

• 药材与资源 •

姜黄苯丙氨酸解氨酶基因的克隆与序列分析

刘建福^{1,2}, 钟书淳², 王明元², 唐源江², 范燕萍^{1*}, 陈钦²

1. 华南农业大学园艺学院, 广东 广州 510642

2. 华侨大学 生物工程与技术系, 福建 厦门 361021

摘要: 目的 对姜黄 *Curcuma longa* 苯丙氨酸解氨酶(phenylalanine ammonia-lyase, PAL)基因(*CurPAL*)全长进行克隆并开展生物信息学分析, 为姜黄次生代谢产物的生物合成机制研究奠定基础。方法 利用RT-PCR和RACE相结合的方法, 以姜黄叶片cDNA为模板, 克隆获得*CurPAL*全长, 进行生物信息学分析。结果 克隆的*CurPAL*(登录号为KJ780359)全长cDNA为1293 bp, 其中包括5'-UTR 243 bp, 3'-UTR 123 bp, 含有1个927 bp的完整开放阅读框(opening reading frame, ORF), 编码308个氨基酸; 预测蛋白质相对分子质量为33 000, 理论等电点pI为5.76; *CurPAL*蛋白性质稳定, 为可溶性蛋白; *CurPAL*编码的蛋白质序列包含与夹竹桃PAL蛋白质相同的脱氨基位点和催化活性位点, 氨基酸序列多重比较发现, *CurPAL*编码的氨基酸序列与中国李、浙江红山茶、拟南芥、辣椒、小果野蕉PAL氨基酸序列的同源性达到75%以上; 与姜目类的植物(如小果野蕉)的PAL聚为一支, 说明两者亲缘关系较近; 通过蛋白质二维、三维结构的预测, 表明*CurPAL*蛋白为全α型蛋白, 由同源四聚体构成。结论 首次克隆并获得*CurPAL*基因全长cDNA, 为姜黄素生物合成途径的阐明和改善中药材品质提供科学依据。

关键词: 姜黄; *CurPAL*; 苯丙氨酸解氨酶; 基因克隆; 序列分析; RT-PCR; RACE

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2014)21-3141-08

DOI: 10.7501/j.issn.0253-2670.2014.21.019

Cloning and sequence analysis of phenylalanine ammonia-lyase gene from *Curcuma longa*

LIU Jian-fu^{1,2}, ZHONG Shu-chun², WANG Ming-yuan², TANG Yuan-jiang², FAN Yan-ping¹, CHEN Qin²

1. College of Horticulture, South China Agricultural University, Guangzhou 510642, China

2. Department of Biological Engineering and Technology, Huaqiao University, Xiamen 361021, China

Abstract: Objective This study aimed at cloning the phenylalanine ammonia-lyase (PAL) gene from *Curcuma longa* (*CurPAL*) and analyzing the bioinformatics. **Methods** *PAL* gene was cloned by RT-PCR and RACE strategy with the template of RNA extracted from *C. longa* leaves. The bioinformatic analysis of this gene and its corresponding protein were performed. **Results** One unique sequence of PAL, named as *CurPAL* (GenBank NO. KJ780359), was cloned from *C. longa*. The full-length of *CurPAL* cDNA was 1 293 bp, including 243 bp of 5'-UTR, 123 bp of 3'-UTR, and 927 bp of ORF encoding 308 amino acids. The molecular weight and theoretical isoelectric point (pI) of the deduced *CurPAL* protein were 33 000 and 5.76, respectively. The protein of *CurPAL* was stable and soluble. The domination sites and catalytic active sites in PAL protein of *Nerium oleander* were also found in *CurPAL*. It was found that the amino acid sequence of *CurPAL* had more than 75% homology with PAL of *Prunus salicina*, *Camellia chekiangoleosa*, *Capsicum chinense*, and *Musa acuminata* via multiple alignments. It revealed that *CurPAL* had closer relationship with PALs from *Zingiberaceae* plants than from other plants by phylogenetic tree analysis. Secondary and tertiary structures indicate that *CurPAL* is a full α protein contained by homotetramer. **Conclusion** The cDNA encoding *PAL* from *C. longa* is cloned and reported for the first time. This work provides a scientific basis for exploring the biosynthetic pathway of the medicinal ingredient and improving its quality in *C. longa*.

Key words: *Curcuma longa* L.; *CurPAL*; phenylalanine ammonia-lyase; gene cloning; sequence analysis; RT-PCR; RACE

收稿日期: 2014-07-03

基金项目: 国家自然科学基金资助项目(31101512); 中央高校基本科研业务费专项资金项目(JB-ZR1151); 泉州市科技计划项目(2014Z111)

作者简介: 刘建福(1975—), 男, 福建泉州人, 博士研究生, 副教授, 硕士生导师, 从事植物生理及代谢调控研究。E-mail: jianfu@hqu.edu.cn

*通信作者 范燕萍, 教授, 博士生导师。Tel:(021) 85288266 E-mail: fanyanping@scau.edu.cn

姜黄 *Curcuma longa* L. 是我国传统中药, 是姜科姜黄属多年生草本植物。姜黄属植物全世界有 60 多种, 盛产于东南亚和澳大利亚北部^[1], 中国主要分布在东南至西南部。姜黄属植物主要包含挥发油和姜黄素类, 其中姜黄素是姜黄的主要成分^[2]。研究表明, 姜黄素具有抗炎、抗氧化、抗癌等药理作用^[3-5], 可以改善糖尿病、代谢性疾病、自身免疫疾病、动脉粥样硬化、关节炎、中风、外周神经病、肠炎和脑外伤等多种疾病症状^[6]。姜黄素对阿尔茨海默病等神经变性病具有很强的神经保护作用^[7]。

