTTTGGGAATCACCATTTGGTAGTGGTC
      F G V G R R S C P G I I L A L P I L G I T I G R L V 464
1483   CAGAATTTCGACTTCTTCCCTCCAGCAGTCTAAACTCGATACTACTGAGAAGGTGGACAATTCAGTTTACV
      Q N F E L L P P P G Q S K L D T T E K G G Q F S L H 491
1561   ATCCTTAAACACTCGACGATGTTTATGAAACAAGAGCCGTTTAAATTTTTCATaagccaagaaagtaaatgact
      I L K H S T I V M K P R A V * 505

```

Figure 3 Nucleotide sequence and amino acid sequence of *LaC4H* gene cDNA. 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 *LaC4H* gene

2.2 保守结构预测 使用 InterProScan 预测的 *LaC4H* 蛋白的保守结构域为细胞色素 P450 结构域 (cytochrome P450) (IPR001128), 此外 *LaC4H* 蛋白还有 1 个保守位点, 位于多肽链 C 端的细胞色素 P450 保守位点 (440~449 位氨基酸残基, IPR017972), 说明 *LaC4H* 蛋白属于细胞色素 P450 超家族成员。

2.3 信号肽、亚细胞定位及跨膜区预测 使用 SignalP 预测 *LaC4H* 蛋白有信号肽的概率是 0.296, 在第 20 位和 21 位氨基酸之间断裂的概率为 0.303, 但是未达到阈值。利用在线软件 TargetP 预测 *LaC4H* 蛋白的亚细胞定位, 结果表明 *LaC4H* 蛋白属于分泌途径, 可能定位于内质网, 同时预测结果也显示 *LaC4H* 蛋白的 N 末端有 20 个氨基酸组成的一段疏水性 α -螺旋结构, 可以将其锚定在内质网膜上。利用 TMHMM 预测 *LaC4H* 蛋白的跨膜区, 结果表明 *LaC4H* 蛋白在 N 末端 1~20 位氨基酸组成一个跨膜区。

2.4 二级结构和三级结构预测 用 PredictProtein 对 *LaC4H* 蛋白的二级结构进行预测, 结果显示在该蛋白中 α -螺旋和无规则卷曲是其主要结构元件, α -螺旋占 50.50%, 无规则卷曲占 41.19%, β -折叠占 8.32%。用 SWISS-MODEL 预测 *LaC4H* 蛋白的三维结构, 以

人的细胞色素 P450 2A6 (human cytochrome P450 2A6) (PDB ID: 3EBS) 为模板, 序列一致性为 25.22%, 在第 30~503 位氨基酸建模, 模型覆盖率 90%, 根据预测结果, *LaC4H* 蛋白可能以单体的形式发挥作用 (图 4)。

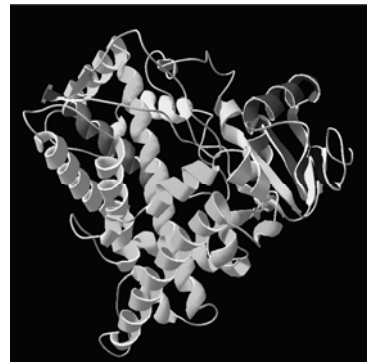


Figure 4 Prediction of three-dimensional structure of *LaC4H* protein

2.5 多序列比对分析 将 *LaC4H* 蛋白的氨基酸序列提交 NCBI 网站通过 blastp 比对 nr 非冗余蛋白数据库, blastp 比对结果显示 *LaC4H* 蛋白与其他植物的 C4H 蛋白有广泛的同源性, 与拟南芥 (*Arabidopsis*