苯丙烷类代谢途径是一切含苯丙烷骨架物质的直接或间接合成途径。苯丙烷类代谢可生成类黄酮、木质素等多种次生代谢产物, 这些次生产物在植物的生长发育、抗病、抗逆反应中起着重要的作用^[8]。姜黄素的生物合成是通过苯丙烷代谢途径进行的, 苯丙氨酸解氨酶 (PAL) 是植物苯丙烷类代谢第一步反应的关键酶和限速酶, 它催化 L-苯丙氨酸解氨酶生成反式肉桂酸, 再由反式肉桂酸生成香豆酸、阿魏酸、芥子酸等中间产物, 进一步形成对香豆酸辅酶 A 和阿魏酸辅酶 A, 然后在姜黄素合成酶催化下与丙二酰辅酶 A 生成姜黄素^[9-10]。

PAL 基因与植物酚类、类黄酮、木质素等次生代谢产物的合成密切相关^[11]。植物 *PAL* 基因为多基因家族所编码, *PAL* 基因的表达受发育和环境信号 (如温度、光质、机械损伤、激素和营养等) 的调节, 表现出严格的组织和发育特异性^[12-13]。目前已从多种植物中克隆到 *PAL* 基因并进行了表达分析^[14-15], 但在姜黄中尚未见到包括 *PAL* 基因在内的苯丙烷类代谢途径中相关酶基因的报道。本研究以姜黄叶片为材料, 开展姜黄 *PAL* 基因 (*CurPAL*) 全长 cDNA 的克隆, 并进行序列分析和系统进化等生物信息学方面的研究, 为 *PAL* 基因功能分析和遗传转化奠定基础, 为姜黄次生代谢产物的生物合成机制提供理论依据。

1 材料

姜黄叶片于 2013 年 10 月采自华侨大学温室, 经唐源江副研究员鉴定为姜黄 *Curcuma longa* L., 清洗干净后迅速放入液氮中冷冻, 置于 -80 °C 保存备用。

M-MLV RTase cDNA Synthesis Kit, pMD19-T 载体, LA Taq Polymerase, DH5α 感受态, DNA 标准相对分子质量 DNA Marker DL2000, Takara 5'-Full RACE Kit with TAP, Takara 3'-Full RACE Core Set with PrimeScript™ RTase 均购于宝生物工

程 (大连) 有限公司; 各引物均由上海生工生物工程技术服务有限公司合成。柱式植物 RNAout 试剂盒购于北京天恩泽基因科技有限公司, TIANScript cDNA 第一链合成试剂盒购自天根生化科技有限公司, 逆转录引物为试剂盒提供的 oligo (dT)₁₅。

2 方法

2.1 总 RNA 提取及 cDNA 第一链合成

取叶样品液氮中研磨, 根据北京天恩泽基因科技有限公司的柱式植物 RNAout 试剂盒说明书提取姜黄总 RNA, -80 °C 保存, 分别用紫外分光光度计和电泳检测其浓度和纯度。以姜黄总 RNA 为模板, 根据 TIANScript cDNA 第一链合成试剂盒进行反转录, 合成 cDNA 第一链, 反转录产物置于 -80 °C 保存, 备用。

2.2 *PAL* 基因 cDNA 片段的扩增

根据小果野蕉 *Musa acuminata* Colla、生姜 *Zingiber officinale* Rosc 等芭蕉目植物 *PAL* 基因的序列, 设计一对保守序列引物 PAL-F 和 PAL-R(表 1), 以上述合成的 cDNA 第一链为模板, 进行 PCR 扩增。RT-PCR 反应体系为 0.4 μL 的 TaKaRa LA Taq (5 U/μL), 引物 PAL-F (10 μmol/L) 和 PAL-R (10 μmol/L) 各 1.0 μL, cDNA 1.0 μL, 2.0 μL 的 dNTP (2.5 mmol/L), 2.5 μL 的 10×缓冲液, 17.1 μL 的 ddH₂O。PCR 反应条件: 94 °C 预变性 5 min; 94 °C 变性 30 s, 53 °C 退火 30 s, 72 °C 延伸 70 s, 30 个循环; 72 °C 延伸 7 min。

2.3 *PAL* 基因 cDNA 的 RACE 扩增

姜黄 *PAL* 全长 cDNA 的 5'、3' 端的克隆采用 RACE 的方法。根据测序得到的 *PAL* 片段序列信息设计 RACE 引物, 在 cDNA 5' 和 3' 端分别设计 2 条特异性引物 (表 1), 并利用 Takara 5'-Full RACE Kit with TAP 和 Takara 3'-Full RACE Core Set with PrimeScript™ RTase 分别进行 2 次巢式扩增获得姜黄 *PAL* 基因的 5' 和 3' 末端序列, 按照试剂盒使用说明书进行实验操作。RACE 反应体系为 cDNA 1.0 μL、dNTP (2.5 mmol/L) 2.0 μL、上、下游引物 (10 μmol/L) 各 1.0 μL、缓冲液 (10×) 2.5 μL、Dream Taq 酶 0.15 μL 和 ddH₂O 17.35 μL。PCR 反应条件为 95 °C 预变性 5 min; 95 °C 变性 35 s, 58 °C 退火 40 s, 72 °C 延伸 2.5 min, 35 个循环, 72 °C 延伸 10 min。

2.4 *PAL* 基因 ORF 扩增

根据拼接全长的 ORF 区域, 设计一对保守序列引物 ORF-F 和 ORF-R (表 1), 以上述合成的 cDNA 第一链为模板, 进行 PCR 扩增。反应体系为 0.75 μL

的 Trans TaqTM HiFi DNA polymerase, 引物 ORF-F (10 μmol/L) 和 ORF-R (10 μmol/L) 各 1.0 μL, cDNA 2.0 μL, 4.0 μL 的 dNTP (2.5 mmol/L), 5.0 μL 的 10×Trans TaqTM HiFi 缓冲液 II, 36.25 μL 的 ddH₂O。PCR 反应条件: 94 °C 预变性 3 min; 94 °C 变性 30 s, 65 °C 退火 30 s, 72 °C 延伸 60 s, 30 个循环; 72 °C 延伸 7 min。

2.5 PCR 产物的回收、克隆与测序

用 DNA 纯化试剂盒回收 PCR 产物并与 PMD19-T 载体 (Takara, 日本) 进行连接, 然后将连接产物转化大肠杆菌 DH5α, 37 °C 过夜培养, 通过在含有 Amp 的 LB 培养基上进行蓝白斑筛选, 随机挑取白色菌落进行 PCR 鉴定, 将获得的阳性克隆进一步提取质粒进行酶切鉴定后送上海生工生物工程技术服务有限公司进行测序。

表 1 姜黄 PAL 基因克隆的引物序列

Table 1 Primer sequence for PAL cloning of *C. longa*

引物名称	引物序列 (5'-3')
PAL-F	ATTGCAAAAGCCGAAGCAGGAC
PAL-R	ACCGGATTGCCAGAAACTGG
3'-PAL-Outer	ACACCCATCGCGTCTCCAT
3'-PAL-Inner	GCTTCGCCACCAAATCCATC
5'-PAL-Outer	AGGGCAATCCGTTGTTGAAAAGTCG
5'-PAL-Inner	TGAGGGGATTATCGTCACGGAGTTG
ORF-F	GCGCG <u>GATCC</u> ATGGACAACACTCGTCTCGCGATC 下划线为 BamHI 酶切位点
ORF-R	GCGCA <u>AAGCTTT</u> AGCAGATGGGAATTGGAGCGCC 下划线为 HindIII 酶切位点

2.6 生物信息学分析

用 DNASTAR 进行基因片段拼接, 利用 NCBI-protein blast 在线分析软件分析姜黄 PAL 基因与其他物种的同源性; 用在线网站 (<http://isoelectric.ovh.org/>) 预测基因编码氨基酸的等电点和蛋白质相对分子质量; 利用 NCBI/ORFFinder 的 ORF 分析核酸序列的开放阅读框架, 用 ExPASy 软件预测编码蛋白的基本性质和疏水性; 应用 TM-HMM 软件分析编码蛋白的跨膜区; 用 SignalP 软件预测编码蛋白信号肽; 用 COILS Server 软件进行 Coil 区分析; 用 TargetP 软件预测编码蛋白的亚细胞定位; 用 PORTER 软件预测蛋白质二维结构; 用 SWISS-MODEL 软件预测蛋白质三维结构; 用 Mega 5.0 软件中的 Clustal W 软件对不同 PAL 氨

基酸序列进行多重序列比对以后, 再使用最大简约法构建系统发育树。

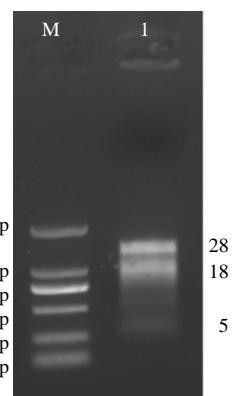
3 结果与分析

3.1 姜黄叶片总 RNA 的提取和检测

取 5 μL 提取后的总 RNA, 在 1.0% 的琼脂糖凝胶上进行电泳检测, 由图 1 可以看出提取的姜黄叶片总 RNA 完整性好; 总 RNA 在波长 260 nm 和 280 nm 处吸光度比值为 2.0, 纯度比较高。DNA 污染少, 可以进行后续的反转录反应。

3.2 姜黄 PAL 基因克隆及序列特征

根据小果野蕉、生姜等姜科植物 PAL 基因的序列, 设计一对引物, 从姜黄叶片 cDNA 中扩增出一条 250~500 bp 长的特异性条带 (图 2), 经测序该片段长度为 419 bp; Blastx 比对分析表明其与多个 PAL 基因有较高的同源性, 确认为姜黄 PAL 基因序列。

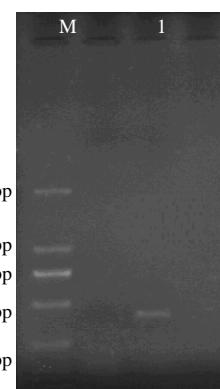


M-Marker 1-姜黄总 RNA 的 PCR 产物

M-Marker 1-PCR product of total RNA

图 1 姜黄叶片总 RNA

Fig. 1 Total RNA of *C. longa* leaves



M-Marker 1-CurPAL 基因的 PCR 产物

M-Marker 1-PCR product of CurPAL

图 2 CurPAL 基因保守序列

Fig. 2 Conserved sequence of CurPAL

采用 RACE 技术进行 cDNA 两端序列的扩增, 经过 3'-RACE 获得 1 条 1 204 bp 的片段(图 3-A), 经过 5'-RACE 获得了 1 条在 212 bp 的片段(图 3-B), 测序和拼接结果表明其为 *PAL* 基因的 3' 端和 5' 端序列。姜黄 *PAL* 基因的 cDNA 全长序列

为 1 293 bp, Blastp 比对结果表明, 该基因推测的氨基酸序列与小果野蕉、菠萝 *Ananas comosus* (Linn.) Merr. 的苯 *PAL* 有 90% 以上的相似性, 推测其为姜黄的 *PAL* 基因。