thaliana, AtC4H, AAC99993.1)、油菜 (*Brassica napus*, BnC4H, XP_013748807.1)、烟草 (*Nicotiana tabacum*, NtC4H, ABC69413.1)、丹参 (*Salvia miltiorrhiza*, SmC4H, AJD25173.1)、马铃薯 (*Solanum tuberosum*, StC4H, ABC69046.1)、亚洲棉 (*Gossypium arboreum*, GaC4H, AAG10196.1) 等植物 C4H 蛋白氨基酸序列的一致性分别为 96%、95%、88%、88%、87%、87%。将独行菜 LaC4H 蛋白与这 6 种植物 C4H 蛋白的氨基酸序列进行多序列比对, 结果见图 5, 可以看出 LaC4H 蛋白与其他植物的 C4H 蛋白一样, 具有 P450 保守序列, 如 PFGVGRSCP (血红素结合域, 含有

保守的 FxxGxRxCxG 序列) 和 PPGMPPI [富含脯氨酸的保守区, (P/I) PGPx (G/P) xP 序列], 同时 LaC4H 具有 CYP73A1 的 6 个特征性底物结合位点 (substrate recognition sites, SRS) 的 5 个 (SRS1、SRS2、SRS4、SRS5 和 SRS6)^[26]。

2.6 系统进化树分析 从 NCBI blastp 的比对结果中选取来源于其他植物的 20 条 C4H 蛋白质序列, 包括药用植物丹参、银杏等, 模式植物拟南芥、水稻、玉米等, 以及苔藓植物小立碗藓和蕨类植物卷柏等绘制系统进化树 (图 6)。从图 6 中可以看出 LaC4H 与来源于十字花科植物的拟南芥和油菜的 C4H 处于同

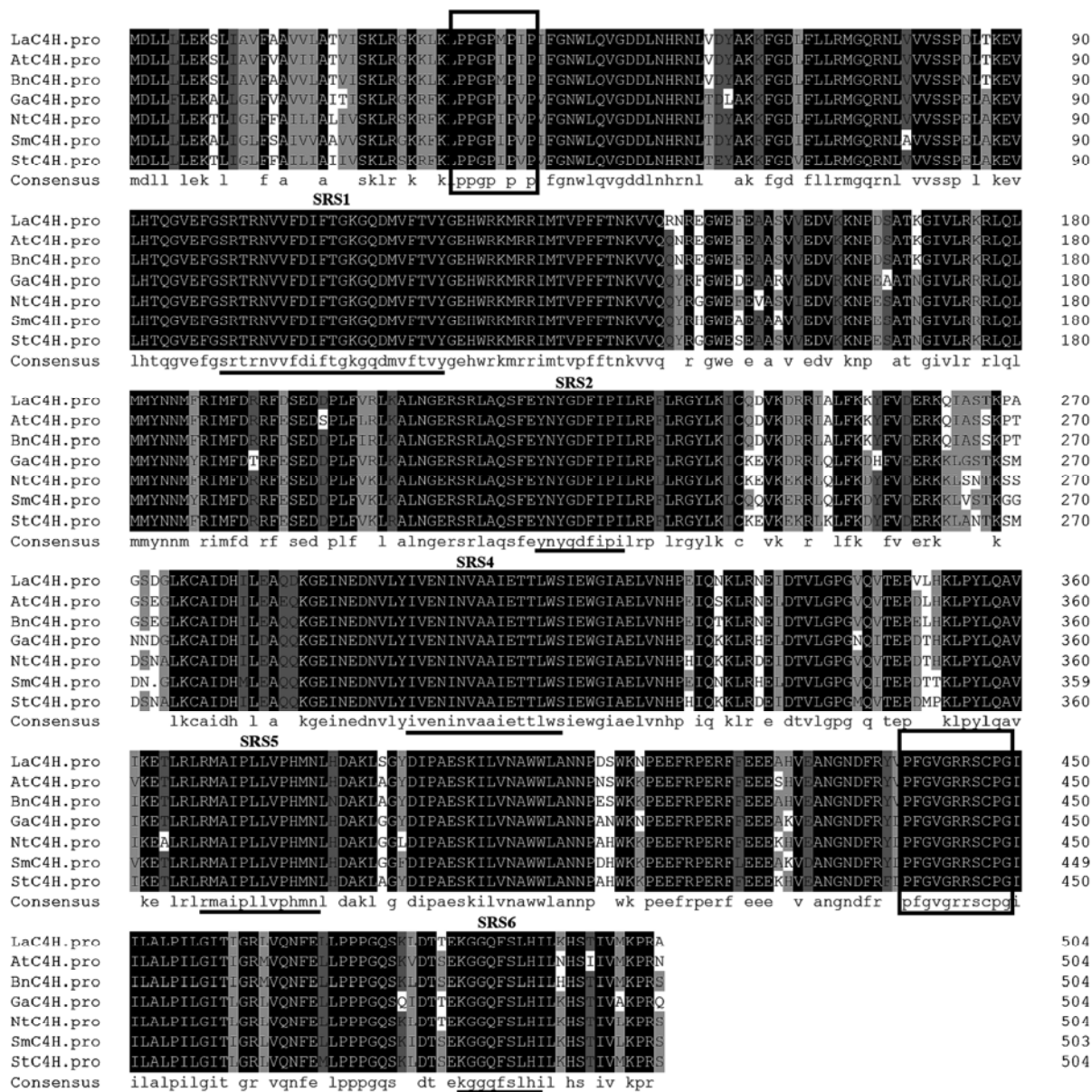


Figure 5 Multiple sequence alignment of LaC4H and C4H from other plant species. The P450 conserved sequences are boxed in black and the underlined regions represent the 5 SRSs; At: *Arabidopsis thaliana*; Bn: *Brassica napus*; Ga: *Gossypium arboreum*; Nt: *Nicotiana tabacum*; Sm: *Salvia miltiorrhiza*; St: *Solanum tuberosum*

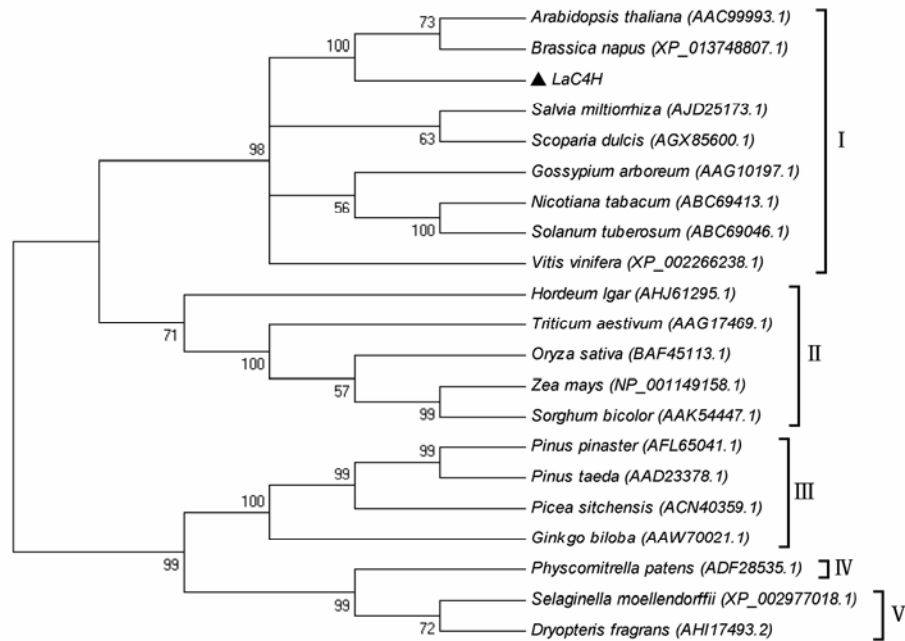


Figure 6 Phylogenetic tree analysis of LaC4H protein with other plant C4H proteins. I: Dicots; II: Monocots; III: Gymnosperms; IV: Mosses; V: Ferns. The accession numbers are showed after the name of species, bootstrap values of the nodes represent the percentage drawn from the bootstrap test

一分支上, 亲缘关系较近, 然后与丹参、亚洲棉、烟草、马铃薯、葡萄等植物的 C4H 聚为一类, 都是被子植物的双子叶植物 (图 6 I); 水稻、玉米、小麦和高粱等禾本科植物的 C4H 聚为一类, 都是被子植物的单子叶植物 (图 6 II); 银杏、北美云杉、海岸松和火炬松的 C4H 聚为一类, 属于裸子植物 (图 6 III); 小立碗藓和香鳞毛蕨、江南卷柏在生物进化上较为原始, 这 3 种植物的 C4H 聚为一类, 分别属于苔藓植物和蕨类植物 (图 6 IV 和 V)。

3 *LaC4H* 原核表达载体的构建

用 *EcoR* 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-*LaC4H*, 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 *EcoR* I 和 *Xho* I 进行双酶切鉴定, 结果见图 7, 电泳结果显示有 6 000 bp 左右的载体条带和 1 500 bp 左右的目的基因条带。将双酶切鉴定正确的单克隆送测序公司测序, 结果显示重组质粒 pET-32a-*LaC4H* 中 *LaC4H* 的序列与目的基因 *LaC4H* 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-*LaC4H* 成功。

4 *LaC4H* 重组蛋白的原核表达与纯化

将测序正确的 pET-32a-*LaC4H* 质粒转化大肠杆菌 BL21 (DE3) 后, 当大肠杆菌生长至 OD_{600} 为 0.6~

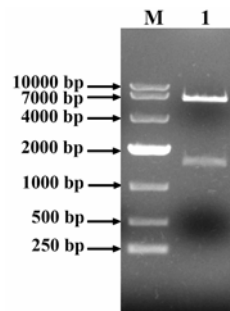


Figure 7 Identification of prokaryotic expression vector pET-32a-*LaC4H* with double digestion. M: DL10000 DNA marker; 1: Double digestion results by *EcoR* I and *Xho* I

0.8 时, 在 28 °C、0.4 mmol·L⁻¹ IPTG、150 r·min⁻¹ 条件下诱导培养 8 h, 对诱导前后的大肠杆菌总蛋白进行 SDS-PAGE 分析。含 pET-32a-*LaC4H* 质粒的表达菌株经 IPTG 诱导后, 在约 75 kD 处出现目的蛋白条带 (图 8A), 因为 *LaC4H* 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), 分子质量应为 77.73 kD。超声破碎菌体后取上清液和沉淀进行蛋白可溶性检查, 结果显示 *LaC4H* 重组蛋白主要以包涵体形式存在于沉淀中 (图 8A)。含 pET-32a 空载体的菌株经 IPTG 诱导后, 未在 75 kD 处出现条带, 而在 20 kD 处出现条带, 因为 pET-32a 空载体的标签序列 (Trx-Tag、His-Tag 和 S-Tag) 在 IPTG 诱导时也会表达融合蛋白, 大小约为 20 kD (图 8B)。结果表明, *LaC4H* 蛋白成

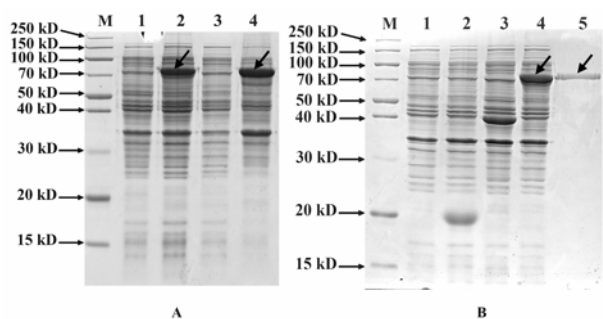