姜黄 *PAL* 基因全长 cDNA 命名为 *CurPAL*, 经 DNASTar 及 ORF finder 软件分析表明, *CurPAL* 的 cDNA 由 243 bp 的 5'-UTR、123 bp 的 3'-UTR, 包括 12 bp 的 polyA 结构及 927 bp 编码 308 个氨基酸的开放阅读框 (ORF) 组成(图 4)。244~246 bp 的 ATG 及邻近序列 (GCGATGG) 符合真核生物起始密码 ATG 预测保守序 (A/GXXTGG); 在 3'-UTR 处有典型的加尾信号 “AATAAA”(图 4); 在 *CurPAL* 序列中, 包含 L6、V23、A7 和 L22 等脱氨基氨基酸残基(图 4); 它们催化活性位点 N29、G30、N145、G146、N300、G301、HNQDV (78~82), 这些活性位点与夹竹桃 *Nerium oleander* L.^[16] 中报道的相同, 说明 *CurPAL* 是 *PAL* 蛋白质家族成员之一。该基因序列在 GenBank 数据库的登录号为 KJ780359。

3.3 CurPAL 氨基酸序列同源性分析及进化树分析

通过 NCBI 上的蛋白序列 Blastp 和 ClustalW 软件分析表明, *CurPAL* 的氨基酸与其他植物 *PAL* 编码的氨基酸序列具有较高的同源性(大于 75%, 图 5), 如与中国李 *Prunus salicina* Lindl.、浙江红山茶 *Camellia chekiangoleosa* Hu 和拟南芥 *Arabidopsis thaliana* (L.) Heynh. 氨基酸序列的一致性达到 78%, 与辣椒 *Capsicum annuum* L. 的一致性达到 79%, 与小果野蕉的一致性达到 84%。

为研究 *CurPAL* 与其他植物 *PAL* 之间的亲缘关系, 用 Clustal W 对氨基酸序列比对后, 利用 Mega 5.0 软件通过 Neighbor-Joining (NJ) 法构建氨基酸序列的进化树(图 6), 结果显示, 姜黄与野蕉 *Musa balbisiana* Colla 和小果野蕉聚为一支, 均为芭蕉目, 其与野蕉的亲缘关系最近; 双子叶植物中罂粟目、大戟目、蔷薇目、管状花目、侧膜胎座目聚为另一支。从进化树上看, 相同科或类群的植物聚为一类, 但不同植物间 *PAL* 也有很高的同源性, 这也可推测 *PAL* 在进化过程中是相当保守的, *CurPAL* 与野蕉在结构和功能上具有一定的相似性。