Figure 8 SDS-PAGE analysis of recombinant LaC4H protein. A: The solubility analysis of recombinant LaC4H protein (M: Protein marker; 1: Uninduced *E. coli* containing pET-32a-LaC4H; 2: Induced *E. coli* containing pET-32a-LaC4H; 3: Soluble protein from induced *E. coli* containing pET-32a-LaC4H; 4: Insoluble fraction from induced *E. coli* containing pET-32a-LaC4H. The arrows show the recombinant LaC4H proteins). B: Prokaryotic expression and purification of recombinant LaC4H protein (M: Protein marker; 1: Uninduced *E. coli* containing pET-32a; 2: Induced *E. coli* containing pET-32a; 3: Uninduced *E. coli* containing pET-32a-LaC4H; 4: Induced *E. coli* containing pET-32a-LaC4H; 5: The purified recombinant LaC4H protein. The arrows show the recombinant LaC4H proteins)

功在大肠杆菌 BL21 (DE3) 菌株中表达。

根据 LaC4H 重组蛋白的诱导条件, 扩大培养体积, 用超声裂解菌体, 裂解液经离心取沉淀, 将沉淀溶于含 $8 \text{ mol} \cdot \text{L}^{-1}$ 尿素的结合缓冲液后, 利用 Ni^{2+} 亲和色谱, 采用 Ni-Agarose His 标签蛋白纯化试剂盒纯化 LaC4H 重组蛋白, 最终得到纯化的 LaC4H 重组蛋白 (图 8B)。

5 LaC4H 基因组织特异性与诱导表达分析

利用荧光定量 PCR 检测 LaC4H 基因的组织特异性表达, 结果显示 LaC4H 基因在独行菜所有组织中均有表达, 在茎中表达最高, 叶和花中次之, 根中表达量最低 (图 9A)。MeJA 诱导后独行菜叶中 LaC4H 基因在不同时间的表达水平, 结果见图 9B, 经 MeJA 诱

导后, LaC4H 表达量上升, 4 h 的表达量是 0 h 表达量的 3.57 倍, 随后在 12 h 和 24 h 表达量下降, 在诱导后 48 h 达到最高值, 是 0 h 表达量的 8.30 倍, 到 72 h LaC4H 表达量仅为 0 h 的 1/2。MeJA 诱导后 LaC4H 的表达模式可能与叶中黄酮的积累有关, 通过测定 48 h、72 h 对照组与 MeJA 处理组叶中黄酮的含量, 发现对照组在 48 h 和 72 h 黄酮含量没有明显变化, 而 MeJA 处理组在 48 h 黄酮含量为 $21.15 \text{ mg} \cdot \text{g}^{-1} \text{ DW}$, 比对照组提高了 43.40% ($P < 0.05$), 在 72 h MeJA 处理组黄酮含量为 $12 \text{ mg} \cdot \text{g}^{-1} \text{ DW}$, 比对照组下降了 13.23% (图 9C)。

讨论