3.4 CurPAL 蛋白理化性质分析

运用 ProtParam 工具分析 *CurPAL* 编码蛋白的理化性质, 结果表明, *CurPAL* 编码蛋白相对分子质量

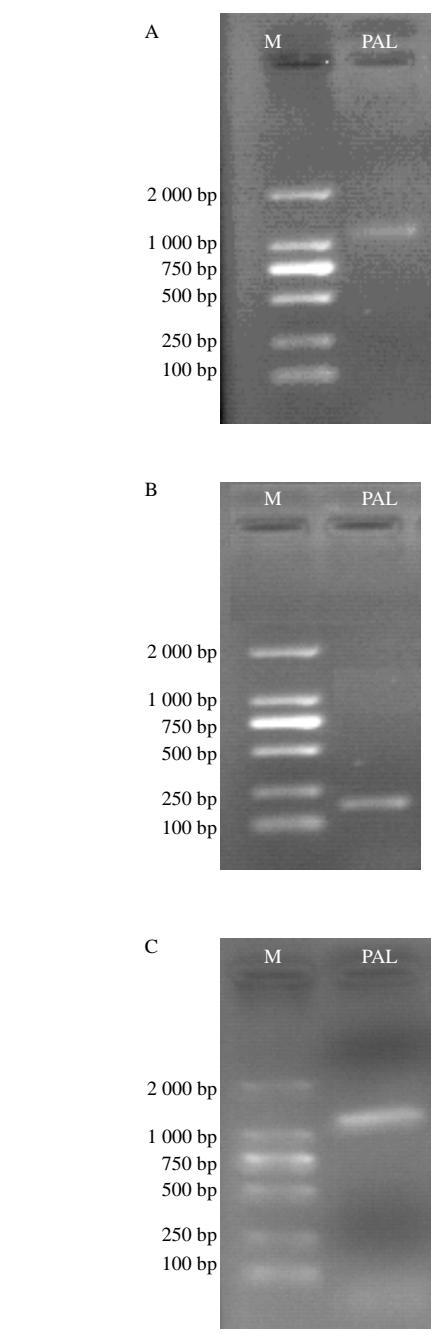


图 3 RACE 扩增的 *CurPAL* 基因

Fig. 3 Result of RACE amplification of *CurPAL* gene

1GAA ACC TAC ATC AAG CAC GCC AAG CTC CAC GAT CAG GAT CCG CTG CAG AAG CCG AAG
 61 CAG GAT CGC TAC GCG CTC CGA ACG TCG CCG CAG TGG CTT GGC CCT CAA ATA GAA GTC CTC
 121 CGC TTC GCC ACC AAA TCC ATC GAA CGC GAG ATC AAC TCC GTG AAC GAT ATC CCC CTC ATC
 181 GAC GTC GCC CGG AAC AAG GCC CTC CAC GGC AAC TTT CAG GGC ACA CCC ATC GG CGTC
 241 TCC ATG GAC AAC ACT CGT CTC GCG ATC GCC GCA ATC GGC AAG CTC ATG TCC GCC CAA TTC
 M D N T R **L**A I A A I G K L M F A Q F
 301 TCC GAA CTC GTC AAC GAC TTT TAC AAC AAC GGA TTG CCC TCT ATT CTC TCC GGC GGA CGG
 S E **L**V N D F Y N N G L P S N L S G G R
 361 AAC CTT AGC CTT GAC TAC GGT TTC AAG GGG GCG GAA ATC GCG ATG GCC TAT TGC TCC
 N L S L D Y G F K G A E I A M A A Y C S
 421 GAG CCC CAG TTT CTG GGC AAT CCG GTG ACC AAC CAC GTT CAG AGT GCG GAG CAG CAT AAC
 E P Q F L G N P V T N H V Q S A E Q **H**N
 481 GAG GAC GTG AAC TCG CTG GGG CTG ATC TCC GCG AGG AAG AGC GCT GAG GCC GTC GAC ATC
Q**D**V N S L G L I S A R K T A E A V D I
 541 CTC AAG CTC ATG TCG GCG ACC TAC TTG GTG GCC CTC TGC CAA GCC ATC GAC CTG CGA CAC
 L K L M S A T Y **L**V A L C Q A I D L R H
 601 CTG GAG GAC CTA AAG CAC GCG GTG AAG GCG GCG GTG AGC CTC GTC GGC AAG CGC GTA
 L E E N L K H A V K A A V S **L**V G K R V
 661 CTC ACC ACT GGC GCC AAC GGA GAG CTC CAC CCC GCG CGT TTC TGC GAG AAG GAC CTC CTC
 L T T G A N G E L H P A R F C E K D L L
 721 ACC GTG GTC GAC CGC GAG CAC GTG CTG AGC TAC GCC GAC GAC CCC TGC AGC TCC GCC TAC
 T V V D R E H V L S Y A D D P C S S A Y
 781 GTG CTG ATG CCC AAG CTG CGG ATG GTA CTG GTG GAG CAC GCC CTC GGC CAC GGC GAC AAG
 V L M P K L R M V **L**V E H A L G H G D K
 841 GAG AAG GAC GCC GCG TCC ATC TTC CAC AAA GTC GCG GCT TTC GAG GAG GAG CTC AAG
 E K D A A A S I F H K V A A F E E E L K
 901 GCG GTG TTG CCG AAG GAG GTG GAA GCG GCG CGC GCC TTG GAG AGC GGC AAC CCG GCC
 A V L P K E V E A A R A A L E S G N P A
 961 ATC AGC ATC CGG ATC AAC GAG TGC AGG TCG TAC CCA CTG TAT CGG TTG GTG AGG GAG CAG
 I S I R I K E C R S Y P L Y R **L**V R E Q
 1021 CTG GGG GCG GCT TAC CTG ACC GGC GAG AAG GTG CGG TCG CCG GGG GAG GAG TTC GAG AAG
 L G A A Y L T G E K V R S P G E E F E K
 1081 GTG TCC GCA GCC ATC AAC GCG GGG CTG GTC ATC GAT CCT CTG CTT CAA TGC TTG AAC GAG
 V S A A I N A G **L**V I D P P L L Q C L N E
 1141 TGG AAT GGC GCT CCA ATT CCC ATC TGC TAA GTG TGA GAT TAT TAA TTA AAG CAA AAA AAA
 1201 W N G A P I P I C *
 GTG TGG AAT TGT GTC CAA TGA AAT CAG AAG CAA GTT ATT TCT TGT TTT TAT TTT CTT TTT
 1261 AAT CGA AAA TAT GTT CTT GTT AAA AAA AAA AAA

粗、斜体表示起始密码子 (ATG); 粗体加*表示终止密码子 (TAA); 脱氨基位点为方框加粗体; 催化激活位点为粗体加下划线; 左侧数字为核苷酸序号

Start codon is in bold and italics, stop codon is indicated by an asterisk and bold, and the deamination site are boxed and in bold, the catalytic active sites are underlined and in bold, numbers on the left margins represent nucleotide sequences

图4 CurPAL基因的核苷酸序列和推测的氨基酸序列

Fig. 4 Nucleotide and deduced amino acid sequences of CurPAL

约为33 000, 理论等电点为5.76, 总共包括4 728个原子, 分子式为 $C_{1484}H_{2376}N_{410}O_{445}S_{13}$ 。CurPAL的负电荷氨基酸残基(Asp+Glu)数为38, 正电荷氨基酸残基(Arg+Lys)数为31; 在组成CurPAL蛋白的20种氨基酸中, 丙氨酸所占的比例最高, 达到13%, 色氨酸所占的比例最低。CurPAL蛋白的不稳定指数为29.55, 脂肪指数为100.49, 根据Guruprasad方法表明CurPAL蛋白稳定; Grand average of hydropathicity (GRAVY)值为-0.022, 肽链表现为亲水性; TM-HMM预测结果显示该蛋白不是膜蛋白, 无跨膜结构域, 定位于细胞质中, 且N端不含有蛋白定位信号肽, 说明CurPAL是一个可溶性蛋白。

3.5 CurPAL蛋白二维和三维结构的预测

运用PORTER在线软件进行姜黄PAL蛋白二维结构的预测(图7), 由预测的结果来看, CurPAL蛋白二维结构组分中, α -螺旋(H)占65.91%>45%,

β -折叠(E)占0.65%<5%, 无规则卷曲(C)占34.44%, α -螺旋和无规则卷曲是其蛋白质二级结构中量最大的结构元件, 故CurPAL属于全 α 型蛋白。运用SWISS-MODEL在线软件进行CurPAL蛋白三维结构的预测(图8), 由预测的结果来看, 本实验中的姜黄PAL蛋白是基于荷兰芹[Petroselinum crispum (Mill.) Hill] PAL蛋白原子结构1w27.1的A链建模的, 两者序列一致性达到78.57%, CurPAL蛋白由同源四聚体构成, 与其他植物相同^[17]。表明本实验克隆得到的PAL基因是正确的。