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, 在植物体内主要通过苯丙烷代谢途径合成, 肉桂酸-4-羟化酶 (C4H) 是植物苯丙烷代谢途径中的第 2 个关键酶, 该酶在植物中的含量可以影响黄酮、香豆素和木质素等次生代谢产物的合成代谢^[27]。本研究从独行菜幼苗的叶片中克隆了 LaC4H 基因, 通过保守结构域预测发现 LaC4H 蛋白含有 CYP450 的保守结构域以及特征性的底物结合位点^[26], 通过氨基酸序列比对结果发现 LaC4H 蛋白与其他植物 C4H 蛋白的序列高度同源, 属于 CYP450 家族成员。在逆境胁迫下 (如 UV-B 胁迫), 苦荞和矮牵牛中 C4H 等基因表达量的增加使得植物组织中黄酮积累量增加, 以应对逆境胁迫^[11, 12]。如果在拟南芥中 AtC4H 基因突变, 会影响苯丙烷代谢途径, 最终导致木质素含量降低, 此外还会引起拟南芥雄性不育、植株矮小等发育问题, 说明 AtC4H 基因对拟南芥正常的发育过程是必须的^[3]。所以 LaC4H 蛋白可能在独行菜黄酮类化合物的生物合成过程中发挥作用, 具

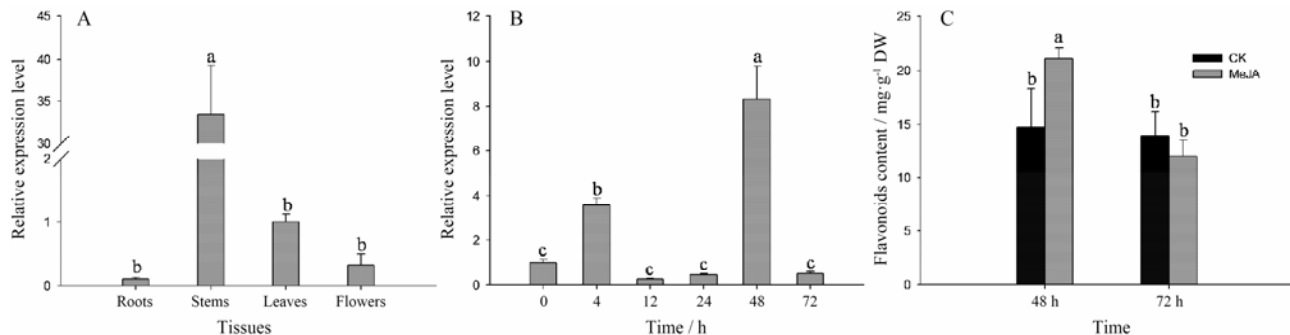


Figure 9 Relative expression level of LaC4H and flavonoids content. A: Relative expression level of LaC4H in different tissues of *L. apetalum*. $P < 0.05$; B: Relative expression level of LaC4H in leaves of *L. apetalum* after MeJA treatment. $P < 0.05$; C: Flavonoids content in leaves of *L. apetalum* at 48 h and 72 h after MeJA treatment. $P < 0.05$. Different lowercase letters indicate significant difference at 0.05 level

有重要的研究意义。植物黄酮类化合物的生物合成一般定位于内质网外膜上, 经生物信息学分析, LaC4H 蛋白的 N 末端, 可能由 20 个氨基酸组成的一段跨膜区, 将 C4H 和 PAL、4CL 形成的多酶复合物锚定在内质网外膜上, 参与黄酮类化合物生物合成途径的催化反应^[28]。植物 C4H 基因的表达具有组织特异性, 拟南芥 *AtC4H* 在茎中表达量最高、根中次之、叶和花中最少^[7], 油菜 *BnC4H* 在茎、叶和花中表达量较高, 根和种子中表达量较低^[29]。荧光定量 PCR 结果显示 *LaC4H* 在独行菜的所有组织中均有表达, 在独行菜茎中的表达量最高, 叶和花中次之, 根中表达量最低, 这与拟南芥和油菜中的研究结果相符, *LaC4H* 在茎中表达量最高可能和茎中的维管组织需要合成较多的木质素单体最终形成木质素有关, *LaC4H* 在叶和花中表达量次之说明独行菜叶和花会积累较多的黄酮类化合物。

MeJA 作为一种在植物体内广泛存在的植物激素, 作为信号分子参与植物的抗逆反应 (如病虫害、机械伤害、低温、干旱等), 还能够提高植物次生代谢水平, 增加黄酮、三萜等次生代谢产物的积累^[30,31]。有报道 MeJA 能够诱导人参中三萜类化合物的生物合成途径, 增加其药效物质人参皂苷的含量^[31]。在贯叶金丝桃的悬浮培养细胞中加入 MeJA, 能够增加悬浮细胞中黄酮的含量^[30]。所以推测 MeJA 能够诱导独行菜黄酮类化合物的生物合成途径, 提高黄酮的含量, 通过检测 MeJA 诱导后独行菜叶中 *LaC4H* 基因的表达量和黄酮的含量, 发现诱导后 *LaC4H* 表达量上升, 随后 12 h 和 24 h 表达量有所下降, 到 48 h 表达量达到最高值, 在 48 h MeJA 处理组的黄酮含量比对照组提高了 43.40%, 而到 72 h *LaC4H* 的表达量下降到 0 h 的 1/2, 而且 MeJA 处理组的黄酮含量比对照组下降了 13.23%, 说明独行菜叶片中 *LaC4H* 的表达量与黄酮积累量之间呈正相关, 这些结果都说明 *LaC4H* 与独行菜黄酮类化合物的生物合成途径相关。

本研究首次在大肠杆菌中表达了独行菜 LaC4H 蛋白, 希望在蛋白水平研究独行菜 LaC4H 蛋白的生物学功能。Zhang 等^[32]通过构建 pET-32a-SmC4H 表达载体, 成功诱导表达了丹参 SmC4H 蛋白, 但是表达的 C4H 蛋白以包涵体的形式存在。Zeng 等^[33]以 pET6×HN 为载体, 在大肠杆菌 *Transtetta* (DE3) 中诱导表达了桂花 OfC4H 蛋白, 但是表达的 OfC4H 融合蛋白的溶解度很低, 主要以包涵体形式存在。本研究通过构建原核表达载体 pET-32a-LaC4H, 成功在大肠杆菌 BL21 (DE3) 菌株中表达了独行菜 LaC4H 蛋

白, 但是重组蛋白 LaC4H 在表达菌中主要以不溶的包涵体形式存在, 这可能是由于 LaC4H 蛋白的 N 末端含有一段作为锚定信号的跨膜区, 导致错误的蛋白折叠, 最终使 LaC4H 重组蛋白在大肠杆菌中以不溶的包涵体形式存在^[34], 而且在大肠杆菌表达系统中不存在维持 C4H 蛋白催化活性必须的 NADPH-细胞色素 P450 还原酶 (cytochrome P450 reductase, CPR), 会造成表达的 CYP450 蛋白没有活性, 最终导致 CYP450 蛋白以包涵体的形式存在, 如果要在大肠杆菌中表达有生物学活性的 C4H 蛋白并且进行功能分析, 就需要构建 C4H 和 CPR 融合蛋白, 使这两种酶在大肠杆菌中共同表达, 利用这种方法, 成功在大肠杆菌中表达了长春花 C4H 蛋白并进行了功能鉴定^[35]。在酵母表达系统中, 酵母具有内源性 CPR, 可以使电子向 CYP450 蛋白传递, 所以 CYP450 蛋白可以在酵母系统中进行表达并且具有催化功能, 目前利用酵母表达系统成功表达了 50 个以上的植物 CYP450 蛋白并进行功能鉴定^[36]。Yao 等^[37]利用 pYES-CsC4H 真核表达载体成功在酿酒酵母中表达了茶树的 C4H 蛋白并且获得了有生物学活性的茶树 C4H 蛋白。此外还有利用杆状病毒-昆虫细胞表达系统表达 CYP450 蛋白并且进行功能鉴定, 但是昆虫表达系统操作程序复杂, 对技术要求比较高, 目前酵母表达系统仍然是 CYP450 蛋白的主要表达系统。下一步研究可以通过包涵体溶解与复性的方法, 或者利用酵母真核表达系统表达 LaC4H 蛋白, 获得有活性的 LaC4H 蛋白, 在体外研究其酶学特征及催化活性, 也可以将纯化的 LaC4H 重组蛋白制备多克隆抗体, 研究 *LaC4H* 在独行菜蛋白水平的表达情况, 为进一步研究 *LaC4H* 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。

References

- [1] Winkelshirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology [J]. *Plant Physiol*, 2001, 126: 485-493.