4 讨论

苯丙氨酸解氨酶广泛存在于各种植物和少数微生物中。它是连接生物初级代谢和苯丙烷类代谢、催化苯丙烷类代谢第一步反应的酶, 是苯丙烷类代谢的关键酶和限速酶, 也是苯丙烷类代谢途径中研究最多的酶。它在类黄酮、木质素、香豆素、水杨酸、鞣质等次生代谢产物的合成过程中也起着很重

AFC372	1	
AAC33966.1	1	
NP_181241	1	MEI NGAHKSNGGGVDAMLCGGDI KTKNNVI NAEDPLNWGAAAEQMKGSHLDEVKRVAEFRPKVVNLGGETLTGQVAALSTI GNSVKVELSETARAGVN
AET41698.1	1	
AAN32867.1	1	
AEX32790.1	1	-MDATNCHGSNKVESFCVS--DPLNWGVAEFLKESHLDEVKRVAEYRKPVYRLGGETLTISQVAALAGREGDVGVLSATARAGVN
AFG2632	1	METTQETGNVLQSFCLKDPLNWEAELSKGSHEEVKRNVDEFRKPLI QLEGATLTISQVAALAR-DSDVVELADSARARK
ACG56648.1	1	-MEFAPKAQVVENGEAFLCKADPLNWI KAAESTGTSHLDEVKRNVVEFRKPVLVRLEGATLTISQVAAVAAARSPVRVELSEARDGVR
KJ780359	1	
consensus	1	
AFC372	1	
AAC33966.1	1	
NP_181241	101	ASSDWMSMNGKIDSYCVITGFGATSHIRITKNCVALQKELIRFLNAGLFGSTKETSHILPHSATRAAALMVRINLLQGSGIRFEELLEAITSPLANNIT
AET41698.1	1	
AAN32867.1	1	-NESMRKGTDTSYCLITGFGATSHIRITKCGGALQELIRFLNAGLFGNGRCSTHTLPHSATRAAALMVRINLLQGSGIRFEELLEAITSPLANNIT
AEX32790.1	86	ASSEWNEMSMSKGTDSYCVITGFGATSHIRITKCGGALQELIRFLNAGLFGNGRCSTHTLPHSATRAAALMVRINLLQGSGIRFEELLEAITSPLANNIT
AFG2632	84	ASSDWMSMNGKDTSYCVITGFGATSHIRITKCGGALQELIRFLNAGLFGNGRCSTHTLPHSATRAAALMVRINLLQGSGIRFEELLEAITSPLANNIT
ACG56648.1	88	ASSEWNEMSMSKGTDSYCVITGFGATSHIRITKCGGALQELIRFLNAGLFGNGRCSTHTLPHSATRAAALMVRINLLQGSGIRFEELLEAITSPLANNIT
KJ780359	1	
consensus	101	
AFC372	1	
AAC33966.1	8	PCLPLRGITIASGDLVPLSYIAQELITGRPNAKAVCPNGEKLNNEEAHRYAGVSGGFELQPKERGLALVNGTAVGSGMASVVLFTBSNLAVLSEVLSVIA
NP_181241	201	PCLPLRGITIASGDLVPLSYIAQELITGRPNAKAVCPNGEKLNNEEAHRYAGVSGGFELQPKERGLALVNGTAVGSGMASVVLFTBSNLAVLSEVLSVIA
AET41698.1	1	
AAN32867.1	95	PCLPLRGITIASGDLVPLSYIAQELITGRPNAKAVCPNGEKLNNEEAHRYAGVSGGFELQPKERGLALVNGTAVGSGMASVVLFTBSNLAVLSEVLSVIA
AEX32790.1	186	PCLPLRGITIASGDLVPLSYIAQELITGRPNAKAVCPNGEKLNNEEAHRYAGVSGGFELQPKERGLALVNGTAVGSGMASVVLFTBSNLAVLSEVLSVIA
AFG2632	184	PCLPLRGITIASGDLVPLSYIAQELITGRPNAKAVCPNGEKLNNEEAHRYAGVSGGFELQPKERGLALVNGTAVGSGMASVVLFTBSNLAVLSEVLSVIA
ACG56648.1	188	PCLPLRGITIASGDLVPLSYIAQELITGRPNAKAVCPNGEKLNNEEAHRYAGVSGGFELQPKERGLALVNGTAVGSGMASVVLFTBSNLAVLSEVLSVIA
KJ780359	1	
consensus	201	
AFC372	1	
AAC33966.1	108	EVMQGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
NP_181241	301	EVMQGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
AET41698.1	1	
AAN32867.1	195	EVMQGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
AEX32790.1	286	EVMQGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
AFG2632	284	EVMQGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
ACG56648.1	288	EVMQGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
KJ780359	1	
consensus	301	
AFC372	1	
AAC33966.1	208	WDNTRLAVAAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
NP_181241	401	EVMSGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
AET41698.1	1	
AAN32867.1	295	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
AEX32790.1	386	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
AFG2632	384	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
ACG56648.1	388	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
KJ780359	1	
consensus	401	
AFC372	1	
AAC33966.1	208	WDNTRLAVAAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
NP_181241	401	EVMSGKPEFTDHILTTLKHHPGQIAAAAEIHELDGSSYMKANCKLHNDPLQPKQDRYALRTSPQWLGPQEVI RAAATRMEREINSVNDNPLIDVS
AET41698.1	1	
AAN32867.1	295	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
AEX32790.1	386	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
AFG2632	384	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
ACG56648.1	388	NKALHGQNFQGTP1GVSNDNTRLAAI GKLNFQAQFSSELVNDLYNNGLPSNLTGGRNPLSDYFGKGAEI AVAS YCS ELQFLANPVTHHVQS AEQHNOVDVN
KJ780359	1	
consensus	401	
AFC372	84	SLGLISSRKTAEALDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
AAC33966.1	308	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
NP_181241	501	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
AET41698.1	93	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
AAN32867.1	395	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
AEX32790.1	486	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
AFG2632	484	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
ACG56648.1	488	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
KJ780359	84	SLGLISSRKTAEAVDI LKLNSSTYLVALCQAI DLRHLEENLRNNTVKNVSQVAKRTLITVGNGELHPSRFCEKEILLRVVDRHVFAYIDDPCCSATYPLVQ
consensus	501	
AFC372	184	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
AAC33966.1	408	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
NP_181241	601	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
AET41698.1	193	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
AAN32867.1	495	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
AEX32790.1	586	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
AFG2632	584	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
ACG56648.1	588	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
KJ780359	184	KLRQLVEHALKNGDAEKNSTSFQKTRAFEEELMLLIPKEVESARRVLESGNPCTPNRKECRSYPLYKIVREELGIGELTGEKALSPGEEDKVFIA
consensus	601	
AFC372	284	MCEGQVI DPPLICLKSWMGAPLP1C
AAC33966.1	508	MCEGQVI DPPLICLKSWMGAPLP1C
NP_181241	701	CEGKQVI DPPLICLKSWMGAPLP1C
AET41698.1	293	CEGKQVI DPPLICLKSWMGAPLP1C
AAN32867.1	595	CEGKQVI DPPLICLKSWMGAPLP1C
AEX32790.1	686	CEGKQVI DPPLICLKSWMGAPLP1C
AFG2632	684	CEGKQVI DPPLICLKSWMGAPLP1C
ACG56648.1	688	CEGKQVI DPPLICLKSWMGAPLP1C
KJ780359	284	CEGKQVI DPPLICLKSWMGAPLP1C
consensus	701	

AET 41698.1-中国李 NP_181241-拟南芥 AAN32867.1-中粒咖啡 AEX32790.1-葡萄 AFC372-浙江红山茶 AAC33966.1-辣椒

AFG2632-肉桂 ACG56648.1-小果野蕉 KJ780359-姜黄

AET 41698.1-P. salicina NP_181241-A. thaliana AAN32867.1-C. canephora AEX32790.1-V.vinifera AFC372-C. chekiangoleosa AAC33966.1-C. annuum AFG2632-C. osmophloeum ACG56648.1-M. acuminata KJ780359-C. longa

图 5 CurPAL 与其他植物 PAL 氨基酸序列多重比对

Fig. 5 Multiple alignments of amino acid sequence between CurPAL and PAL from other plants

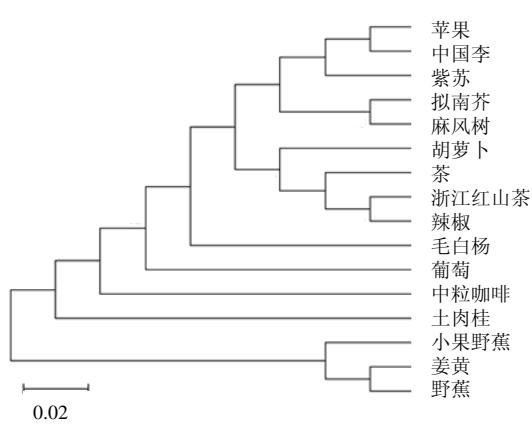


图6 CurPAL与其他植物PAL氨基酸序列的系统进化分析

Fig. 6 Phylogenetic relationship of amino acid sequences between CurPAL and PAL from other plants

```

MDNTRLAIAAGKLMFAQFSELVNDFYNNGLPSNLSGGRNLSLDYGFKAEIAMAAYCSE
CCHHHHHHHHHHHHHHHHHHHHHHHHHHHCCCCCCCCCHHHCCCCCHCCCCCHHHHHHHHHHHHHH
PQFLGNPTNHVQSSEQHNQDVNSLGLISARKTAEAVDILKLMSTATYLVALCQAIDLRLH
HHHHCCCCCCCCCECCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
EENLKHAVKAAVSLVGKRVLTGANGELHPARFCEKDLLTVDRHVLSSYADDPCSSAYV
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
LMPKLRMVLVHALGHGDKEKDAASIFHKVAAFEEELKAVLPKEVEAARALESGNPAI
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
SIRIKECRSYPLYRLVREQLGAAYLTGEKVRSPEFEKVSAAINAGLVIDPLLQLCNEW
CCHHHHCCCCHHHHHHHHHHHHCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
NGAPIPIC
CCCCCCCC