- [2] Park N, Park J, Park S. Overexpression of cinnamate 4-hydroxylase gene enhances biosynthesis of decursinol angelate in *Angelica gigas* hairy roots [J]. *Mol Biotechnol*, 2011, 50: 114-120.
- [3] Schilmiller A, Stout J, Weng J, et al. Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis* [J]. *Plant J*, 2009, 60: 771-782.
- [4] Russell DW, Conn EE. The cinnamic acid 4-hydroxylase of

- pea seedlings [J]. Arch Biochem Biophys, 1967, 122: 256–258.
- [5] Teutsch H, Hasenfratz M, Lesot A, et al. Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway [J]. Proc Nat Acad Sci USA, 1993, 90: 4102–4106.
- [6] Mizutani M, Ward E, Dimairo J, et al. Molecular cloning and sequencing of a cDNA encoding mung bean cytochrome P450 (P450 C4H) possessing cinnamate 4-hydroxylase activity [J]. Biochem Biophys Res Commun, 1993, 190: 875–880.
- [7] Mizutani M, Ohta D, Sato R. Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis* and its expression manner in planta [J]. Plant Physiol, 1997, 113: 755–763.
- [8] Sewalt V, Ni W, Blount J, et al. Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of *L*-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase [J]. Plant Physiol, 1997, 115: 41–50.
- [9] Millar DJ, Marianne L, Georgina D, et al. Introduction of sense constructs of cinnamate 4-hydroxylase (CYP73A24) in transgenic tomato plants shows opposite effects on flux into stem lignin and fruit flavonoids [J]. Phytochemistry, 2007, 68: 1497–1509.
- [10] Belllalong D, Cusumano J, Meyer K, et al. Cinnamate-4-hydroxylase expression in *Arabidopsis* (regulation in response to development and the environment) [J]. Plant Physiol, 1997, 113: 729–738.
- [11] Chen HH, Yuan MQ, Li SJ, et al. Cloning of cinnamate 4-hydroxylase gene (*C4H*) from tartary buckwheat (*Fagopyrum tararicum*) and its tissue-specific expression under UV-B stress during seed germination [J]. J Agric Biotechnol (农业生物技术学报), 2013, 21: 137–147.
- [12] Ryan KG, Swinny EE, Markham KR, et al. Flavonoid gene expression and UV photoprotection in transgenic and mutant *Petunia* leaves [J]. Phytochemistry, 2002, 59: 23–32.
- [13] Singh K, Kumar S, Rani A, et al. Phenylalanine ammonia-lyase (PAL) and cinnamate 4-hydroxylase (C4H) and catechins (flavan-3-ols) accumulation in tea [J]. Funct Integr Genomics, 2009, 9: 125–134.
- [14] Li HW, Zheng XK, Gong JH, et al. Research progress in chemical constituents of *Lepidium apetalum* and *Descurainia sophia* and their pharmacological activities [J]. Drug Eval Res (药物评价研究), 2013, 36: 235–240.
- [15] Zhou XD, Tang LY, Zhou GH, et al. Advances on *Lepidium semen* and *Descurainiae semen* [J]. China J Chin Mater Med (中国中药杂志), 2014, 39: 4699–4708.
- [16] Cao WG, Liu ZQ, Shao Y, et al. A progress in pharmacological research of flavonoids [J]. Acta Bot Boreal-Occid Sin (西北植物学报), 2003, 23: 2241–2247.
- [17] Zhou Q, Zhao HX, Li PP, et al. *De novo* characterization of the seed transcriptome of *Lepidium apetalum* Willd [J]. J Chin Biotechnol (中国生物工程杂志), 2016, 36: 38–46.
- [18] Yuan JW, Huang DH, Xu DS, et al. Cloning and analysis of seed coat mucilage-related gene *MUM4* from *Lepidium perfoliatum* [J]. Acta Bot Boreal-Occid Sin (西北植物学报), 2013, 33: 1940–1952.