```

C-无规则卷曲 H- α -螺旋 E- β -折叠
C-coil H- α -helix E- β -extend

图7 CurPAL蛋白二维结构预测

Fig. 7 Forecast for 2D structure of CurPAL protein

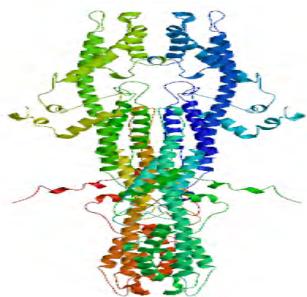


图8 CurPAL蛋白三维结构预测

Fig. 8 Forecast for 3D structure of CurPAL protein

要的作用。这些次生物质对植物生长发育、抗病害、防御紫外线等方面都有帮助。近年来，有关PAL基因的研究也成为研究热点。

本研究以姜黄的近缘物种的PAL基因序列设计简并引物，从姜黄成熟叶片中克隆了1个PAL基因

片段并结合RACE的方法获得了5'和3'端序列，首次从姜黄中获得了1个PAL基因全长的cDNA序列，命名为CurPAL基因。生物信息学分析表明，CurPAL基因推测的氨基酸序列与已报道植物的PAL基因氨基酸序列有很高的一致性（大于75%），基于推导氨基酸序列与其他已报道植物PAL基因氨基酸序列多重比对和系统进化树分析，CurPAL基因和小果野蕉聚为一支，CurPAL序列包含L180、V181、L230、A231等脱氨基氨基酸残基，其催化活性位点分别为N234、G235、N356、D357、N358、H370、HNQDV(460~464)，这些活性位点与水稻和玉米的相同，说明它们在结构和功能上有一定的相似性。可见，克隆得到的CurPAL基因是苯丙氨酸解氨酶基因。通过预测二维和三维结构表明CurPAL与荷兰芹PAL蛋白的结构一致性达到78.57%，由同源四聚体构成，不存在导肽和信号肽，无跨膜结构域，定位于细胞质中，肽链表现为亲水性； α -螺旋和无规则卷曲是其蛋白质二级结构中量最大的结构元件， β -转角和延伸链散布于整个蛋白质中。

姜黄素是通过苯丙烷代谢途径生成的，而苯丙氨酸解氨酶是苯丙烷代谢途径的一个关键酶和限速酶，其表达强度和酶活性将会影响姜黄素的合成。因此，本研究结果可为姜黄素类化合物代谢的分子调控机制奠定基础，为提高姜黄素含量以及姜黄遗传改良提供理论依据。

参考文献

- [1] 赵欣, 王爱里, 袁园, 等. 姜黄中姜黄素、去甲氧基姜黄素、双去甲氧基姜黄素的光稳定性分析 [J]. 中草药, 2013, 44(10): 1338-1341.
- [2] 刘建福, 范燕萍, 王明元, 等. 发光二级管不同光质对姜黄光合特性和姜黄素含量的影响 [J]. 生态学杂志, 2014, 33(3): 631-636.
- [3] 刘红艳, 王海燕, 叶松. 姜黄素药理作用及其机制研究进展 [J]. 中国现代医学杂志, 2012, 22(6): 48-51.
- [4] 王笑晴. 基于DPPH自由基清除能力的姜黄提取物抗氧化活性评价 [J]. 药物评价研究, 2011, 34(5): 360-363.
- [5] 罗廷顺, 李洪文, 刘正文, 等. 姜黄素的提取分离与药理作用研究进展 [J]. 现代药物与临床, 2011, 26(2): 102-107.
- [6] Aggarwal B B, Harikumar K B. Potential therapeutic effects of curcumin, the anti-inflammatory agent, against neurodegenerative, cardiovascular, pulmonary, metabolic,

- autoimmune and neoplastic diseases [J]. *Int J Biochem Cell Biol*, 2009, 41(1): 40-59.
- [7] Begum A N, Jones M R, Lim G P, et al. Curcumin structure-function, bioavailability, and efficacy in models of neuroinflammation and Alzheimer's Disease [J]. *J Pharmacol Exp Ther*, 2008, 32(1): 196.
- [8] 吴松权, 祖元刚, 管清杰, 等. 膜荚黄芪苯丙氨酸解氨酶基因的克隆与序列分析 [J]. 中草药, 2010, 41(3): 456-460.
- [9] Katsuyama Y, Kita T, Horinouchi S. Identification and characterization of multiple curcumin synthases from the herb *Curcuma longa* [J]. *FEBS Lett*, 2009, 583(17): 2799-2803.
- [10] del Carmen Ramirez-Ahumada M, Timmermann B N, Gang D R. Biosynthesis of curcuminoids and gingerols in turmeric (*Curcuma longa*) and ginger (*Zingiber officinale*): Identification of curcuminoid synthase and hydroxycinnamoyl-CoA thioesterases [J]. *Phytochemistry*, 2006, 67: 2017-2029.
- [11] Xu H, Park N, Li X H, et al. Molecular cloning and characterization of phenylalanine ammonia-lyase, cinnamate 4-hydroxylase and genes involved in flavone biosynthesis in *Scutellaria baicalensis* [J]. *Bioresour Technol*, 2010, 101: 9715-9722.
- [12] 孙梓健, 汤青林, 宋明, 等. 红叶芥低温胁迫下苯丙氨酸解氨酶活性及其基因的克隆表达 [J]. 西南大学学报: 自然科学版, 2010, 32(2): 90-94.
- [13] 江汉民, 王楠, 赵换, 等. 花椰菜苯丙氨酸解氨酶基因的克隆及黑腐病胁迫下的表达分析 [J]. 南开大学学报: 自然科学版, 2012, 45(4): 87-92.
- [14] 高杰, 谷巍, 周娟娟, 等. 黑三棱苯丙氨酸解氨酶基因克隆与序列分析 [J]. 中草药, 2014, 45(3): 403-409.
- [15] 徐莉莉, 童再康, 林二培, 等. 杉木苯丙氨酸解氨酶基因 CIPAL 的克隆与表达分析 [J]. 林业科学, 2013, 49(12): 64-72.
- [16] 王燕, 许峰, 杜何为, 等. 夹竹桃苯丙氨酸解氨酶的基因克隆与序列分析 [J]. 华北农学报, 2007, 22(4): 19-24.
- [17] 黄志伟, 郑亚凤, 许明, 等. 粳稻苯丙氨酸解氨酶基因的克隆与原核表达 [J]. 江苏农业学报, 2012, 28(5): 1194-1196.