- [19] Cao J, Xu DS, Huang DH, et al. Cloning, characterization, and functional analysis of seed coat mucilage-related gene *TTG1* from *Lepidium perfoliatum* [J]. Plant Sci J (植物科学学报), 2014, 32: 371–382.
- [20] Zhao QZ, Shi QJ, Li Q. Cloning and plant expression vector construction of transcription factor CBF gene in *Lepidium apetalum* Willd [J]. Seed (种子), 2012, 31: 4–6.
- [21] Ma LG, Zhao L, Li YC, et al. Cloning, sequence analysis and prokaryotic expression of GGPS gene from *Lepidium apetalum* [J]. World Sci Technol/Mod Tradit Chin Med (世界科学技术: 中医药现代化), 2015, 17: 485–491.
- [22] Czechowski T, Stitt M, Altmann T, et al. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis* [J]. Plant Physiol, 2005, 139: 5–17.
- [23] Ge FW, Tian YZ, Zeng WJ, et al. Cloning and RT-PCR analysis of eEF-1a gene fragment from *Lepidium apetalum* Willd [J]. J Xianjiang Norm Univ (Nat Sci) (新疆师范大学学报: 自然科学版), 2014, 33: 22–26.
- [24] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method [J]. Methods, 2001, 25: 402–408.
- [25] Zheng XK, Hou QW, Li M, et al. Extracting technology of total flavones from *Selaginella tamariscina* (Beauv) spring [J]. Chin J New Drugs (中国新药杂志), 2011, 20: 1509–1513.
- [26] Schoch G, Attias R, Ret M, et al. Key substrate recognition residues in the active site of a plant cytochrome P450, CYP73A1 [J]. Eur J Biochem, 2003, 270: 3684–3695.
- [27] Fahrendorf T, Dixon RA. Stress responses in alfalfa (*Medicago sativa* L.) XVIII: molecular cloning and expression of the elicitor-inducible cinnamic acid 4-hydroxylase cytochrome P450 [J]. Arch Biochem Biophys, 1993, 305: 509–515.
- [28] Saslowsky DE, Ujwala W, Winkel BSJ. Nuclear localization of flavonoid enzymes in *Arabidopsis* [J]. J Biol Chem, 2005,

- 280: 23735–23740.
- [29] Chen AH, Chai YR, Li JN, et al. Molecular cloning of two genes encoding cinnamate 4-hydroxylase (C4H) from oilseed rape (*Brassica napus*) [J]. *J Biochem Mol Biol*, 2007, 40: 247–260.
- [30] Wang J, Qian J, Yao L, et al. Enhanced production of flavonoids by methyl jasmonate elicitation in cell suspension culture of *Hypericum perforatum* [J]. *Bioresour Bioproc*, 2015, 2: 1–9.
- [31] Choi DW, Jung J, Ha YI, et al. Analysis of transcripts in methyl jasmonate-treated ginseng hairy roots to identify genes involved in the biosynthesis of ginsenosides and other secondary metabolites [J]. *Plant Cell Rep*, 2005, 23: 557–566.
- [32] Zhang TT, Wang CL, Han RL, et al. Construction of pET32a prokaryotic expression vector of C4H gene in *Salvia miltiorrhiza* [J]. *J Northwest Agric Forest Univ (Nat Sci)* (西北农林科技大学学报:自然科学版), 2011, 39: 158–162.
- [33] Zeng XL, Zheng RR, Luo J, et al. Cloning and characterization of cinnamate 4-hydroxylase (C4H) genes from *Osmanthus fragrans* [J]. *Acta Hort Sin* (园艺学报), 2016, 43: 525–537.
- [34] Singh SM, Panda AK. Solubilization and refolding of bacterial inclusion body proteins [J]. *J Biosci Bioeng*, 2005, 99: 303–310.
- [35] Hotze M, Schröder G, Schröder J. Cinnamate 4-hydroxylase from *Catharanthus roseus* and a strategy for the functional expression of plant cytochrome P450 proteins as translational fusions with P450 reductase in *Escherichia coli* [J]. *FEBS Lett*, 1995, 374: 345–350.
- [36] Tu J, Zhu J, Cheng KL. Heterologous expression systems of plant cytochrome P450 [J]. *J Chin Biotechnol* (中国生物工程杂志), 2003, 23: 32–37.
- [37] Yao SB, Wang WZ, Li MZ, et al. The gene cloning and expression analysis of C4H in tea plant (*Camellia sinensis*) [J]. *J Tea Sci* (茶叶科学), 2015, 35: 35–44.